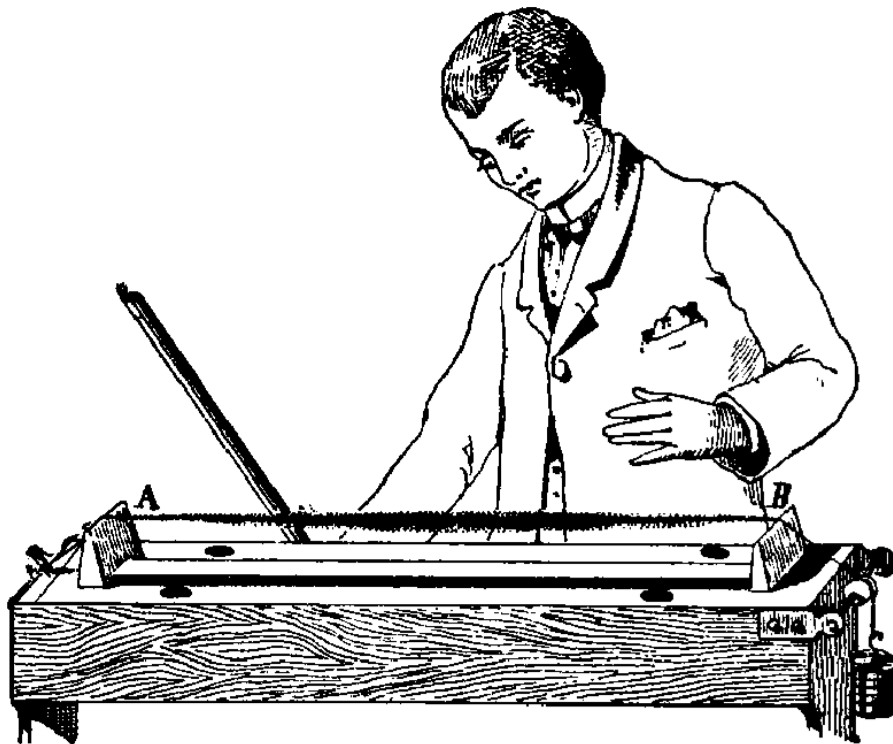




Proceedings of the Stockholm Music Acoustics Conference 2013



KTH Royal Institute of Technology
Sound and Music Computing Group

Proceedings of

SMAC 2013

4th Stockholm Music Acoustics Conference

July 30 – August 3, 2013

KTH Royal Institute of Technology

Stockholm, Sweden

Edited by Roberto Bresin & Anders Askenfelt

Cover by Kjetil Falkenberg Hansen

ISBN 978-91-7501-830-0

The website of this book is <http://smacsmc2013.smcsweden.se/>

Contents

Technical program committee viii

Reviewers ix

Preface xi

Violins 1

Playability of bowed-string instruments 3

Jim Woodhouse (invited)

On the effective material properties of violin plates 9

Evan Davis (invited)

Acoustic measurements in the workshop 16

George Stoppani (invited)

On perceptual evaluation of instruments: The case of the violin put into perspective 24

Claudia Fritz (invited)

Effect of holding the violin and soundpost removal on violin radiation via the dynamic filter model 32

George Bissinger

The acoustics of the Hardanger fiddle 38

Anders Buen

Acoustical features of complex bowing patterns in violin performance 46

Matthias Demoucron, Erwin Schoonderwaldt, Marc Leman

Analysis of bridge mobility of violins 54

Benjamin Elie, François Gautier, Bertrand David

String ‘after length’ and the cello tailpiece: Acoustics and perception 60

Eric Fouilhé, Anne Houssay

Vibrational modes of the violin family 66

Colin Gough

Acoustic characterization of violin family signature modes by internal cavity measurements 75

Colin Gough

Sensor Based Fatigue Recognition in Violin Playing 82

Tobias Großhauser

New devices for testing the stiffness characteristics of free violin plates 88

Ruiqing Jia, Ailin Zhang, Lei Fu

Enhanced simulation of the bowed cello string 94

Hossein Mansour, Jim Woodhouse, Gary P. Scavone

Digital modeling of bridge driving-point admittances from measurements on violin-family instruments 101

Esteban Maestre, Gary P. Scavone, Julius O. Smith

Evaluating violin quality: A comparison of player reliability in constrained vs unconstrained tasks 109

Charalampos Saitis, Gary P. Scavone, Claudia Fritz, Bruno L. Giordano

Acoustical constraints and individual preference in the coordination of complex bowing patterns 115

Erwin Schoonderwaldt, Matthias Demoucron, Eckart Altenmüller

Conceptualization of violin quality by experienced performers 123

Charalampos Saitis, Claudia Fritz, Catherine Guastavino, Gary P. Scavone

The influence of plate arching and thickness on the second and fifth mode on violin tops 129

Mats Tinnsten, Peter Carlsson

Probabilistic modeling of bowing gestures for gesture-based violin sound synthesis 133

Akshaya Thippur, Anders Askenfelt, Hedvig Kjellström

Violin quality evaluation: Examining the role of auditory and vibrotactile feedbacks 140

Indiana Wollman, Claudia Fritz, Jacques Poitevineau, Stephen McAdams

The influence of different driving conditions on the frequency response of bowed-string instruments 147

Ailin Zhang, Jim Woodhouse

Guitars and other plucked instruments 153

Analysis of the harpsichord plectrum-string interaction 155

Delphine Chadeaux, Jean-Loïc Le Carrou, Sandie Le Conte, Michèle Castellengo

Bio-inspired robot to study stringed instruments: Application to the harp 161

Delphine Chadeaux, Alexandre Roy, Jean-Loïc Le Carrou, Marie-Aude Vitrani, Benoît Fabre

Acoustic radiation of a split carbon composite soundboard guitar compared to wooden classical guitars 167

Sylvie Le Moyne, François Ollivier, Joël Frelat, Charles Besnainou

Vibroacoustic behavior of a vihuela 172

Sylvie Le Moyne, François Ollivier, Sandie Le Conte

Eigenspectrum of the sarode membrane 176

Manaswi Mishra, Gowtham Suresh, Ronojoy Adhikari

Ebony vs. rosewood: Experimental investigation about the influence of the fingerboard on the sound of a solid body electric guitar 182

Arthur Paté, Jean-Loïc Le Carrou, Benoît Fabre

High-speed camera displacement measurement (HCDM) technique of string vibrations 188

Niko Plath

Finite element model of a kantele with improved sound radiation 193

Henna Tabvanainen, Jyrki Pölkki, Henri Penttinen, Vesa Välimäki

The Bolivian Charango – An acoustic study 199

Owen Woods, Jim Woodhouse

Singing voice 203

Supranormal voices in singing 205

Ken-Ichi Sakakibara (invited)

Acoustic characteristics of vibrato in different singing styles 206

Noam Amir, Irit Ronen, Ofer Amir

Evaluation of pitch detection algorithms: Case of monophonic vocal performance 211

Robertas Budrys, Rytis Ambrazevičius

Formant tuning in Byzantine chant 217

Georgios Chrysochoidis, Georgios Kouroupetroglou, Dimitrios Delviniotis, Sergios Theodoridis

Glitch free FM vocal synthesis 224

Chris Chafe

Testing a new protocol to measure tuning response behaviour in solo voice ensemble singing 231

Helena Daffern, Jude Brereton

Assessment of the acoustical impact of piriform sinuses in MRI based vocal tract replicas 238

Bertrand Delvaux, David Howard

Vibrato analysis in Byzantine music 242

Dimitrios Delviniotis, Georgios Kouroupetroglou, Sergios Theodoridis

Parametrization of Byzantine chant ethos through acoustic analysis: From theory to praxis 250

Anastasia Georgaki, Achilleas Chaldaikis, Takis Tzevelekos

Interference-free observation of temporal and spectral features in ‘shout’ singing voices and their perceptual roles 256

Hideki Kawahara, Masanori Morise, Ken-Ichi Sakakibara

Formant frequencies of sung vowels intonated by six traditional Japanese Shigin singers. Part I: Dataset construction and analysis method 264

Masashi Nakayama, Kosuke Kato, Masaru Matsunaga

Formant frequencies of sung vowels intonated by six traditional Japanese Shigin singers. Part II: Results of analysis and statistical investigations 269

Kosuke Kato, Masashi Nakayama, Masaru Matsunaga

A pilot study of vibration pattern measurement for facial surface during singing by using scanning vibrometer 275

Tatsuya Kitamura, Hiroaki Hatano, Takeshi Saitou, Yui Shimokura, Eri Haneishi, Hiroko Kishimoto

Postural sway in vocal duets 279

Erik Koopmans, Caroline Palmer, Frances Spidle

Experimental study of the frequency leap interval produced by the change of laryngeal vibratory mechanism during sustained notes 280

Sylvain Lamesch, Boris Doval, Michèle Castellengo

An attempt to develop a singing synthesizer by collaborative creation 287

Masanori Morise

A study of speaker dependent formant space variations in karaoke singing 293

Mahnoosh Mehrabani, John H.L. Hansen

Temporal coordination in vocal duet performances of musical rounds 299

Caroline Palmer, Frances Spidle, Erik Koopmans, Peter Schubert

A method of division of soprano ranges and confirmation of their voice transformation point based on harmonics analysis 300

Ge Qu

Diverse resonance tuning strategies for women singers 306

John Smith, Joe Wolfe, Nathalie Henrich, Maëva Garnier

Power control for the second harmonic at high pitches in soprano singing: A case study 311

Hironori Takemoto, Seiji Adachi, Takeshi Saitou, Kiyoshi Honda, Eri Haneishi, Hiroko Kishimoto

Generating singing voice expression contours based on unit selection 315

Marti Umbert, Jordi Bonada, Merlijn Blaauw

Woodwinds – reeds and flutes 321

The player–wind instrument interaction 323

Joe Wolfe (invited), André Almeida, Jer-Ming Chen, David George, Noel Hanna, John Smith

Determination of 2D quasi incompressible flow around a recorder labium: A comparison between different methods 331

Roman Auvray, Pierre-Yves Lagrée, Benoît Fabre

Comparison of two methods of sound power measurements of flue organ pipes 338

Judit Angster, Katrin Hoge, Andras Miklos

Prediction of the dynamic oscillation threshold of a clarinet model: Comparison between analytical predictions and simulation results 344

Baptiste Bergeot, André Almeida, Christophe Vergez, Bruno Gazengel

An attempt at predicting the variation in playing frequencies for clarinets 350

Whitney Coyle, Jean Kergomard, Philippe Guillemain, Christophe Vergez, Alexis Guilloteau

On reeds and resonators: Possible explanations for cyclic spectral envelopes in the case of double reed instruments 358

Sandra Carral, Christoph Reuter

External pipe resonators and harmonica acoustics 365

James Cottingham, Casey Brock

Influence of pad ‘resonators’ on saxophone 371

Pauline Eveno, Marthe Curtit, Jean-Pierre Dalmont, René Caussé

A study of sound characteristics of a new bassoon as compared to modern German bassoon 378

Timo Grothe, Peter Wolf

Numerical modeling of a recorder in three dimensions 386

Nicholas Giordano

Investigation of bassoon directivity 391

Timo Grothe, Malte Kob

Evaluating a wavelet-based analysis of sensor reed signals for performance research 398

Alex Hofmann, Werner Goebel, Michael Weilguni

The voice of the mechanical dragon 403

Michel Hirschberg, Oleksii Rudenko, Gunes Nakiboglu, Ad Holten, Jan Willems, Avraham Hirschberg

Graph-based models for woodwinds 409

Georges Le Vey

Vocal tract effects on the timbre of the saxophone 415

Weicong Li, Jer-Ming Chen, John Smith, Joe Wolfe

Embracing the digital in instrument making: Towards a musician-tailored mouthpiece by 3D printing 419

Valerio Lorenzoni, Zjenja Doubrovski, Jouke Verlinden

Simulations of modal active control applied to the self-sustained oscillations of the clarinet 425

Thibaut Meurisse, Adrien Mamou-Mani, René Caussé, David Sharp

'In vivo' and 'in vitro' characterization of single cane reeds 432

Alberto Munoz, Bruno Gazengel, Jean Pierre Dalmont

A digital bagpipe chanter system to assist in one-to-one piping tuition 440

Duncan Menzies, Andrew McPherson

An experimental study of temperature variations inside a clarinet 446

Daniel Noreland

Study of the perceived quality of saxophone reeds by a panel of musicians 451

Jean-Francois Petiot, Pierric Kersaudy, Gary Scavone, Stephen McAdams

Numerical analysis of the mean flow effect on the sound directivity pattern of cylindrical ducts 458

Yong Shi, Andrey Da Silva, Gary Scavone

Is the jet-drive flute model able to produce modulated sounds like Flautas de Chinos? 465

Soizic Terrien, Christophe Vergez, Patricio de La Cuadra, Benoît Fabre

The design of a chromatic quena: How can linear acoustic help? 473

Camille Vauthrin, Benoit Fabre, Patricio de La Cuadra

Brass instruments 481

Lip motion, the playing frequency of the trombone and the upstream and downstream impedances 483

Henri Boutin, Neville Fletcher, John Smith, Joe Wolfe

Timpani-horn interactions at the player's lips 490

Jer-Ming Chen, John Smith, Joe Wolfe

The use of the input impedance for characterising historical serpents 495

Pauline Eveno, Sandie Le Conte

Trombone sound simulation under varying upstream coupling conditions 502

Vincent Fréour, Gary P. Scavone

Sensor based hand and lip weight and pressure measurements in trombone playing 509

Tobias Großhauser, Gerhard Tröster, Bernhard Hufnagl, Adina Mornell

Time domain simulation of standing waves in brass wind instruments taking non-linear wave steepening into account 514

Wilfried Kausel, Clemens Bernhard Geyer

Control of an artificial mouth playing a trombone and analysis of sound descriptors on experimental data 521

Nicolas Lopes, Thomas Hélie, René Caussé

Muscle activity in playing trumpet: The dependence on the playable pitch region and the experience of a non-trumpet brass instrument player 529

Shogo Matsukata, Hiroko Terasawa, Masaki Matsubara, Tetsuro Kitahara

Pitch bending techniques on early horns by manipulation of the embouchure: A comparison between measured and predicted data 534

Lisa Norman, Jonathan Kemp, John Chick, Murray Campbell

Percussions 541

The role of damping in steel pan manufacture 543

Claire Barlow, Soren Maloney, Jim Woodhouse

An objective approach for assessing the tuning properties of historical carillons 549

Vincent Debut, Miguel Carvalho, José Antunes

Nonlinear vibrations of steelpans: Analysis of mode coupling in view of modal sound synthesis 557

Mélodie Monteil, Cyril Touzé, Olivier Thomas

Time-resolved interferometry and phase vocoder analysis of a Caribbean steelpan 563

Andrew Morrison, Daniel Zietlow, Thomas Moore

Numerical experiments with non-linear double membrane drums 569

Alberto Torin, Stefan Bilbao

Experimental study of coupled drumhead vibrations using electronic speckle-pattern interferometry 577

Randy Worland

Physical modeling 583

Acoustics of pianos: Physical modeling, simulations and experiments 585

Antoine Chaigne (invited), Juliette Chabassier, Nicolas Burban

Large scale physical modeling sound synthesis 593

Stefan Bilbao, Brian Hamilton, Alberto Torin, Craig Webb, Paul Graham, Alan Gray, Kostas Kavoussanakis, James Perry

Simulated effects of combined control applied to an experimentally identified soundboard 601

Simon Benacchio, Baptiste Chomette, Adrien Mamou-Mani, François Ollivier

Sound synthesis of gongs obtained from nonlinear thin plates vibrations: Comparison between a modal approach and a finite difference scheme 607

Michele Ducceschi, Cyril Touzé, Stefan Bilbao

A structured approach to using a rectangular brace to design a soundboard section for a desired natural frequency 613

Patrick Dumond, Natalie Baddour

A new method for the identification of the original modes of damped three-dimensional axi-symmetric structures subjected to constraining boundary conditions 619

Vincent Debut, Miguel Carvalho, José Antunes

Modeling a vibrating string terminated against a bridge with arbitrary geometry 626

Dmitri Kartofelev, Anatoli Stulov, Heidi-Maria Lehtonen, Vesa Välimäki

Coupled modes and time-domain simulations of a twelve-string guitar with a movable bridge 633

Miguel Marques, José Antunes, Vincent Debut

Distributed piano soundboard modeling with common-pole parallel filters 641

Stefano Zambon

Computing virtual acoustics using the 3D finite difference time domain method and

Kepler architecture GPUs 648

Craig Webb

Music acoustics education 655

Music acoustics education at the Erich Thienhaus Institut in Detmold 657

Malte Kob

The Musical Acoustics Research Library (MARL): Fully digital & online 661

Gary Scavone, Jerry McBride

Activities for a course of physics of bowed instruments 663

Jesus A. Torres

Teaching physics via the web using music acoustics 668

Joe Wolfe, George Hatsidimitris, John Smith, John Tann

Author index 673

Technical program committee

Anna Rita Addessi, University of Bologna, Italy

Anders Askenfelt, KTH, Sweden

Federico Avanzini, University of Padova, Italy

Roberto Bresin, KTH, Sweden

Murray Campbell, University of Edinburgh, UK

Sofia Dahl, Aalborg University in Copenhagen, Denmark

Matthias Demoucron, Ghent University, Belgium

Gianpaolo Evangelista, Linköping University, Sweden

Emilia Gomez, UPF, Barcelona, Spain

Benoit Fabre, UPMC, France

Anders Friberg, KTH, Sweden

Leonardo Fuks, Rio de Janeiro Federal University, Brasil

Bruno L Giordano, University of Glasgow, UK

Werner Goebel, University of Music and Performing Arts Vienna, Austria

Kjetil Hansen, KTH, Sweden

Alexander Refsum Jensenius, University of Oslo, Norway

Malte Kob, Hochschule für Musik Detmold, Germany

Petri Laukka, Stockholm University, Sweden

Erwin Schoonderwaldt, Hanover University of Music, Drama and Media, Germany

Stefania Serafin, Aalborg University in Copenhagen, Denmark

Patrick Susini, IRCAM, France

Sten Ternström, KTH, Sweden

Vesa Välimäki, Aalto University, Finland

Reviewers

Adhitya, Sara	Chatron, Jacques	Guastavino, Catherine
Ahlbäck, Sven	Chen, Alex	Gómez, Daniel
Alexandre, Enrique	Chick, John	Hacihabiboglu, Huseyin
Allison, Jesse	Chon, Song Hui	Hattwick, Ian
Almeida, André	Château, Noël	Hellmer, Kahl
Anders, Torsten	Comajuncosas, Josep M	Herrera, Perfecto
Andrade Esquef, Paulo	Curtin, Joseph	Himberg, Tommi
Antonio	Cuthbert, Michael Scott	Hirschberg, Avraham
Antonacci, Fabio	Daffern, Helena	Hofmann, Alex
Antunes, José	Dalmont, Jean-Pierre	Houix, Olivier
Arar, Raphael	Dannenberg, Roger	Howard, David
Ashley, Richard	Daudet, Laurent	Hunt, Andy
Atienza, Ricardo	Davies, Matthew	Härmä, Aki
Aucouturier, Jean-Julien	de Francisco, Martha	Jackowski, Dariusz
Auvray, Roman	de Götzen, Amalia	Janer, Jordi
Avanzini, Federico	De La Cuadra, Patricio	Jansson, Erik
Baldan, Stefano	De Poli, Giovanni	Jensen, Kristoffer
Bank, Balázs	De Quay, Yago	Johnson, Rose
Barrass, Stephen	Debut, Vincent	Katz, Brian
Barthet, Mathieu	Delle Monache, Stefano	Kausel, Wilfried
Bello, Juan P.	Demoucron, Matthias	Kemp, Jonathan
Benetos, Emmanouil	Depalle, Philippe	Kereliuk, Corey
Beskow, Jonas	Ege, Kerem	Kergomard, Jean
Bevilacqua, Frederic	Elblaus, Ludvig	Kirke, Alexis
Bilbao, Stefan	Erkut, Cumhur	Klapuri, Anssi
Birkett, Stephen	Essl, Georg	Kleimola, Jari
Bishop, Laura	Eveno, Pauline	Klien, Volkmar
Bissinger, George	Farrùs, Mireia	Koehly, Rodolphe
Blauert, Jens	Fernstrom, Mikael	Kojs, Juraj
Bonada, Jordi	Flossmann, Sebastian	Kretz, Johannes
Bosch, Juan J.	Fontana, Federico	Kronland-Martinet, Richard
Bosi, Mathieu	Freour, Vincent	Lartillot, Olivier
Boutillon, Xavier	Friberg, Anders	Lavandier, Mathieu
Boyer, Eric	Fritz, Claudia	Le Carrou, Jean-Loic
Bozkurt, Baris	Germain, Francois	Lee, Edward Jangwon
Buen, Anders	Geronazzo, Michele	Lehtonen, Heidi-Maria
Buick, James	Gilbert, Joel	Lembke, Sven-Amin
Burger, Birgitta	Gingras, Bruno	Lindborg, Permagnus
Burgoyne, J. A.	Giordano, Nicholas	Lindström, Erik
Cambouropoulos, Emilios	Godøy, Rolf Inge	Lopez Arteaga, Ines
Canazza, Sergio	Gough, Colin	Lorenz-Kierakiewitz, Klaus-
Caramiaux, Baptiste	Gouyon, Fabien	Hendrik
Carlsson, Peter	Grani, Francesco	MacRitchie, Jennifer
Causse, René	Grassi, Massimo	Maculewicz, Justyna
Chaigne, Antoine	Großhauser, Tobias	Madison, Guy
Chandra, Arjun	Grothe, Timo	Malloch, Joseph

Mamou-Mani, Adrien	Saitis, Charalampos	Woolley, Alan
Mansour, Hossein	Savioja, Lauri	Yeo, Woon Seung
Marchand, Sylvain	Scavone, Gary	Zambon, Stefano
Mauch, Matthias	Scherrer, Bertrand	Zeiner-Henriksen, Hans
Mauro, Davide Andrea	Schmidt, Erik	Özbek, Erdal
McAdams, Steve	Schön, Daniele	
McLennan, John	Schwarz, Diemo	
Meunier, Sabine	Senturk, Sertan	
Misdariis, Nicolas	Serrà, Joan	
Moelants, Dirk	Sharp, David	
Moore, Thomas	Sinclair, Stephen	
Mores, Robert	Skogstad, Ståle A.	
Morreale, Fabio	Smith, Julius	
Myers, Arnold	Smyth, Tamara	
Naveda, Luiz	Sordo, Mohamed	
Nederveen, Kees	Spagnol, Simone	
Newton, Michael	Sprechmann, Pablo	
Nicol, Rozenn	Sujbert, László	
Nieto, Oriol	Sundberg, Johan	
Noisternig, Markus	Taillard, Pierre André	
Nymoen, Kristian	Temperley, David	
O'Modhrain, Sile	Thoret, Etienne	
Overholt, Daniel	Timmers, Renee	
Pabon, Peter	Toiviainen, Petri	
Papadopoulos, Helene	Torin, Alberto	
Papetti, Stefano	Torresen, Jim	
Parker, Julian	Touzé, Cyril	
Parncutt, Richard	Trail, Shawn	
Parseihian, Gaëtan	Trento, Stefano	
Pauletto, Sandra	Turchet, Luca	
Pearce, Marcus	Tørresen, Jim	
Penttinen, Henri	Umbert, Marti	
Percival, Graham	Upham, Finn	
Peters, Nils	Van Nort, Doug	
Petiot, Jean-Francois	van Vugt, Floris	
Pätynen, Jukka	Van Walstijn, Maarten	
Rabenstein, Rudolf	Vempala, Naresh	
Raphael, Christopher	Vergara, Sossio	
Rasamimanana, Nicolas	Vergez, Christophe	
Richardson, Bernard	Volpe, Gualtiero	
Rigaud, Francois	Wanderley, Marcelo	
Roads, Curtis	Weinreich, Gabriel	
Rocchesso, Davide	Widholm, Gregor	
Rodà, Antonio	Wingstedt, Johnny	
Roebel, Axel	Winters, Michael	
Rämö, Jussi	Wolfe, Joe	
Saari, Pasi	Woodhouse, Jim	

Preface

Welcome to the fourth Stockholm Music Acoustics Conference (SMAC 2013), July 30 – August 3, 2013. This time SMAC is run in parallel with the Sound and Music Computing Conference (SMC 2013). SMAC 2013 continues the series of music acoustics conferences in Stockholm, started 30 years ago. Following the tradition of SMAC 83, SMAC 93, and SMAC 03, SMAC 2013 covers the traditional fields of music acoustics, including musical instruments, singing voice, perception, and physical modeling. The fields of music perception and performance are this time covered by SMC 2013, as well as the rapidly growing research in sonic interaction design, sound processing and music information retrieval

A good reason for keeping the broad perspective from earlier SMACs is to offer the possibility of at least a partial overview of the many fascinating research areas which address the wonderful combination of performing arts, physics, creativity, and life experience called music. In order to give this perspective SMAC 2013 and SMC 2013 feature a number of invited presentations (the exact number being 13), in which outstanding researchers, old and young, will present overviews of their areas up to and including the research frontier.

SMAC 03 was run in two parallel sections. In the 2013 edition, one of the two sessions is the SMC conference. SMC is rapidly establishing as one of the most important conference series in the field of sound and music computing. This year SMC celebrates the 10th edition. SMC 2013 is jointly hosted by the Sound and Music Computing Research Group at the Royal Institute of Technology (KTH) and the Department of Composition, Conducting and Music Theory at the Royal College of Music (KMH) in Stockholm. KTH is responsible for the scientific part and KMH will host the music performances.

The theme for SMAC and SMC this year is “*Sound Science, Sound Experience.*” During the past five decades, the domain of music acoustics has widened from studies of the acoustics of musical instruments and voice, including basic elements of musical perception and performance, to investigations of how humans experience and interact with sounds and music. Increasingly, the knowledge is put into industrial, societal and psychological perspectives. The age-old dream of bridging science and art has found new and bountiful ground in the field of Sound and Music Computing.

Besides all scientific sessions SMAC and SMC 2013 will include many memorable events, including three concerts organized by KMH, Rencon Performance Rendering Contest, an electroacoustic pub, a Swedish summer night banquet in the archipelago of Stockholm, and above all, numerous occasions to meet friends and colleagues, old and new.

WE WELCOME YOU ALL TO SMAC 2013 & SMC 2013 !

Sten Ternström, Roberto Bresin, Anders Friberg, Anders Askenfelt

Violins

PLAYABILITY OF BOWED-STRING INSTRUMENTS

Jim Woodhouse

Department of Engineering
University of Cambridge
jw12@cam.ac.uk

ABSTRACT

Theory and measurements relating to the motion of a bowed string are surveyed critically, to see where they may shed light on judgments by players about the relative “ease of playing” of different violins. Particular attention is paid to what a player can do to achieve a range of tone colour, and how that perception of tonal range might vary between instruments. Some current lines of research are reviewed, and suggestions made for future physical and psychophysical investigation.

1. INTRODUCTION

What does a violinist mean when they say that one instrument is “easier to play” than another? In broad terms, they presumably mean “easier to get the particular sound I am trying to achieve”. Some aspects of being “easy to play” are no doubt associated with simple matters of set-up: string height and spacing, for example. Those are not of interest here: instead, we concentrate on questions involving the mechanics of bowing, the vibration of the bowed string, and the vibration and sound radiation of the violin body. For the past few decades, musical acousticians have been trying to pin down some aspects of the elusive idea of “playability”, so that a link could be made between acoustical measurements, simulation models, and the player’s perception of what is “easy to play”. In a pattern that is familiar throughout musical acoustics, a question which initially seemed quite simple has proved to be increasingly complicated the more we try to study it. This paper will review some aspects of what is known and discuss some currently active lines of enquiry.

2. WHAT IS “RANGE OF TONE”?

Players and instrument makers use a variety of terms when talking about playability issues. Some seem to be quite general, others refer to specific bowings. Some seem to refer mainly to quality of sound: terms like “good ring”, “projection” or “core”. Others seem to involve, at least in part, the response to bowing: “support”, “playing through/running out”, “range of tone”, “resistance”, “cushion”. To engage these players and makers in useful

dialogue, it is essential to make the attempt to understand what quantifiable meaning, if any, can be attributed to terms like these. I will mainly concentrate on “range of tone”, which seems quite a clear concept but which in fact raises a serious puzzle.

This puzzle stems from what might be called the paradox of Helmholtz motion. It is “well known” that the violin has very flexible and versatile sound: it is often said that of all instruments it comes closest to the flexibility of the human voice. However, it has been known since Helmholtz that virtually all normal notes on a violin are based on more or less the same string motion, the famous “Helmholtz motion” that produces a waveform of force driving the violin body which has a sawtooth form (see Figure 1a). So how can a player produce “variation of tone” if the string motion is always the same? Certainly there will be variations of sound as the body response changes significantly from note to note, governed by the frequency response function of the instrument, in turn determined by the frequencies, mode shapes and radiation efficiencies of the body modes. But the player has virtually no control over these factors once a given instrument has been chosen, and this really doesn’t seem to be the main thing players are talking about when “range of tone” is discussed.

So what can players do to vary “tone” by the details of the way they bow the strings? There are relatively few options.

- 1: Make subtle changes to the sawtooth form: the main mechanisms for this are to vary the sharpness/roundedness of the “Helmholtz corner” [1], and to modify the sawtooth wave via “Schelleng’s ripples” [2].
- 2: Switch to an entirely different regime of periodic string vibration: a bowed string is certainly capable of many kinds of motion in addition to the Helmholtz motion.
- 3: Do something that makes the Helmholtz motion non-periodic in one way or another.
- 4: Apply vibrato: this rather particular kind of modulation of a steady periodic note is of course much used by violinists, and perhaps the perceptual effects are best described in terms of modifying tone quality.
- 5: Influence, in some way, the perception of the quality of sound via differences in the major starting and ending transients of the notes, via the range of bowing gestures that players spend so long acquiring and practicing.

Do any or all of these options have potential to give rise to perceived differences in playability between different instruments? For clarity, we can imagine different violins strung with the same strings, set up with the same bridge top curve, string heights and so on, and played by the same player with the same bow, brand of rosin,

shoulder rest arrangement etc. We can take the options from the list in turn and ask whether each is likely to be influenced significantly by the acoustical behaviour of the body of the violin, and if so whether we know how to model the effect and whether we can plan psychoacoustical tests to explore it. We should keep in mind one persistent thread in the anecdotal evidence: it is often said that some instruments are challenging to play, but that they can reward a sufficiently expert player by offering a greater range of tone colours. The price of “tonal range” may be a harder instrument to control. This suggests conflicting aspects of “playability”, and indeed that there is no single scale along which instrument can be ranked, but that players of different abilities and different styles might well choose different instruments.

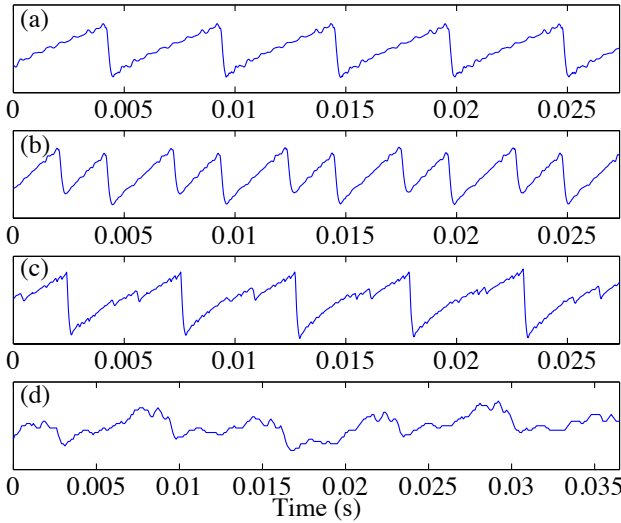


Figure 1. Typical bridge-force waveforms of bowed-string regimes: (a) Helmholtz motion, with Schelleng’s ripples clearly visible; (b) double-slip motion; (c) Helmholtz motion with irregular “spikes” caused by differential slipping; (d) “patchy” Helmholtz motion. (a-c) are a violin open G string, (d) is a cello open D string: time scales have been adjusted to show the same number of cycles.

3. SCHELLENG’S DIAGRAM

Schelleng made the first systematic attempt to say something about item 1 [2]. He examined the limits on the normal force between the bow and the string beyond which it is not possible to maintain Helmholtz motion. The range of bow force is important because the bow force determines the sharpness of the Helmholtz corner, and thus the high frequency content of the sound, which can be quantified for example by the spectral centroid [3]. At least one aspect of “range of tone” must surely involve how wide a range of bow force the player is able to apply, so we should ask how that range is influenced by the behaviour of the violin body (and also by the skill of the player).

Schelleng’s results were summarised in a famous diagram, sketched in Figure 2: for the moment, disregard the dashed line in the figure. As the contact point between bow and string moves closer to the bridge, both the min-

imum and the maximum force increase, but the minimum force increases faster and so the range of force available to the player narrows, and eventually pinches off at the apex of the triangular region. Schelleng’s diagram can be of great help to beginners on a bowed instrument, giving a clear image of a basic bow-control skill they are trying to acquire. It is less obvious whether the diagram has very much to say to expert players: does it tell us anything about the subtleties of playability? To address that we need to consider what physical events determine the maximum and minimum bow force lines.

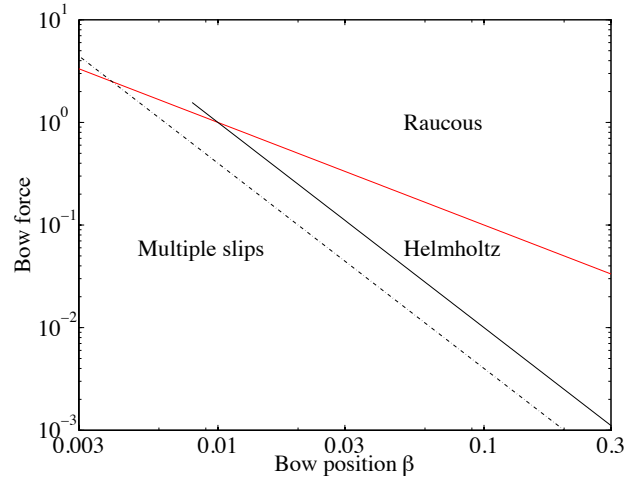


Figure 2. Schelleng’s diagram showing maximum (red line) and minimum (black line) bow force as a function of bow position expressed as a fraction β of the string’s length. Both axes are plotted on logarithmic scales. The dashed line indicates an alternative position for the minimum line: see text.

The minimum bow force is determined by the condition that the string is no longer able to stick to the bow-hairs throughout the period required by the Helmholtz motion. It gives way to motion in which the string slips twice or more per cycle, described by players as “surface sound” or “not getting into the string”. Experimental measurements confirm that this really is what happens to the Helmholtz motion at this limit, at least provided the bow is not too far away from the bridge [4,5]: see for example Figure 1b. Crucially, this limiting condition depends very strongly on the response of the violin body, removing energy from the vibrating string by allowing the string’s termination at the bridge to move as the body vibrates. The original analysis by Raman [6] and Schelleng [2] described this energy loss by idealising the body as a simple mechanical resistance or dashpot, but Schelleng’s argument can be extended to allow the note-by-note variation in predicted minimum bow force to be calculated from a measurement of the frequency response function at the violin bridge [7]: an example is shown in Figure 3. Work is currently under way to explore the extent to which the peaks revealed by such plots correlate with playability problems, but there is no doubt that some high peaks can have such implications since they mark the position of the “wolf note”.

For the maximum bow force, things are much less clear. Schelleng’s formula is based on the limiting bow force such that the arrival of a Helmholtz corner at the sticking

bow will not be sufficient to trigger a transition to sticking. This condition is governed by the properties of friction from the rosin on the bow hair and by properties of the string, but the violin body behaviour plays no role. On the face of it, then, maximum bow force would not be expected to vary between different violins provided they had the same strings and were played with the same bow and rosin.

Taken at face value, that means that we might expect different violins, or different notes on the same violin, to show the same maximum force line in Schelleng's diagram, but different minimum force lines: the dashed line in Figure 2 indicates an alternative minimum bow force. This does not change the maximum force at a given bow-string contact position, but it does affect the position of the apex of the triangle, and so it might affect the largest possible bow force by changing how close to the bridge it is possible to bow the string while maintaining Helmholtz motion. Is this a route for variation in "range of tone"? That depends on whether the apex of Schelleng's triangle really exists in the simple form suggested by the diagram. There have been some experimental determinations of Schelleng diagrams [4,5] using automated bowing machines, but unfortunately neither machine used so far was able to apply high enough force to probe the apex region directly: this could be a useful target for future research.

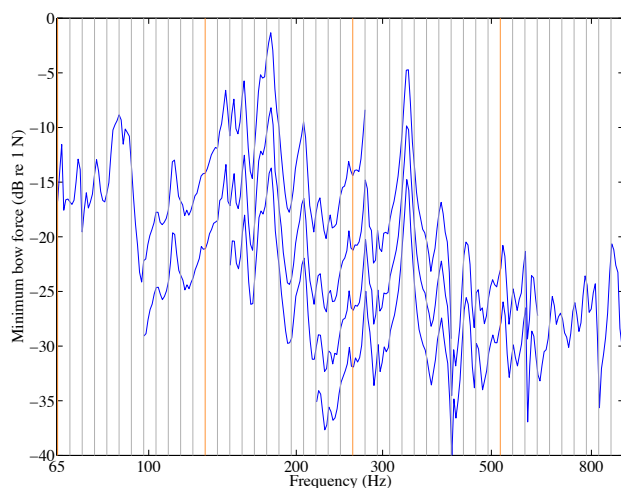


Figure 3. Minimum bow force as a function of played note frequency, deduced from the measured response of a cello for bow speed 0.1 m/s and $\beta = 0.1$. The four curves show values for the C, G, D and A strings, over a range of 2 octaves on each string starting from its open note. Vertical lines denote equal-tempered semitones; coloured lines mark the notes C.

But in any case, there is a snag with this interpretation of Schelleng's diagram. The maximum force criterion he used suggests a breakdown of Helmholtz motion to some kind of raucous "crunch", or perhaps to a so-called "anomalous low frequency" [8] in which a periodic note is produced at a pitch lower than the string's fundamental frequency. But there are other things that can go wrong at high bow force, and these need to be factored in. When the bow is a long way from the bridge, the Helmholtz motion begins to play systematically flat with high

bow force [9], and that may limit what force a player is able to use.

More relevant to the Schelleng apex, when the bow is close to the bridge it becomes essential to consider the finite width of the ribbon of bow hair in contact with the string. The player has some control over this width by tilting the bow, but it is always a few millimetres. When the bow is so close to the bridge that this width is a significant fraction of the bow-bridge distance, then the simple picture of stick-slip alternation breaks down. The string may continue to stick to some of the bow hairs, while slipping over others [10]. These "differential slips" may be regular, but usually they are rather irregular: an example is shown in Figure 1c. Audible noise can be heard accompanying the note, and what level of this noise might be acceptable depends on the musical context: it can often be heard in concerto performances but much more rarely in string quartets.

Some of the descriptions of playability differences between instruments by players and makers give a tantalising hint that they may be talking about variability in this noise associated with differential slipping. It is certainly sometimes observed, in measurements and in simulations, that the differential slipping events can be rather regular: what Pitteroff called a "miniature Helmholtz motion" can be set up between the bridge and the further edge of the bow [11]. It is an open research question what exactly determines the degree of regularity. Perhaps it is influenced by body behaviour or setup details around the bridge notches, and perhaps a sufficiently skilled player can learn to control it, and thus extend the range of bow force in the apex region?

4. OTHER WAYS TO INFLUENCE TONE

Returning to the list of ways that a player can influence range of tone, item 2 can be rather rapidly dealt with: transitions to non-Helmholtz regimes are not ordinarily used by violinists except under special circumstances. They explain the effects of *sul ponticello* and *sul tasto* playing, but normally if a violinist strays into another regime they are immediately told off by their teacher. It seems unlikely that these regimes contribute significantly to "range of tone" in the sense being discussed here.

Item 3 on the list seems a much more promising candidate. In the world of music synthesis, there is a widespread appreciation that very regular sounds don't make good music because they give an impression of being mechanical. Various tricks may be used in that context to introduce irregularities: some of these are to reflect the variations of gesture by a human performer, but in the context of real violin playing it would not be surprising if aspects of the instrument, or string, or bow, influence the degree of non-periodicity of string response in a way that has a perceptually important impact. We have already met one aspect of non-periodicity, in the irregular differential slips associated with finite bow width, but there are other possibilities.

The most familiar from the literature is "jitter", cycle-by-cycle variations of period. If real violin playing is analysed while "steady" notes are being played without vibrato, a spread of periods is found (see e.g. Cremer [1]).

For notes played in the player's "comfort zone", for example in the middle range of the Schelleng diagram, there seems to be no theoretical reason why a bowed string could not vibrate in an accurately periodic way, and the most careful measurements performed with this aim in view seem to confirm that [10].

The observed spread of periods during normal playing appears to be due to the reaction of the string to small external irregularities, from the player's gestural control or from spatial non-uniformity of the rosin. A bowed string is a nonlinear dynamical system that can show the kind of sensitive dependence on input conditions which has become familiar in other settings (for example the global weather, the "butterfly effect"). This tendency may well become greater near the boundaries of the usable regions in the player's parameter space, for example near the maximum bow force limit. Whether the tendency to sensitivity is influenced by the properties of the violin body is at present an open question, one that would probably best be approached via simulation studies: see section 5.

A phenomenon that may be related to this idea of sensitivity occurs in the lower right-hand corner of the Schelleng diagram, at low forces when bowing far from the bridge. Here, measurements suggest that the lower limit on Helmholtz motion is not caused by a transition to double slipping, but by a progressive loss of periodicity through what has been called "patchy" Helmholtz motion [4]. An example, obtained with a bowing machine with no human input of irregularity in the bowing gesture, is shown in Figure 1d. An underlying sawtooth wave is still discernable, but modified both by systematic effects and by details that vary from cycle to cycle.

It is entirely possible that there are other types of subtle departure from periodicity in bowed strings, which under some circumstances may be desirable: for example, one sometimes hears talk of instruments that allow the player to produce a "foggy tone" when desired. Exactly what this means is not obvious, but non-periodicity might surely be involved. Does the waveform of Figure 1d sound "foggy"? Careful psychoacoustic testing would be needed to address that question.

A different illustration is provided by recent experience trying to investigate item 4 on the list. Players sometimes talk of "lively" or "sparkling" instruments, and one theory to explain this involves sensitivity to vibrato. The frequency response of a violin has large peaks and dips, so that in response to vibrato the radiated sound from each harmonic of the note is modulated in amplitude as well as frequency, each one differently. The whole pattern changes from note to note. It is easy to imagine that some instruments might produce this effect more strongly than others, because of detailed differences of frequency response, and that this might be perceived by the player as a difference in responsiveness to vibrato.

However, recent efforts to test this idea psychoacoustically ran into difficulties [12]. It proved to be remarkably hard to design a test that gave the "right" answer, perhaps because the answer really is not right, but perhaps for a different reason. To give suitably controlled variations in vibrato and instrument response, it is necessary to use synthesized sounds. One might have thought that a short

note using a synthesized sawtooth wave, modulated by vibrato, as input to a filter with the required frequency response function, would sound quite like a violin note. However, most musically-trained listeners when asked a question like "Which of these two sounds is more lively?" responded by saying "Neither of them: they both sound horrible and mechanical". Efforts were made to incorporate transient details and irregularities of various kinds into the synthesis, but to no avail: the impression of an artificial, mechanical sound remained. This suggests that something important for the perception of naturalness is missing, and this must surely be something to do with irregularity in the sound, but the right kind of irregularity has yet to be identified. Pragmatic tricks may already have been developed to achieve natural-sounding synthesis in some computer-music contexts, but we are interested here in identifying the actual effects that matter in physical instruments, which might turn out to be quite a different matter and to have some importance for violin acoustics.

5. THE ROLE OF TRANSIENTS

Item 5 on the list takes us into different territory. The response of a string to a particular bowing gesture is a complicated matter, and players spend a long time mastering the control of these bowing transients. There is no doubt at all that the transient details of a played note can make a large difference to a listener's perception, and it has been shown that only a rather narrow range of transient responses is deemed acceptable [13]. Whether these transient details influence what is here described as "range of tone" is another matter, but it is certainly a question worth investigating.

One can envisage ways in which an influence might arise. There are two stages: the effect of body behaviour on transients, and the effect of transients on perception by the player, leading to differences in judged "playability", and specifically to a sense by the player of more or less "range of tone". The first issue can be explored by simulation and measurement. The first theoretical model of the bowed string, due to Raman, is nearly 100 years old now [6]. We have had computer simulation models capable of producing vaguely plausible transients for nearly 40 years [9]. Quite extensive studies relating to playability questions have been made using these models: see for example [14,15]. Unfortunately, the models all show sensitive dependence of transient details both on the assumed bowing gesture, and also on the assumptions about underlying physics incorporated into the model. The best current models show behaviour encouragingly close to that observed in measurements, and yet tantalizingly not close enough for full confidence [4,16]. There is without doubt further research waiting to be done in this area: an experimentally-validated model capable of predicting transients reliably could be used to explore many interesting questions. Some good methodologies have already been established, and these should be revisited with improved simulation models.

Experimental investigations are also useful. Figure 4 shows an example: a measured "Guettler diagram" [4,17] showing the length of pre-Helmholtz transient for a two-

dimensional family of bowing gestures. This plot gives some ideas about what kind of effects might be considered as possible sources of playability variations between different violins, associated with transient response to the bow. First, there are questions similar to those raised earlier about the Schelleng diagram. As Guettler predicted [17], the region of this plane where Helmholtz motion can be established quickly is confined to a wedge, and the influence of body parameters on the boundaries of that wedge can be studied.

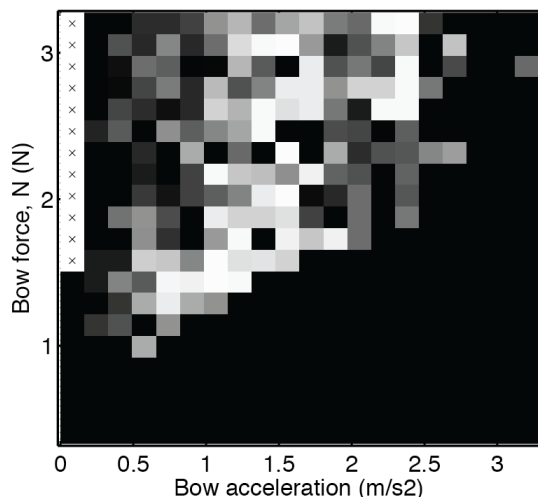


Figure 4. Measured “Guettler diagram” for the open D string of a cello, showing transient length before Helmholtz motion is established for a range of bow gestures with constant bow force and with the bow accelerated from rest with constant acceleration. Black pixels have not reached Helmholtz motion after 20 periods, white ones showed “perfect starts”, and the grey scale indicates the range in between.

But there is a second issue that may turn out to be more important. Within the wedge the pattern is “spotty”, and if the sequence of tests is repeated, the details of the spottiness change. This is a direct visualization of the sensitive dependence mentioned earlier: even with a bowing machine to apply nominally identical bow gestures, the exact details of transients can vary. It seems quite possible that the extent of this sensitivity might vary between different instruments, and that would surely be perceived by a player in terms of an instrument being more benign or more challenging to play. Such a variation might fit with the anecdotal evidence cited earlier: some instruments are said to be hard to control, but to reward a sufficiently skillful player with a wider tonal range.

To address this question seriously would be quite challenging: the first thought might be to use a combination of simulations with psychoacoustical testing. But listening tests are probably not enough: the issue here is the feedback to the player, as they try to control the bow to produce transients with a particular sound in the presence of the sensitive dependence. A test might require some kind of hybrid “hardware-in-the-loop” system, in which a

player held something feeling sufficiently like a violin, but which measured the bow actions and then fed them into a real-time simulation. The simulation could embody changes in properties of “string” and “body”, and be used to explore the player’s perceptions and preferences.

6. CONCLUDING DISCUSSION

Theory and measurements relating to bowed-string motion have been reviewed, to ask what light they might shed on questions of playability, with a particular emphasis on how one instrument might offer a player a larger range of tone colour. There have been more questions than answers, but at least an attempt has been made to organize the questions into sufficiently specific form that future research might be guided into productive directions.

One such thread of future work needs to involve simulation models, and the necessary first step is to enhance current models to the point where they can be convincingly demonstrated to give good predictions of transient details, and to contain all physical effects that seem likely play an important role. The more subtle the question to be addressed, the better the simulation model needs to be. Earlier workers such as Schelleng and Guettler managed to extract very useful information from relatively crude models, but to investigate questions like the propensity to sensitivity, or to subtle non-periodicity of response, might need rather a good model.

One factor that is known to require more work is the characterization of the frictional behaviour of rosin [16,18]. But there is also work currently going on to include additional physical features into simulations models: for example string motion in the plane perpendicular to the bow-hair, and dynamic motion of the bow hairs and stick. If an enhanced model can be developed, that should open the way to a new generation of parameter-plane studies, to assess the relative importance of the many model parameters.

Experimental work is also needed: further detailed observations of bowed string motion are likely to be useful, but exciting work has also been done recently into measuring what players actually do with their bows, and indeed with their whole bodies [19,20].

Finally there is the perceptual dimension of the problem. Psychoacoustical and psychophysical testing is easy to speculate about, but remarkably challenging to carry out in this subject area. One very characteristic problem is a conflict between the demands of statistical rigour (to do sufficiently many repeat trials, for example) and the requirement to retain at least a vestige of musical context. A typical listening test might involve spending an hour listening to rather similar sounds through headphones, being asked to make subtle judgments about musical preference. Even with the statutory breaks between test segments, is it realistic to expect the subjects to continue to make exactly the same judgments as they might make in a more normal musical context such as when rehearsing a piece in a quartet, or choosing between instruments to buy?

But all the most important questions surely fall in this perceptual area. How do we synthesise natural-sounding

violin music, even in tiny fragments? Can we pin down a reproducible meaning for such terms as “foggy tone quality”, or “good projection”, or “playing through” a string? And what DO players mean when they say that a particular favoured instrument offers a large range of tone quality? There is no shortage of challenging but potentially fruitful avenues to pursue before SMAC2023.

7. REFERENCES

- [1] L. Cremer, *The physics of the violin*, MIT Press, 1985.
- [2] J. C. Schelleng, “The bowed string and the player”, *J. Acoust. Soc. Amer.*, vol. 53, pp. 26–41, 1973.
- [3] E. Schoonderwaldt, “The violinist’s sound palette: Spectral centroid, pitch flattening and anomalous low frequencies,” *Acta Acustica united with Acustica*, vol. 95, pp. 901–914, 2009.
- [4] P. Galluzzo, “On the playability of stringed instruments”, Doctoral dissertation, University of Cambridge, 2004.
- [5] E. Schoonderwaldt, K. Guettler and A. Askenfelt, “An empirical investigation of bow-force limits in the Schelleng diagram,” *Acta Acustica united with Acustica*, vol. 94, pp. 604–622, 2008.
- [6] C. V. Raman, “On the mechanical theory of vibrations of bowed strings”, *Indian Assoc. Cult. Sci. Bulletin*, vol. 15, pp. 1–158, 1918.
- [7] J. Woodhouse, “On the playability of violins: Part 2”, *Acustica*, vol. 78:, pp. 137–153, 1993.
- [8] K. Guettler, “Wave analysis of a string bowed to anomalous low frequencies”, in *Journal of the Catgut Acoustical Society*, 1994, 2:8–14.
- [9] M.E. McIntyre and J. Woodhouse, “Fundamentals of bowed-string dynamics”, *Acustica*, vol. 43, pp. 93–108, 1979.
- [10] M.E. McIntyre, R.T. Schumacher and J. Woodhouse, “Aperiodicity in bowed-string motion”, *Acustica*, vol. 49, pp. 13–32, 1981.
- [11] R. Pitteroff and J. Woodhouse, “Mechanics of the contact area between a violin bow and a string, Part II”, *Acta Acustica united with Acustica*, vol. 84, pp. 744–757, 1998.
- [12] C. Fritz, J. Woodhouse, F. P. H. Cheng, I. Cross, A. F. Blackwell and B. C. J. Moore, “Perceptual studies of violin body damping and vibrato,” *J. Acoust. Soc. Amer.*, vol. 127, pp. 513–524, 2010.
- [13] K. Guettler and A. Askenfelt, “Acceptance limits for the duration of pre-Helmholtz transients in bowed-string attacks”, *J. Acoust. Soc. Amer.*, vol. 101, pp. 2903–2913, 1997.
- [14] R. T. Schumacher and J. Woodhouse, “The transient behaviour of models of bowed-string motion”, *Chaos*, vol. 5, pp. 509–523, 1995.
- [15] R. T. Schumacher and J. Woodhouse, “Computer modelling of violin playing”, *Contemporary Physics*, vol. 36, pp. 79–92, 1995.
- [16] J. Woodhouse and P. M. Galluzzo, “The bowed string as we know it today”, *Acustica*, vol. 90, pp. 579–589, 2004.
- [17] K. Guettler, “On the creation of the Helmholtz motion in bowed strings”, *Acustica*, vol. 88, pp. 970–985, 2002.
- [18] J. Woodhouse, “Bowed string simulation using a thermal friction model”, *Acta Acustica united with Acustica*, vol. 89, pp. 355–368, 2003.
- [19] E. Schoonderwaldt and M. Demoucron, “Extraction of bowing parameters from violin performance combining motion capture and sensors,” *J. Acoust. Soc. Amer.*, vol. 126, pp. 2695–2708, 2009.
- [20] E. Schoonderwaldt, “The violinist’s sound palette: Spectral centroid, pitch flattening and anomalous low frequencies,” *Acta Acustica united with Acustica*, vol. 95, pp. 901–914, 2009.

ON THE EFFECTIVE MATERIAL PROPERTIES OF VIOLIN PLATES

Evan B. Davis

Boeing Commercial Airplanes
EBDsince1955@gmail.com

ABSTRACT

The graceful arch and thickness variations of the violin family plates are replaced by a rectangular block of an effective material. The effective material properties are estimated from the violin's free plate modal frequencies. The dynamics of a rectangular block of the effective material will match the dynamics of the violin plate in the low frequencies.

The violin free plate mode shapes have the characteristic ring and X shapes of a dynamically-square plate. Dynamically-square plates balance the bending stiffness in two directions with the plate dimensions.

The Poisson's ratio of the effective material is determined by the ratio of the frequency of the ring mode to the X mode. The observed tuning ratio of 2.3 (M5:M2) yields a Poisson's ratio of one. Violin free plate data from old masters and modern makers cluster near the 2.3 tuning ratio.

A Poisson's ratio of one implies that the free plate has dynamic properties similar to a spherical cap. A spherical cap model can be used to gain insights into roles of arch-height and plate-thickness in plate tuning.

1. INTRODUCTION

Fifty years ago, John Shelleng published the benchmark paper "the Violin as a Circuit" in which he developed scaling rules for the violin family using flat plate theory [1]. In this paper, the graceful arch and thickness variations of the violin family plates are replaced by a rectangular block of an effective material. A rectangular plate built with the effective material will have the same dynamics as a violin plate at low frequencies. The dimensions of the effective plate will be the maximum length of the violin plate, the width of the upper bout and a thickness of the plate arch height.

The effective material properties are estimated using the modal frequencies of the free plates [2]. The effective properties are computed using the relationships asso-

ciated with dynamically square plates. Dynamically square plates will always have the characteristic X and ring mode shapes. The nodal patterns of the free violin plate and the corresponding dynamically-square rectangular plate are shown in figure 1.

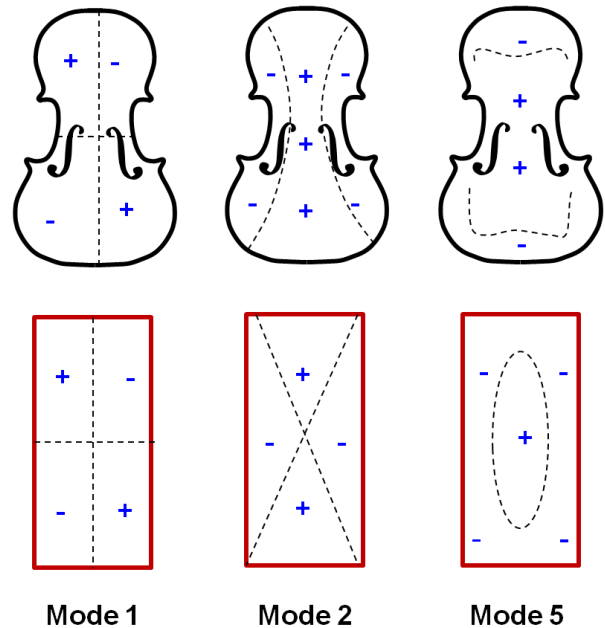


Figure 1. The mode shapes used in the estimation of the effective properties of the violin family plates. The frequencies associated with these mode shapes will be notated as M1, M2 and M5.

Any material can be made into a dynamically square plate by scaling the plate dimensions. Plates made of isotropic materials are dynamically square when the plate is square. Plates made of orthotropic materials can be made dynamically square by the proper selection of the plate's length-to-width ratio.

The average modal spacing will be used to order the data sets. The average modal spacing is the expected difference in frequency between adjacent modal frequencies in units of Hz/mode. The average modal spacing parameter is the reciprocal of the average modal density used in Statistical Energy Analysis. It will be shown that the average modal spacing can be computed directly from the free plate X (M2) and ring (M5) modal frequencies.

Selected ‘old master’ free plate data will be highlighted in the figures throughout the paper [3].

2. EFFECTIVE PLATE PROPERTIES

The use of effective material models, such as replacing ribs of a rib-stiffened plate with an effective material, has a long history in engineering design. The effective material models of rib-stiffened plates must be abandoned when the bending wavelengths approach the dimensions of the inter-rib-spacing. Unfortunately the violin plate does not have such a simple marker on when to abandon the effective plate representation. The general guidance would be to use the effective plate properties model for modes that have bending wavelengths on the order of the plate’s dimensions.

2.1 Selection of Dimensions

The ring and X mode pair exist only in rectangular plates that are dynamically square. The dynamically square plate has a fixed relationship between the bending stiffness and plate dimensions. The bending stiffness D_{length} is aligned with the plate’s major dimension L_{length} and is usually associated with the properties in the wood grain direction. The bending stiffness D_{width} is aligned with the width L_{width} or minor dimension of the plate and associated with the cross grain direction.

$$\frac{L_{length}}{L_{width}} = \sqrt[4]{\frac{D_{length}}{D_{width}}} \quad (1)$$

In practice, the ring and X modes exist in modified forms over a modest range of length ratios. Experimental results on nominally square isotropic plates show that the ring and X mode phenomena requires the aspect ratio to be within ten percent of square [4].

The selection of the dimension for a violin family plate is somewhat arbitrary but needs to be consistent in the study. The calculated bending stiffness parameters will be dependent on selected lengths. The plate’s maximum length, upper bout width and arch height are used to define the effective plate’s geometry. The product plate length and upper bout width is a reasonable estimation of projected plate area.

2.2 Calculation of Stiffness Parameters

The bending stiffness parameters can be computed from the modal frequencies $M1$, $M2$ and $M5$ of the twist, X and ring mode of the plate and the plate’s mass M_{plate} and area A_{plate} [2].

$$D_{length} = (M5^2 + M2^2) \frac{M_{plate} A_{plate}}{26} \left(\frac{L_{length}}{L_{width}} \right)^2 \quad (2)$$

$$D_{width} = (M5^2 + M2^2) \frac{M_{plate} A_{plate}}{26} \left(\frac{L_{width}}{L_{length}} \right)^2 \quad (3)$$

$$D_{couple} = 2.95 (M5^2 - M2^2) \frac{M_{plate} A_{plate}}{26} \quad (4)$$

$$D_{shear} = 1.78 (M1^2) \frac{M_{plate} A_{plate}}{26} \quad (5)$$

The plate’s effective properties and the selected geometry must be viewed as a matched set. The effective plate model can be used to compute deflections under loads as well as frequencies of the plate with various boundary conditions.

2.3 Estimation of Poisson’s Ratio

To complete the characterization of the effective material an effective Poisson’s ratio μ is needed. The frequencies of the ring (M5) and X (M2) modal pair alone provide the required information [2].

$$\mu = 1.48 \left[\frac{(M5/M2)^2 - 1}{(M5/M2)^2 + 1} \right] \quad (6)$$

The ring to X mode tuning ratio determines the Poisson’s ratio of the effective material.

2.4 Average Modal Spacing

The average modal spacing is computed from the plate’s area, mass and average bending stiffness.

$$\mathcal{F}_{\infty} = \frac{2}{A_{plate}} \sqrt{\frac{D_{length} D_{width}}{M_{plate} / A_{plate}}} \quad (7)$$

The average bending stiffness of an orthotropic plate can be taken to be the geometric average of the bending stiffness in the two orthogonal directions. The average bending stiffness can be expressed in terms of the free plate modal frequencies.

$$\sqrt{D_{length} D_{width}} = (M5^2 + M2^2) \frac{M_{plate} A_{plate}}{26} \quad (8)$$

Substituting the average bending stiffness expression into the average modal spacing equation leads to the important finding that the average modal spacing is a function of only the free plate X (M2) and ring (M5) modal frequencies.

$$\mathcal{F}_{\infty} = \sqrt{\frac{2}{13} (M5^2 + M2^2)} \quad (9)$$

The average modal spacing is independent of the characteristic dimensions of the plate.

The bending stiffness parameters can be computed from the plate's average modal spacing, mass and geometry.

$$D_{length} = M_{plate} A_{plate} \left(\frac{\mathcal{F}_{\infty}}{2} \frac{L_{width}}{L_{length}} \right)^2 \quad (10)$$

$$D_{width} = M_{plate} A_{plate} \left(\frac{\mathcal{F}_{\infty}}{2} \frac{L_{length}}{L_{width}} \right)^2 \quad (11)$$

The plate mass and the tap tone frequencies is a complete data set that allows for the comparison of plates in terms of the effective plate properties.

3. FREE TOP PLATE DATA

3.1 Free Plates Frequencies

The free plate data has been collected from literature and the violin makers attending the Oberlin Violin Acoustics workshop in the summer of 2012 [3, 5-7]. The bulk of the Oberlin data was contributed by makers Douglass Cox and Joseph Curtin. Jim Woodhouse supplied data from a set of six Rubio violins in his care. The majority of the data is violin data however a cello, a balsa violin and a few violas are included.

In the following figures, four old masters violin plates are highlighted with solid symbols. The “old master” plates, in order of increasing average modal spacing, are the Kreutzer Stradivari of 1727, the Booth Stradivari of 1716, the Allard Stradivari of 1728 and the Stretton Guarneri Del Gesu of 1726 [3].

There are a number of uncontrolled variables in the available data set 1) the environments in which the measurements were made are unknown with regard to temperature and humidity, 2) the test techniques and instrumentation varied by data collectors and 3) finished old master plates data has been combined with modern maker ‘in the white’ data. All the plates are in a “ready to assemble condition” with bass-bars and f-holes.

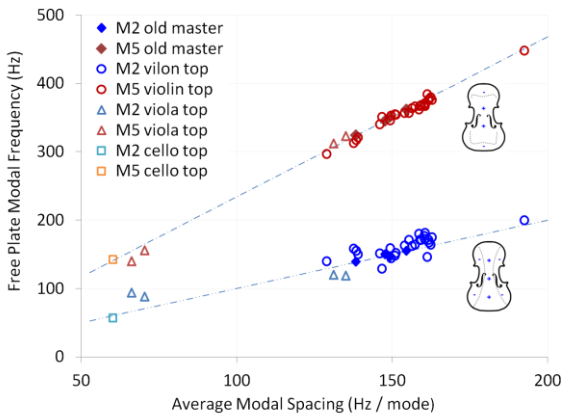


Figure 2. Free Plate Modal frequencies ordered by average modal spacing. The notation M2 refers to the frequency of the X-mode and M5 refers to the frequency of the ring-mode. The old master violin plates are the solid symbols.

High correlation between the average modal spacing and the free plate M5 frequencies can be anticipated from equation 9 and the large separation in frequency between the M2 and M5 modes. The fact that the M2 frequency is simply scaled to the average modal spacing is the more interesting result. It is also interesting that the average modal spacing parameter orders the data across all members of the violin family.

The free plate modal frequencies fall along straight lines in figure 2. The straight lines are due to the fact that the M5:M2 ratio is almost constant for the free plates of the violin family. Data from eight old master violin top plates has been plotted on an expanded scale in figure 3.

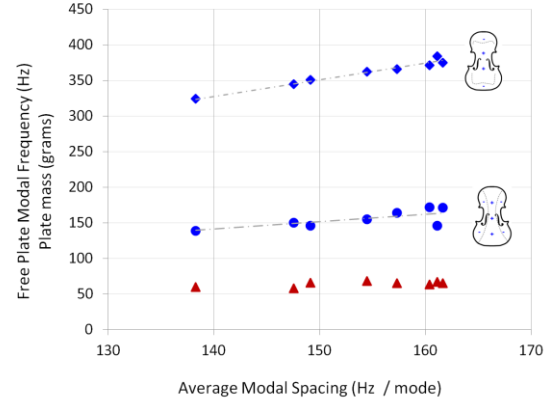


Figure 3. Old master free plate modal frequencies and plate masses in grams (triangles) ordered by average modal spacing [3].

Four of these old master plates have been highlighted in the previous figures. In order of increasing average modal spacing are the Kreutzer Stradivari of 1727, the Booth Stradivari of 1716, the Allard Stradivari of 1728, and the Stretton Guarneri Del Gesu of 1726, a Carlo Testore, a Carlo Landolfi of 1762, Carlo Tononi ca. 1730 and Francesco Rugeri of 1685[3].

The effective Poisson's ratio of the plates is determined by the tuning ratio between the ring (M5) and X (M2) modes.

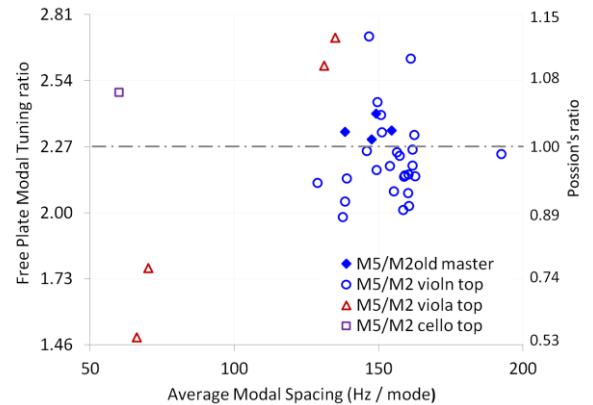


Figure 4. Free plate ring (M5) to X (M2) mode tuning ratio as a function of modal spacing with the corresponding effective Poisson's ratio. The old master violin plates are the solid symbols.

The measured violin plates have a nominal effective Poisson's ratio of one. An effective Poisson's ratio of one suggests a hemispherical shell-like structure. The average effective Poisson ratio of the old master plates was 1.01 which corresponds to a tuning ratio of 2.32. The cluster of data shows that there are a number of modern plates that match the tuning ratio of the old master plates.

The viola plates showed a much wider range in the tuning ratio than the violins. With a sample of one, not much can be said with confidence about the cello.

The tuning ratio objective of an octave is attributed to Carleen Hutchins [7-9]. The octave tuning ratio was based on the measurement of a Stradivari top with the caution "Undue emphasis should not be placed on only one test of this kind..." [8]. There are plates in the data set that are tuned by the octave rule, this may or may not be due to the influence of Carleen Hutchins [9,10]. The six top plates in the Rubio violins were octave tuned. The Rubio violins were made over the period 1991-1992.

The old master and modern plate data support the view that an octave tuning rule is a first order estimate. A refinement to a M5:M2 = 2.3 (octave plus 2 1/4 semitones) tuning rule is a result of more old master free plate data.

The effective material plate theory suggests that the design objective is a Poisson's ratio of one. In theory, a Poisson's ratio of one corresponds to a tuning ratio M5:M2=2.27. The design objective of a Poisson's ratio of one is well within standard violin making practices.

3.2 Drive Point Impedance and Plate Mass

Plate mass is required in the evaluation of impedance. Unfortunately, not all the free plate modal frequency data sets came with plate mass data. The data used in this section is a sub-set of the data used in the previous section.

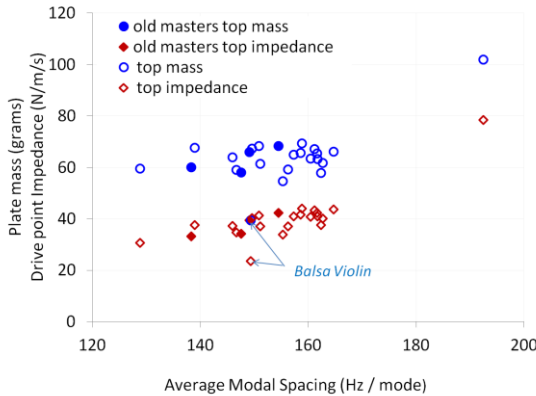


Figure 5. Violin plate masses and characteristic impedances ordered by average modal spacing.

The characteristic drive-point impedance Z_{∞} is the impedance of a system averaged over all possible drive points. The real part of the characteristic drive-point impedance is a function of the average modal spacing and the plate's mass M_{plate} .

$$\Re(Z_{\infty}) = 4M_{plate}\mathcal{F}_{\infty} \quad (12)$$

The masses of the old master plates ranged from 58 to 68 grams. The heaviest modern spruce violin plate was 102 grams. The lightest violin plate was a balsa violin at 39.4 grams. There is some correlation between plate mass and average modal spacing.

3.3 Summary of the Free Top Plate Data

The design objective appears to be an effective Poisson's ratio of one. The tuning ratio between the ring (M5) and X mode (M2) required to achieve this objective in theory would be 2.27 and is 2.32 by old master practice. To two significant figures, the M5:M2 tuning ratio is 2.3 in either case. The 2.3 tuning ratio should be viewed as a revision to the Hutchins octave rule, based on further evidence, not as a repudiation of it.

4. FREE BACK PLATE DATA

4.1 Back Plate Data

Free back plate data has been made available from modern makers. No data from old master back plates was available to the author.

The free back plate data has the same basic data trends as top plate data (figure 2) and cluster at very similar frequencies, but with more variation. Three of the back plates have closely spaced ring (M5) and X (M2) modes.

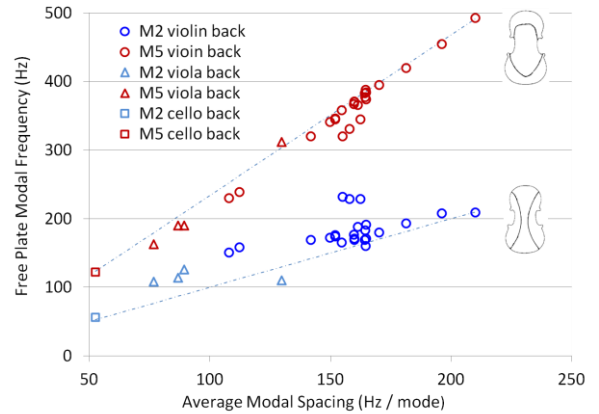


Figure 6. Back plate data from modern violin family instruments ordered by average modal spacing.

To have the maple back plate and spruce top plate act so similarly requires a sophisticated design. The top is traditionally made with book-matched quarter-sawn spruce and has a bass-bar and f-holes. The maple back plate is a solid plate with a thick center. To further complicate the design, quarter-sawn as well as slab-cut maple backs have been used. Making the back plate thicker in the C-bout area appears to be part of a design solution that compensates for the differences between maple and spruce.

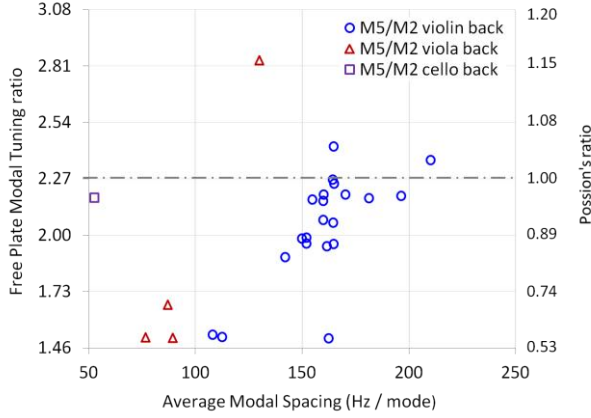


Figure 7. Back plate data from modern violin family instruments ordered by average modal spacing.

The back-plate design objective appears to be a Poisson's ratio of one, like the top-plate. The required M5:M2 tuning ratio of 2.3 is a good tuning strategy for the back plate and is achievable. There are six octave tuned Rubio back plates in this data set.

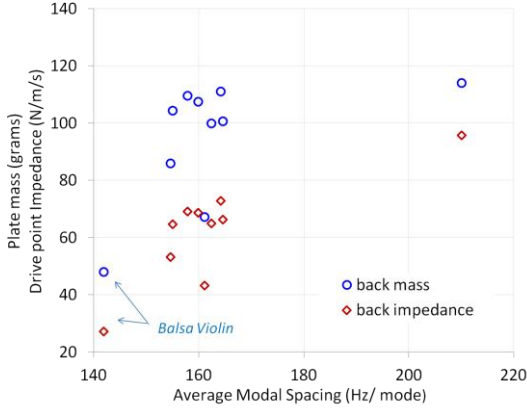


Figure 8. Violin free back plate mass and real part of the characteristic impedance.

The characteristic impedance scales with the plate mass. The higher mass of the backs relative to the tops gives the back plates higher drive point impedances.

The ratio of the mass of the back plate to the mass of the top plates ranged from approximately unity to two. Only the very light tops were at half the back plate weights. Most of the top to back pairs had a 2:3 top to back mass ratio. The ratio of the density of spruce to maple is approximately 2:3, suggesting the volume of wood in both plates is very similar in most cases.

5. ANALYSIS USING THE EFFECTIVE PLATE PROPERTIES

The propose for developing the effective plate properties was look at 1) how the free plate would react to different boundary conditions and 2) if sandwich plate theory could give some insights to top and back plate matching strategies.

5.1 Clamping the Edge of the Free Plate

The effective properties of a free plate can be used to compute frequencies of a mounted plate. The natural frequencies of plates mounted with various boundary conditions can be written as a function of the average modal spacing.

$$\left(\frac{f_{n,m}}{\mathcal{F}_\infty}\right)^2 = \frac{\pi^2}{16} \left\{ G_{length,n}^4 \left(\frac{L_{width}}{L_{length}}\right)^2 \sqrt{\frac{D_{length}}{D_{width}}} + G_{width,m}^4 \left(\frac{L_{length}}{L_{width}}\right)^2 \sqrt{\frac{D_{width}}{D_{length}}} \right. \\ \left. + \left(H_{length,n} H_{width,m}\right) \frac{2D_{couple}}{\sqrt{D_{length} D_{width}}} \right. \\ \left. + \left(J_{length,n} J_{width,m} - H_{length,n} H_{width,m}\right) \frac{4D_{shear}}{\sqrt{D_{length} D_{width}}} \right\} \quad (13)$$

The boundary conditions of the parallel edges of the plates determine the parameters $G_{length,n}$, $H_{length,n}$, $J_{length,n}$ and $G_{width,m}$, $H_{width,m}$, $J_{width,m}$ [11, see table 11-5].

Limiting the boundary conditions to simply supported or clamped and employing the dynamically square relationships simplifies the frequency equation.

$$\left(\frac{f_{n,m}}{\mathcal{F}_\infty}\right)^2 = \frac{\pi^2}{16} \left\{ G_{length,n}^4 + G_{width,m}^4 + \mu \left(H_{length,n} H_{width,m} \right) \right\} \quad (14)$$

The effective material model should be restricted to the study of plate modes with only a few nodal lines. The fundamental plate mode is of primary interest and the effective material modeling approach is well suited for that study. In the limit of high frequencies, it will be the actual wood properties and thickness profile that govern the dynamics of the carved plate.

A clamped plate is a more efficient structural acoustic design than a simply supported plate for fixed dimensions in a resonant frequency constrained design [12, 13]. The violin appears to have the expected clamped edge boundary condition [14].

$$MO_{clamped} = 1.17 M5 \sqrt{1 + 0.259 \left(\frac{M2}{M5}\right)^2} \quad (15)$$

The clamped plate resonance frequency (MO) is in the frequency range of the M5 or ring mode frequency. Using the suggested M5:M2=2.3 tuning ratio, the clamped plate natural frequency estimate can be written in terms of the M5 modal frequency or average modal spacing.

$$MO_{clamped} \approx 1.19 M5 \approx 2.80 \mathcal{F}_\infty \quad (16)$$

The assembled violin signature modes have a mix of “body bending” and “body breathing” motions. The clamped plate model is a good candidate as the source of the “body breathing” motion in the signature modes.

5.2 Coincidence Frequency

The coincidence frequency is the frequency at which the speed of the bending waves in the plate travel at the speed of sound c_0 .

$$f_{\text{coincidence}} = \frac{c_0^2}{\pi A_{\text{plate}} \delta f_{\infty}} \quad (17)$$

Above the coincidence frequency, the plate has reached its maximum radiation efficiency. Below the critical frequency, the radiation efficiency is proportional to the driving frequency divided by the critical frequency. For top and back violin plates the coincidence frequency is approximately 4200 Hz.

5.3 Sandwich Plate Body Model

The sandwich plate model is an idealization in which the top and back plates are restricted to move together as if they were attached by a large number of mass-less sound posts. The sound posts fix the separation distance between the two plates.

The sandwich plate's bending stiffness is determined by the effective plate properties and the distance separating the plates. The separation distance between the plates has been taken to be the height of the ribs h_{rib} plus one half of the arch height of each plate $h_{\text{top-arch}}$ and $h_{\text{back-arch}}$.

$$D_{\text{box}} \cong D_{\text{top}} \left[1 + \frac{D_{\text{back}}}{D_{\text{top}}} + \frac{\left[\frac{1}{2} \left(1 + \frac{h_{\text{back-arch}}}{h_{\text{top-arch}}} \right) + \frac{h_{\text{rib}}}{h_{\text{top-arch}}} \right]^2}{1 + \frac{D_{\text{top}}}{D_{\text{back}}} \left(\frac{h_{\text{back-arch}}}{h_{\text{top-arch}}} \right)^2} \right] \quad (18)$$

The violin family rib-height to top plate arch-height ratio varies by family member. Expressing the top arch to rib height in mm's the typical violin has a ratio of approximately 15:30, the viola 17:40 and the cello 27:116.

The sandwich panel ring and X modes can be tuned for a given set of plates with the rib height.

$$M2_{\text{box}} \cong \sqrt{\frac{13 \sqrt{D_{\text{length-box}} D_{\text{width-box}}} - 4.4 D_{\text{couple-box}}}{(M_{\text{top}} + M_{\text{back}}) A_{\text{plate}}}} \quad (19)$$

$$M5_{\text{box}} \cong \sqrt{\frac{13 \sqrt{D_{\text{length-box}} D_{\text{width-box}}} + 4.4 D_{\text{couple-box}}}{(M_{\text{top}} + M_{\text{back}}) A_{\text{plate}}}} \quad (20)$$

In the special case of matched tuned tops and backs, the sandwich box modes can be expressed in terms of the free plate modes. Match tuning the top and back plates is a common, but not universal, design strategy.

$$M2_{\text{box}} \cong M2_{\text{top}} \sqrt{1 + \left(\frac{1 + h_{\text{rib}}/h_{\text{arch-top}}}{1 + M_{\text{top}}/M_{\text{back}}} \right)^2} \quad (21)$$

$$M5_{\text{box}} \cong M5_{\text{top}} \sqrt{1 + \left(\frac{1 + h_{\text{rib}}/h_{\text{arch-top}}}{1 + M_{\text{top}}/M_{\text{back}}} \right)^2} \quad (22)$$

Assuming a back-to-top mass ratio of 3:2 and using typical rib-height to arch-height ratios leads to the conclusion that the M2 mode of the sandwich plate is similar to the clamped plate frequency MO.

5.4 Combining Component Modes

The component modes only exist as mathematical idealization. The component modes are clamped plate modes of the tops and backs and X mode of the top-back sandwich. The sandwich body mode can be viewed as a global mode. The clamped plate modes can be viewed as local modes. The local and global mode shapes combine to generate system mode shapes. The relative phases of the global and local modes lead to mode shapes that resemble the body bending B0 mode and a precursor B1 (T1) mode. The B1 precursor mode has been observed in violins without a sound post [14].

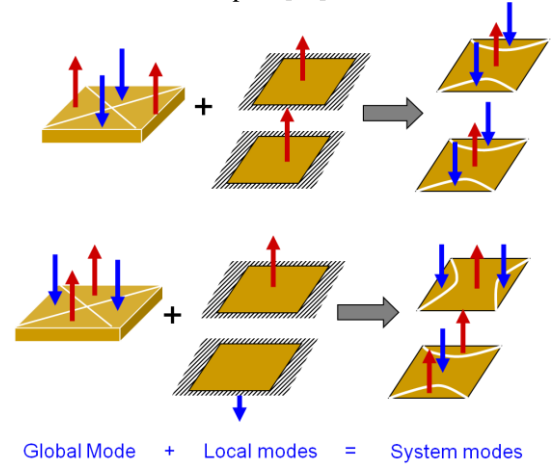


Figure 9. A schematic sketch of how the component modes combine to create system modes associated with the violin B0 and B1 modes.

The sketch of modal motion is as far as one can take this analysis without data on the ribs. The ideal sandwich plate model does not address the in-plane motions of the box [15]. A rib mode is needed for that.

5.5 On Plate Tuning

The effective-properties flat plate model treats the carved violin plate as a rectangular block of an effective material. The effective properties give no guidance on how or where to carve the plates to adjust or tune them. The effective Poisson's ratio of the effective material suggests it is acting as if it were a spherical cap. Using the dynamics of a spherical cap some intuition on the roles of the plate's arch and thickness can be developed.

The following discussion is based on a model for a circular plate with a shallow spherical arch. The radius of curvature of the arch is R_{arch} . The circular plate has a diameter d , thickness t , modulus E and density ρ . The frequencies of the arched plate are a function of the flat

plate frequencies f_{flat} and a term associated with the radius of curvature of the arch.

$$f_{shell}^2 = f_{flat}^2 + f_{arch}^2$$

$$f_{flat}^2 = \frac{1}{4\pi^2} \frac{\lambda_{i,j}^4}{d^2} \frac{16}{12} \left(\frac{h}{d}\right)^2 \frac{E}{\rho(1-\mu^2)} \quad (23)$$

$$f_{arch}^2 = \frac{1}{4\pi^2} \frac{1}{R_{arch}^2} \frac{E}{\rho}$$

Expanding the above expressions and stating the radius of curvature in terms of the arch height h_{arch} allows the examination of the roles of arch design and thickness.

$$f_{shell} \cong f_{arch} \sqrt{1 + \frac{\lambda_{i,j}^4}{12} \left(\frac{t}{h_{arch}}\right)^2 \left[1 + 4\left(\frac{h_{arch}}{d}\right)^2\right]^2} \quad (24)$$

The modal frequency parameter lambda squared $\lambda_{i,j}^2$ is tabulated in many reference books and depends on the boundary conditions and mode shapes [10]. Lambda squared for a free circular plate is 9.084 and is a ring mode. Lambda squared for a clamped circular plate is 10.22 and has a node at the edge of the plate.

The plate is primarily tuned by the arch frequency and adjusted with the thickness-to-arch-height ratio. John Shelleng also discussed a plate-thickness to plate-arch-height scaling rule [1, footnote 26].

Generalizing to the violin plates, the arch frequency would be a function of the arch profile and body shape. Plate thinning has a diminished effect as the plate approaches the arch frequency.

6. CONCLUSIONS

The interpretation of free plate data in the form of a set of effective properties provides insight into the design of the violin family of instruments. The breathing components of the violin's signature modes are a result of the clamped plate modes. The body bending modes arise from the X-mode of a top-plate, back-plate sandwich. The modal frequency of these ideal modes can be computed from the effective plate properties. The free plate data, by itself, does not have enough information to completely define the dynamics of the assembled violin.

The free plate M5:M2=2.3 tuning rule is based on data from old master plates and a theoretical model. The effective plate material theoretical model has been used to argue that the observed M5:M2 tuning is the result of a very efficient structure. The fact that generations of creative violin makers have not significantly changed the design of the violin supports the assertion that the violin shape and its materials are an optimized system.

Acknowledgments

Thanks to the violin makers at Oberlin Violin Acoustics workshop who made this project possible. These skilled makers generously shared their data and experiences after asking the hard questions.

7. REFERENCES

- [1] John Shelleng, "The Violin as a Circuit", *J. Acoust. Soc. Am.* 35(3) pp. 326-338, 1963
- [2] M. E. McIntyre and J. Woodhouse, "on measuring the Elastic and Damping constants of Orthotropic sheet Materials", *Acta metall.* Vol. 36, No. 6 pp 1397-1416, 1988
- [3] Joseph Curtin, "Tap Routine", *The Strad magazine*, 48-54, October 2006
- [4] G. B. Warburton, "The Vibration of Rectangular Plates", *Proc. Inst. Mech. Eng.*, ser A, vol. 168, no. 12, pp. 371-385, 1954
- [5] N. E. Molin, L-E. Lindgren and Erik V. Jansson, "Parameters of violin plates and their influences on the plate modes", *J. Acoust. Soc. Am.* 83(1) 281-291, 1988
- [6] Chris Waltham, "A Balsa Violin", *Am. J. Phys.* 77(1), pp 30-35, 2009
- [7] Robert A. Wilkins, "Cello Modes Throughout Three Stages of Construction", *CASJ Vol 4*, (series II), 64-71, 2003
- [8] Carleen Hutchins, "Stradivius Plate Tests" *CAS NL*. #37 p 30, 1982
- [9] Carleen Hutchins, "Plate Tuning for the Violin Maker" *CAS NL*. #39, pp 25-32, 1983
- [10] Carleen Hutchins, "Mode Tuning for the Violin Maker" *Catgut Acoust. Soc. J.* Vol 2(4) series II, pp 5-9, 1993
- [11] Robert D. Blevins, Formulas for Natural Frequency and Mode Shape, Robert E. Krieger Publishing, 1984
- [12] Evan B. Davis, *Designing Soundboards with Flexural Disk Models*, *Proceeding of Meeting on Acoustics*, Vol. 12, *Acoust. Soc. Am.* 2011
- [13] Bernard Richardson, *Guitar Making – an Acoustician's Tale*, *Proceedings of the Second Vienna Talk*, University of Music and Performing Arts Vienna, Austria, Sept. 19-21, 2010
- [14] Erik V. Jansson and Benedykt Niewczyk "Experiments violin plates and different boundary conditions", *J. Acoust. Soc. Am.* 86(3) 895-901, 1989
- [15] Henrik O. Saldner, Nils-Erik Molin and Erik V. Jansson, "Vibration modes of the violin forced via the bridge and action of the soundpost", *J. Acoust. Soc. Am.* 100(2) pt. 1168-1177, 1996
- [16] Anna Runnermalm, Nils-Erik Molin, Erik Jansson, "On operating deflection shapes of the violin body including in-plane motions", *J. Acoust. Soc. Am.* 107(6) 3452-3459, 2000

ACOUSTIC MEASUREMENTS IN THE WORKSHOP

George Stoppani
Violin and string maker,
Manchester
george@stoppani.co.uk

ABSTRACT

Violin makers in the past have been remarkably successful with minimal knowledge of sound and vibration. We know some things about their woodworking methods but nothing about how they thought about the relationship between the structure and the sound it produced. The trade has tended to insulate itself from science, preferring to refine the empirical and intuitive methods and build on the success or failure in their own work and of previous generations. The intention of incorporating acoustic knowledge and measurements into the making process is to add to the existing body of expertise rather than to replace it. Structural vibration and sound radiation measurements can be helpful in making decisions about the initial design, during the making process and after the instrument is completed. Measurements can also be useful for diagnosing problems with older instruments that are performing below expectation and even for a decision whether or not to buy an instrument.

Acoustic measurements are often disappointing when the intention is to explore the effect of usual workshop adjustments such as soundpost moves or bridge trimming. However, for a more global view of where a particular instrument lies on a map of mass, stiffness and frequency characteristics or to examine the relationship between substructures such as neck/fingerboard or tailpiece modal analysis provides insight that cannot be obtained by traditional methods.

1. INTRODUCTION

It is a logical argument that since the purpose of musical instruments is to radiate sound it is the radiated sound that should be the focus of attention and it does not matter what underlying mechanisms exist to create that sound. Recent experiments suggest that many modern makers are producing violins that sound as good or better than many old ones that are considered excellent and very valuable [1]. Whether or not this result is due to the nature of the experiments or that the superiority of old Italian instruments is, in fact entirely myth does not significantly affect the goals of modern makers. Both old and new instruments have better or worse attributes and the goal is to increase the proportion of desirable attributes and minimise those that are not desirable. Recently the

idea of “playability” has emerged as a key area where more understanding would be helpful. It is another logical view that playing differences must have something to do with how energy passes between the string and the body of the instrument which, in turn must have a lot to do with the vibrating modes of that instrument. Therefore it is measurements of structural vibration that are potentially most useful to a maker seeking to improve the quality of their work because structural vibration determines both the radiated sound and playing feel.

We have developed good protocols for measuring sound radiation spectra and this data is particularly relevant for psychoacoustic exploration of quality and preferences. Where the intention is to have control during the making process structural measurements give us insight into the underlying causes of the sound and playing feel and therefore suggest ways to effect changes.

2. A FUTURISTIC VIEW

If we had software that could process data from CT scans of violins, experimental modal analysis, sound radiation and do finite element modeling, then predicting the effect of changes might be quite precise. Important steps in that direction were taken by Bissinger: see for example [2]. For that matter, why not identify the necessary modifications for a desired sound and playing feel? There is nothing totally impossible here, only extreme complexity. Similar things are routine in industry for simpler problems. We are already trying to do this but in parts rather than as an integrated operation. We are trying to reduce the purely empirical element by combining a general understanding of how violins work as sound radiating machines with the traditional expertise that comes from a long experience of making.

3. APPROACH TO MAKING

3.1 Some objectives

Acoustic measurements are needed for two allied purposes: to look for quantifiable features that correlate with perceptions of quality and to build these features into new instruments. It hardly needs to be mentioned that neither task is simple. Anything to do with quality inevitably has a large subjective component. While others are working in this area, makers cannot wait for the results but must bulldoze through the problem and focus on some basic assumptions that will probably turn out to be more or less right. For building desirable attributes into new instruments we focus on where we have some understanding

and ability to regulate. The situation is not much different for makers who work in a traditional way, those who use elementary acoustic measurements and those that have the capability to do modal analysis, pioneered by Schleske [3]. There must be some notion of quality and a box of tricks for achieving it. However, all other things being equal, a maker with a good general knowledge of violin acoustics and measurement equipment can potentially be more effective.

In fact, the list of basic assumptions (not yet properly verified) about quality is rather long so we will limit it to a few points. Top of the list is that there should be plenty of sound in relation to the bowing effort. If not the player will be trapped in the *forte* dynamic region and unable to deliver enough volume when it is needed. There must be both good low and high frequency output: a strong air resonance A0 and a bridge hill. There should not be deep, wide valleys in the frequency response or very large isolated peaks. It must also be possible to play very quietly.

Too high minimum bow force is to be avoided (see section 6 below). This is a condition where the energy losses from the string due to the compliance at the bridge notch are sufficient to disrupt the normal Helmholtz motion of the string. At worst this would result in a wolf note but more commonly it marks a note that jumps up an octave with a light bow stroke or feels as though it is difficult to make the bow grip the string. It can be a serious problem for playing *pianissimo*.

A very condensed acoustic description of the violin could be that it is made from carved plates, rib garland (of blocks and bent strips), neck and scroll and fingerboard, all made separately then glued together. The free plates can be weighed and have familiar modes, but these are transformed by attaching to the ribs. This is because the plate edges are now connected to each other and there is an enclosed air volume. The addition of the neck and fingerboard further perturb the new set of modes as does the insertion of the offset soundpost and addition of bridge, strings and tailpiece. We have several components which are not especially acoustically complicated on their own but together they form a very complicated system with multiple oscillating substructures coupled with varying degrees of strength. We are ultimately interested in how the whole instrument behaves when fully assembled but need to be able to assess the contribution of each separate part.

3.2 A dilemma

A basic conflict is that the conditions that might lead to a minimum bow force problem are the same as those required to have a large sound output. Light wood and thin plates are good for power but at the expense of having excessive compliance and consequently large peaks in the bridge mobility. The design of the violin allows for low impedance values at low frequencies while also being strong enough to withstand the string tensions.

In the 17th and 18th centuries violins often had total string tensions in excess 30 kg. Extra wood was left around the end blocks for structural strength, plates were generally thicker and bridges were lower. We can con-

veniently divide the violin into a corpus and the *montage*. The *montage* (borrowed from the French) means all the parts that are not the corpus box itself or the scroll but should include the bass bar. It is the length, inclination and shaping of the neck handle, the fingerboard, tailpiece, bridge, strings and chin rest (if applicable). While the general shape of the corpus itself has been rather stable over the history of the violin the *montage* has varied considerably with time, region and individual musicians. Rather than all the features being a mix and match they fall into distinct packages of all components of the *montage*. That is to say that particular combinations of strings, bridge etc. work well together to suit the taste of the time and place. There is no single way to setup a violin that is right or wrong but if one deviates far from the specifications for a particular historical or modern setup there is a risk of something going wrong. A contemporary example: due to ethical problems about the supply of ebony used for fingerboards some modern makers have experimented with alternative materials and constructions. This can result in very different frequencies for the modes of the neck/fingerboard and unsatisfactory results.

Towards the end of the 18th and through the 19th centuries many instruments were regraduated (thinned) to suit preferences for lighter stringing, with graded tensions and taller bridges. Today string tensions are again getting higher to get more power and the old instruments that are now particularly desirable for soloists are the ones that have had little or no regradiation. Many old violins have been patched inside to recover plate strength lost through thinning or to deal with multiple cracks. Therefore when a maker is copying an old instrument it may be necessary to disregard the current plate thicknesses and think in terms of building a corpus that will sit under a standard bridge, strings and the rest of the *montage* and have impedance levels that are appropriate for these conditions. The historical legacy of surviving instruments and former practices clearly provides a remarkable database of success and failure and possible ways to achieve particular characteristics.

3.3 Prediction

The most basic way to have some kind of prediction of the impedance is by looking at the free plate frequencies and weights. There is much literature about free plate modes and in particular the use of #2 and #5 as indicators of cross-grain and long-grain stiffness respectively [4]. Looking at the nodal lines of these two modes it does indeed appear that #2 is bending across the width while #5 is bending across the length and it is probably true that the predominant bending components for each are like this. However, this is not quite right. #2 has anticlastic while #5 has synclastic bending. The anticlastic mode is certain to occur at a lower frequency as it requires much less energy. Both are a mixture of cross-wise and length-wise bending with some twist as well and indeed some stretching, especially for #5. One cannot deform an arched plate without invoking all types of bending. Nevertheless, free plate tuning has been part of violin making culture for a long time and there is still good reason to take note of the free plate modes.

The modes we see are a consequence of the mechanical properties of the wood, the graduation scheme, the arching and the f-hole cutouts. Therefore the mode frequencies tell us more if we take account of these things as well. If the frequencies are higher or lower than expected we ask why before deciding to make changes. It is better, but more difficult, to look at the signature modes of the completed instrument, particularly B1+, also taking into account the wood properties and arching. One very convenient way is to look at the modes of the corpus as soon as the plates are glued to the ribs but before the neck is fitted. With a soundpost fitted in the normal position, the first five modes can easily be identified simply by tapping the corpus and looking at the peaks of the FFT of a microphone signal. They will be well separated in frequency, in the right order and with no complications due to interactions with the neck, fingerboard and tailpiece substructures. Allowances need to be made for the frequency shifts that will result from adding the substructures but this is no more of a problem than using free plate frequencies while being a very much more reliable predictor. The transition from free plates to closed box is the largest step change and the least predictable in the progress to the finished instrument. It is also far more convenient to remove the top at this stage for modifications than when the instrument is complete. This is even more true for makers who complete the edgework and purfling after the corpus is closed.

3.4 Mode configuration

If a violin has a wolf note it is most likely to be a note close to B1+ but it is sometimes near B1- (see section 5 below for these shapes). If we can choose the configuration of CBR, B1- and B1+ we would probably prefer to have a reasonable gap between CBR and B1- and to have B1- and B1+ at approximately the same amplitude. This does not always happen. One common scenario is that CBR is high and close to B1- which has much larger amplitude than B1+. There will very likely be a wolf note near B1- and worse, probably close to the open A string. Another is that B1+ is quite a lot higher than B1- and has much higher amplitude. There may well be a wolf on the C natural or C sharp.

There are two main strategies to deal with this kind of situation. One would be to try to split the larger peak into a pair of “mirrored modes” each with lower amplitude than the original by coupling to a neck/fingerboard mode. A “mirrored mode” can happen when two parts of a structure are strongly coupled. Looking at only one part of the structure (e.g. the corpus) it appears that the same mode shape has occurred twice, but if the other part (e.g. the neck/fingerboard) is inspected it can be seen that it has the opposite phase in each of the mirrored pair. For B1- the most likely candidate would be the neck/fingerboard mode where there is large scroll twisting motion, and for B1+ one dominated by twisting of the free end of the fingerboard. The other strategy would be to try to correct the amplitude balance between B1- and B1+.

Before such a strategy can be considered we need to know why the modes are configured in this way and

some theory is needed. In the case of the first scenario a simple explanation might be that it is just the proximity of the two modes that makes B1- large because the operating deflection shape is the sum of all modes but dominated by the B1- mode plus large residuals from the nearby CBR. If the two modes can be separated more in frequency the problem will probably be solved. Alternatively, CBR could be coupled to a neck/fingerboard mode and split into a mirrored pair, each with lower amplitudes.

A much more challenging approach is to consider Colin Gough’s finite element model [5]. In this model B1- and B1+ are both modes that have a bending component along the length of the corpus and a volume change component due to antisymmetric plate motions. Taking a step back, arched plates, when they bend have both flexural and longitudinal deformations, or in different words, normal and in-plane motion. Because they are curved, bending across the width causes an increase in width as the arch is being flattened and a decrease as it gets higher (wrapping and un-wrapping). Differential width change of the two plates naturally leads to a bending along the length of the corpus.

It can easily be seen that B1- has anticlastic bending of the top but synclastic bending for the back and mostly bending along its length. Therefore as the volume of the box is increasing the width across the top edges at the waist is narrower than those at the back edges, and the corpus bends away from the bridge. For B1+ it is the opposite way round. With a soundpost there are other differences between these mode shapes, but as they occur in the corpus without a soundpost they look like coupled modes that differ only in the phase of the bending component. A consequence of the way these modes are coupled by the strength of the rib structure is that the pair can occur in three different configurations. The optimum would be close to a situation where both have an equal proportion of bending and breathing components and equal sound radiation. If the coupling is too small B1+ will have large volume change, less corpus bending and stronger radiation whereas B1- will be dominated by corpus bending and have weaker radiation. If too large it will be B1- that is the stronger radiator while B1+ has more bending than breathing. In the last two cases the frequency separation will be greater. The model thus provides us with some indicators as to which of these configurations we are looking at and possible causes. We can get an idea of the bending/breathing ratio from the sound radiation measurement or directly from modal analysis measurements by looking at the sum and difference of the two plates or calculating the volume change component.

In practice it has been possible to identify which configuration is present but there has been only limited success in changing the configuration at a later date. Clearly, this is an area where synchronising the FEA model with the measured data could provide a very powerful tool. In the meantime the approach has been to monitor how these configurations have turned out for particular violin models and where they have been less than ideal to try to build the next one in a slightly different way. For example, we would try to adjust the rib coupling strength by using different thicknesses of ribs, linings of different

dimensions and different thickness for plate edges. The bending component stiffness might be more associated with thick, stiff backs than stiff tops. We might associate the frequency of the breathing component with wood density and thickness of the plates away from the edges. There has been some success in avoiding the less desirable configurations.

Rather than having a set formula for dealing with such scenarios it seems to be better to look at the modal analysis data and everything else about the violin. We should know which areas are more robust and can afford to lose some strength and we should be able to see if a frequency shift for a substructure is practical. We would not embark on a modification process solely on the basis of something that looks wrong or unexpected in the acoustic measurements unless there was also a perceived problem in the sound or playing feel. If a problem is observed we would mine all the available data for possible underlying causes. Most often a problem has a cluster of symptoms and there will be evidence to support a diagnosis in both the acoustic measurements and known details of the design and materials. For example: “too” stiff or “too” compliant is a property that we can learn to recognize from the bowing feel; it is often confirmed by the signature mode frequencies.

4. SOME EXAMPLES OF INTERVENTIONS

Regraduation in order to lower mode frequencies is the most common intervention after an instrument is finished. Being woodworkers we habitually leave allowances that can be trimmed at a later time. As we get more skilled and experienced, the allowances get smaller. In violin making a large proportion of the necessary decisions are about whether or not to cut away a bit more wood and where to remove it from. The making process could be described as a long chain of decisions, many of which have to be updated as the work progresses and many of which could compromise the result if they are wrong. It is unlike an industrial scenario where a product is designed by engineers and the manufacture is executed according precise specifications by other people. This would not work for making high quality violins because it would be inevitable that opportunities to adjust would be missed. It would be fair to say that many of the judgements are fuzzy but that is one of the things that violin makers are sometimes good at.

It is by far a preferable scenario to open the box and remove some wood than to have to add wood to raise the frequencies. It is not so easy to judge exactly the optimum mode frequencies at the outset since frequencies are not the only thing that matters. Attempts to improve low frequency mode shapes by local thinning have little success. The shapes are determined mostly by arching and more or less any wood removal lowers the frequencies.

Occasionally it is necessary to change the angle of the neck set. Adjustments to the neck and fingerboard and to the tailpiece can sometimes be very effective for a small amount of effort and can be reversed much more easily than regraduation.

4.1 To regraduate or not?

A violin based on the “Titian” Stradivari had free plate frequencies that were unexceptional, while a Guarneri del Gesù model with a thick back had a high lying #5 for the top and very high #2 and #5 for the back (table 1.) However, the CBR, B1- and B1+ were higher for the Stradivari model (table 2.) It was later decided to regraduate to lower these frequencies because it seemed too soprano and slightly lacking in depth (table 3). The Guarneri del Gesù model was left as it was. If only the free plate frequency information had been available then the decision would probably have been to reduce the back thickness of this violin.

Signature modes tend to rise in frequency over the first 2 or 3 years so if the instrument exhibits characteristics of too high mode frequencies at the outset it can be expected to get worse over time.

	Top #2	Top #5	Back #2	Back #5
Titian Stradivari copy 2011	169	376	161	376
Del Gesù model 2013	158	398	178	448

Table 1. Free plate frequencies for two violins.

	A0	CBR	B1-	B1+
Titian Stradivari copy 2011	275	436	502	621
Del Gesù model 2013	277	432	476	592

Table 2. Corpus only with soundpost.

	A0	CBR	B1-	B1+
Titian Stradivari copy with strings	258	417	462	557
Titian Stradivari copy with strings after regraduation	271	406	443	537

Table 3. Titian Stradivari copy with strings before and after regraduation.

4.2 A small adjustment to a fingerboard

Another “Titian” copy that had lower signature mode frequencies developed a wolf note around the C on the A string after about 18 months. This was not only an irritation when playing that note, particularly up on the G string but the C on the G string an octave below was uncooperative, feeling as though there was not enough rosin on the bow. An inspection of the relationships between B0 (flapping of the end of the fingerboard) and B1+ and the fingerboard twisting mode showed that B0 was a few Hz higher than the peak in the internal air cavity (A0) and that the fingerboard twisting mode was creating an additional peak on the shoulder of the B1+ peak. This was believed to the cause of a minimum bow force problem. A closer A0/B0 match would be unlikely to make matters worse, while the fingerboard twisting mode has been seen to be associated with wolf notes in other cases. An obvious intervention to try was to lower both these finger-

board frequencies by removing wood from underneath the fingerboard close to the neck root. Removal of a remarkably small amount of wood lowered both frequencies by several Hz. Both playing problems were much improved and there was a perception of bowing being a more satisfying experience on all notes.

When A0 and B0 are close in frequency it may be difficult to tell which is which. An effective way to inspect the situation is to place a small microphone inside the corpus and an accelerometer at the end of the fingerboard. Two peaks will be observed in both responses. However, one peak will usually be larger for the internal air and the other larger for the fingerboard. If the peaks are the same amplitude for each transducer then the two modes are well coupled. The shapes will be the same except that the fingerboard phase is flipped.

4.2 An experiment for playability

The specialist cello maker, Robin Aitchison did an extended experiment in which he made the plates of a cello deliberately too thin and then glued external pieces of wood to put back stiffness into selected areas. (This was similar to Sam Zygmuntowicz's "gluey" experiment.) He worked with a professional cellist to evaluate the effect of these modifications. As a pilot experiment for the 2012 Cambridge Acoustics Group meeting we continued the experiment. Robin was able to quickly switch the cello between better and worse states by adding or removing bars. The minimum bow force plots were derived from laser measurements of the bridge mobility. We also evaluated the playing feel at each stage and found a strong correlation between the minimum bow force plots and the perceived bowing difficulty.

We went a stage further, did a modal analysis and extracted the local bending from the data instead of the displacements. Looking at the bending plots for modes close to problem notes we selected areas of large bending as sites to place stiffening bars. Not all selected areas had the desired effects but some did lower the minimum bow force and improved the playing feel. In the light of this experiment Robin has reviewed his graduation schemes.

5. SIGNATURE MODES

The idea behind the term "signature modes" is that they characterize an instrument in terms of a set of modes that we expect to see on all instruments, with recognizable shapes that occur in the same frequency sequence in much the same way as #2 and #5 of the free plates are taken as indicators for the completed instrument. It is a useful concept but needs some qualification. It could be argued that there is no special reason why this group need be confined to the modes up to C4. There are several more modes that are also recognizable in most violins though they do not always occur in the same sequence. In fact, for cellos CBR is often higher in frequency than B1-. For most fully assembled instruments at least one of this set is split into a pair. In cellos we often see, instead of CBR and B1-, a collection of mixtures of two. We

also sometimes see an alternative version of CBR with the plates in the opposite phase relationship. A sound radiation spectrum can leave uncertainty as to the origin of the response peaks.

Without the montage the complications go away and without a soundpost we can see the immediate effect of gluing the plates to the ribs. In this state it is far easier to compare the FEA model with the measured data. The following plots show data measured by the author. With exception of the C4 plots they are averaged from violins by several different makers. The convention adopted here is that both plates are seen from the outside like opening a book and seeing the front and back covers at the same time. The colour code shows positive motion in warm colours, negative in cold colours and nodal lines in white.

5.1 Signature modes without soundpost

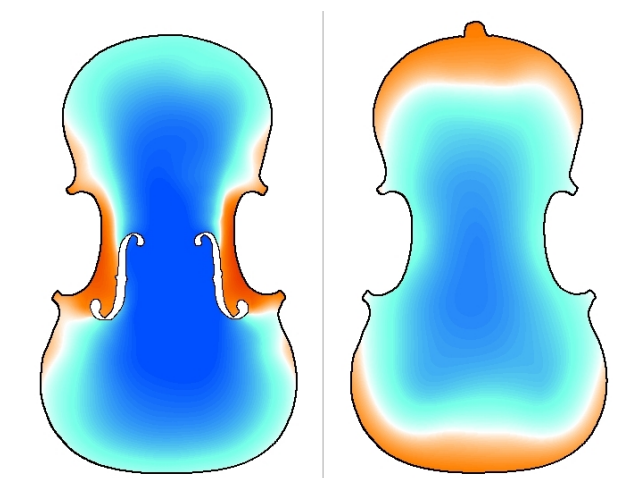


Figure 1. Typically at around 245 Hz

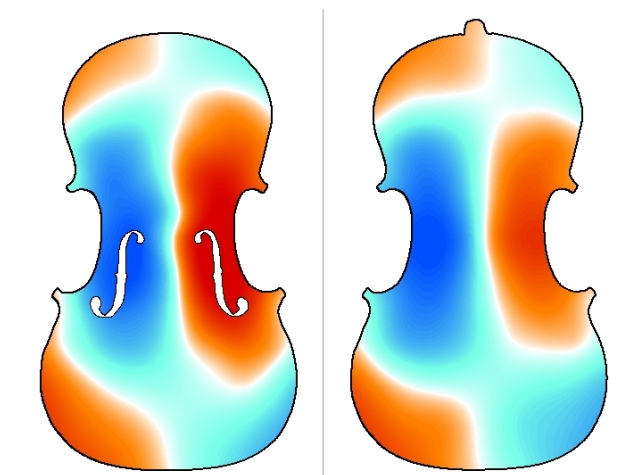


Figure 2. CBR typically around 400 Hz, (hardly affected by the soundpost)

The modes in figures 3 and 4 at around 400 Hz may be switched in the frequency sequence but always seem to occur very close in frequency. The "island twist" and the bending/breathing mode at around 585 Hz (figures 4 and 5) may also be switched and are also always close in frequency. These two pairs are each of one symmetric and

one antisymmetric mode. The asymmetry due to the bass bar sometimes seems to couple the pair weakly, giving mode shapes where each has a little of the other mixed in. No reason is known why these two pairs always appear close together or if it is a good or bad feature. They can rarely be distinguished directly from the operating deflection shape and need to be curve-fitted.

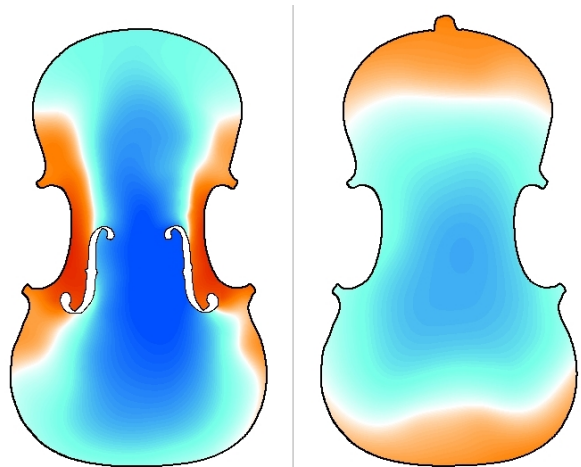


Figure 3. Typically around 400 Hz. The same bending/breathing phase relationship as B1-.

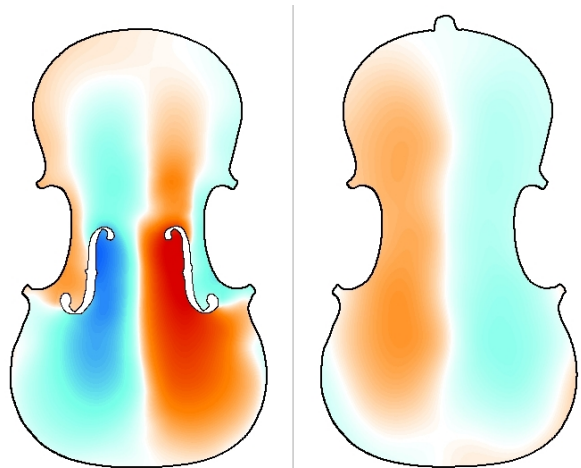


Figure 4. "Island twist", typically around 570 Hz

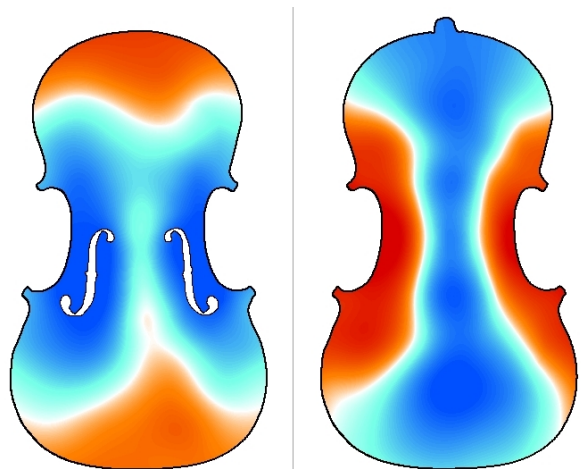


Figure 5. Typically around 585 Hz. The same bending/breathing phase relationship as B1+.

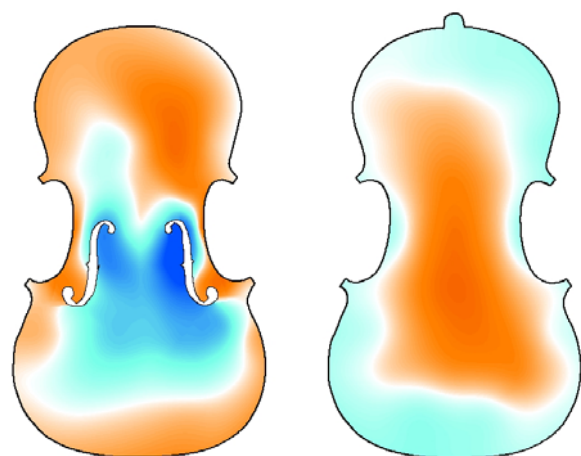


Figure 6 Typically around 725 Hz. Recognisably the C4 shape.

Note that the lowest mode in figure 1 is almost identical to that in figure 3. The difference is the phase of the internal air. The finite element model predicts a "breathing" in the absence of enclosed air at around 300 Hz. These two modes are therefore a result of coupling of the enclosed air to the breathing mode. Note also the opposite bending phase of figure 1 and 3 compared to figure 5. The bending phase is a means to distinguish B1- and B1+ in cases where there is ambiguity.

5.2 Signature modes with soundpost

The frequency shifts are related to how much the nodal lines are moved. The modes in figures 4 and 5 appear to have amalgamated to make B1+ whereas those in figures 1 and 3 are the same those in figures 7 and 9 except for the shift in the nodal line towards the soundpost position. CBR is minimally affected as the soundpost does not perturb the shape. It is not clear why C4 goes down in frequency. Sometimes the shape is minimally affected by the soundpost.

Adding the neck and fingerboard causes a general downward shift even when there is no sign of strong coupling to neck/fingerboard modes.

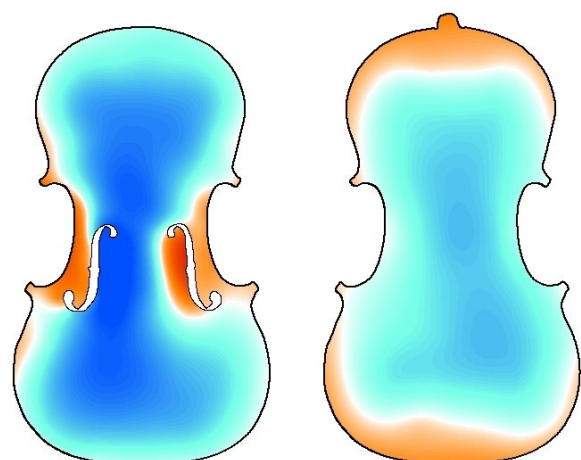


Figure 7. A0 Typically around 275 Hz.

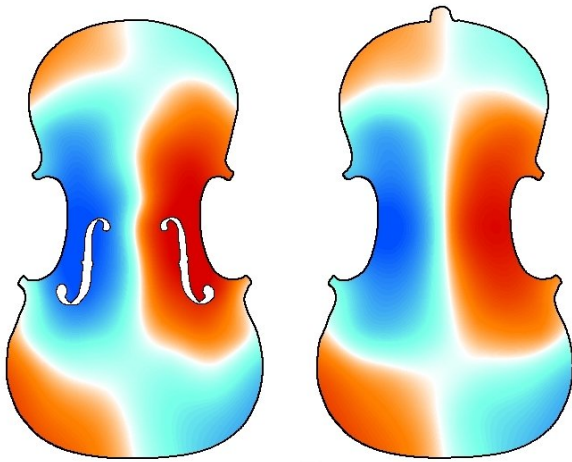


Figure 8. CBR still typically around 400 Hz

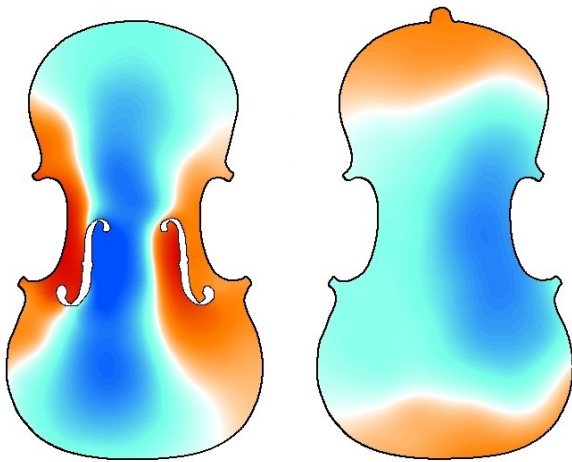


Figure 9. B1- typically around 450 Hz

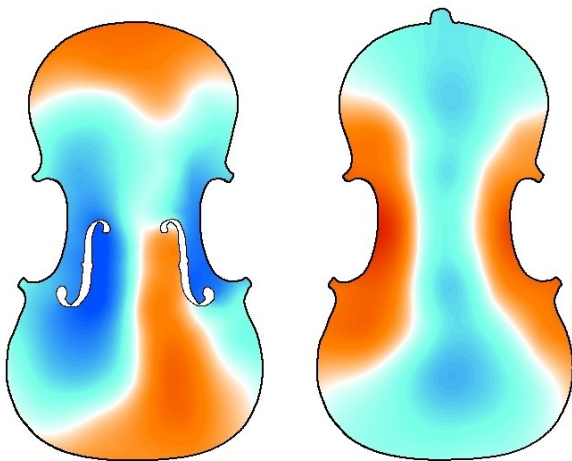


Figure 10. B1+ typically around 550 Hz

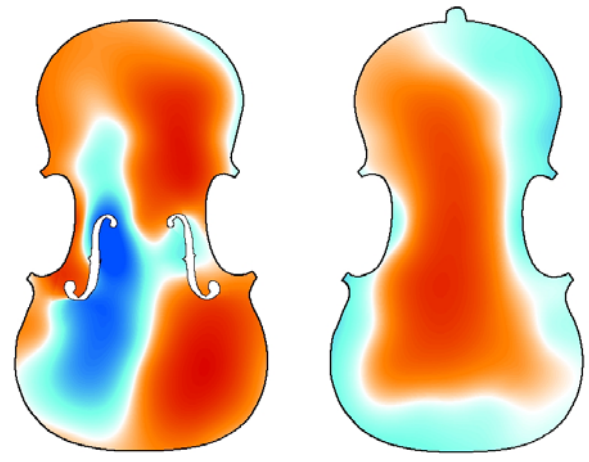


Figure 11. C4 often comes down in frequency with soundpost.

6. MINIMUM BOW FORCE

The calculation referred to here is the one suggested by Woodhouse [6]. It develops the idea behind Schelleng's bow force plot to take account of the sawtooth waveform of the string and the frequency dependence of the losses at the bridge (rather than just a constant). Ideally, the mobility would be measured at the string notch on the bridge but an ordinary lateral bridge mobility measurement is an adequate approximation. If the scaling operation for a known string impedance is omitted the calculation still gives a good indication for where there may be a problem. When we say that a low minimum bow force is a positive attribute it does not necessarily mean that an overall lower minimum bow force is any better. It is more a matter of not having notes for which it is high enough to cause a problem. A very low minimum bow force suggests that the string energy is being rejected and therefore there is also very weak sound radiation. Even the most highly rated violins have a very ragged spectrum for both the bridge mobility and the radiated sound but we might like to be able to redistribute the energy by shaving a little off the peaks and filling the valleys. It is an interesting learning exercise to play an instrument where the minimum bow force plot is available.

7. CONCLUSIONS

In recent years it has become practical for makers to do experimental modal analysis in their own workshops and to make high quality sound radiation measurements. The first application of these tools was to collect data on instruments that were considered excellent and might serve as models for new instruments. Almost anyone can be taught how to do the measurements, but the process of interpreting the data in terms of what constitutes good or bad in a particular instrument, or whether particular features are of any significance has proved far more difficult. Progress has been made on all fronts, though it is still, and perhaps always will be, "work in progress".

Two main areas seem to offer good prospects: further development of the FEA model and synchronisation with

the measured data, and collaborative projects between makers, musicians and scientists.

Acknowledgments

There are too many people who have been helpful in their various ways to mention them all. For specialist knowledge I thank in particular Professor Jim Woodhouse, Cambridge University Engineering Department, Colin Gough, Emeritus Professor of Physics at the University of Birmingham and Evan B. Davis Ph.D. Technical Fellow, Acoustics Technology, Boeing Commercial Airplanes. Among fellow makers: Joseph Curtin, Samuel Zygmuntowicz and Terry Bormann. I should also mention Martin Schleske who started working in this area years before the rest of us even thought about it. The workshops organized by the Violin Society of America in Oberlin have been very valuable for their freely flowing information and sharing of ideas.

8. REFERENCES

- [1] C. Fritz, J. Curtin, J. Poitevineau, P. Morrel-Samuels, and F-C Tao, "Player preferences among new and old violins", *Proc. National Acad. Sciences*, vol. 109, pp. 760-763, 2011
- [2] S. Zygmuntowicz, G. Bissinger, "Strad 3D", 2009: see web site <http://strad3d.org/cms/>.
- [3] M. Schleske, "Empirical Tools in Contemporary Violin Making: Part I. Analysis of Design, Materials, Varnish and Normal Modes", *Catgut Acoust. Soc. Journal*, vol. 4, no. 5, 2002; and "Empirical Tools in Contemporary Violin Making: Part II: Psychoacoustic Analysis and Use of Acoustical Tools" *Catgut Acoust. Soc. Journal*, vol. 4, No.6, 2002.
- [4] A. Buen, "Two models for predicting the violin timbre from the material properties of the top and back plates", Brekke & Strand akustik as, Norway, 2004.
- [5] C. Gough, (elsewhere in this volume), 2013.
- [6] J. Woodhouse, "On the playability of violins: Part 2", *Acustica*, vol. 78, pp. 137-153, 1993.

On perceptual evaluation of instruments: The case of the violin put into perspective

Claudia Fritz

LAM, Institut Jean Le Rond d'Alembert, UMR 7190, CNRS / Université Paris 06
claudia.fritz@upmc.fr

ABSTRACT

An object of study in mechanics for more than three hundred years, the violin has only recently been scientifically studied from a perceptual point of view. The range of experiments which have been conducted since 2005, complemented by a few studies of other instruments, offers a nice illustration of possible methodologies, and serves as a basis for discussing their respective advantages, as well as their limitations and issues. While there is no general recipe (since methodological choices depend largely on the goals of the study), ecological validity is certainly one of the key prerequisites for a perceptual experiment seeking to provide meaningful results, in the sense that they can be generalised from laboratory settings to “real life” situations.

1. INTRODUCTION

The violin is one of the most culturally important instruments in Western society, and has therefore been extensively studied by scientists for more than three centuries [1]. Most of these studies were anchored in mechanics and focused on physically modeling the instrument and its control by the player (e.g. recently [2–9]).

A long-standing goal of violin research has been to correlate the violin's playing and sound qualities with specific attributes of physical structure and dynamic behavior, which could be, in particular, extracted from admittance and/or radiation measurements. For instance, Alonso Moral and Jansson [10] suggested the importance of the signature modes below 600 Hz and the bridge hill in the 2–3 kHz range for violin sound quality, based on bridge admittance measurements on 24 violins, which had previously been played and rated on tonal quality by two professional violinists. Hutchins [11] suggested a correlation between the spacing between the A1 and B1 modes measured on 37 violins and comments about the quality of these violins. These comments were however only made by the respective player or owner of each violin. Dünwald [12] measured the sound output of 700 violins from a single microphone position and derived a combined parameter based on spectral considerations which, along with the

level of the A0 mode, allowed him to categorize the 700 violins into classes, and to separate the good violins from the bad violins. However, what is meant by good and bad, and who decided which violins were good or bad, is not specified. It seems this was solely based on the reputation of the makers and/or the owners (famous soloists for instance). Langhoff *et al.* [13] conducted experiments in which violin performances were filtered digitally. They used one violin as a baseline and then modified its frequency response curve (and therefore its impulse response) in several ways, to give enhancement of the Helmholtz resonance and of mid-range frequencies (around 1.7 kHz), and creation of a smoother decay towards higher frequencies. This experiment did show that it was possible to compare violin spectra by listening to digitally filtered signals but it did not address the question of how people perceived the different sounds created. Their paper only reports the “subjective impressions” of one of the authors and no other participant was involved. Finally, Bissinger's wide range of vibration and radiation measurements [14] of 17 violins with quality ratings from bad to excellent showed no significant quality differentiators except for the Helmholtz-like cavity mode A0, the radiation of which was significantly stronger for excellent violins than for bad. However, the quality ratings were provided by just a single professional player for 12 of the violins, and by Bissinger himself for 5. Few details on the rating procedure were provided.

This search for correlations between “subjective” and “objective” properties is not uncommon in the musical acoustics field, especially in the last five decades, and has been growing for the last fifteen years. Examples can be found for all kinds of instruments: starting with the piano in 1962 [15], then followed by the trombone [16], French horn [17], trumpet [18–21], flute [22], bassoon [23], clarinet and saxophone reeds [24–27], didgeridoos [28, 29], cellos [30], oboe [31], electric guitar [32],... However, what distinguishes these examples from the violin studies mentioned previously is that they contain controlled (at least to some extent) objective perceptual studies. Even though most of the effort was generally put into the mechanical measurements, and thus some of the perceptual studies were rather small scale, and their design and control could have been improved, they have the merit of having led to perceptual properties which were not just the views of the authors, and which could therefore serve as an objective basis for correlations with mechanical properties. In the case of the violin, no well-controlled study on how violinists rate and evaluate violins has been done until recently (2005),

and thus the main problem with the attempts at correlating dynamic measurements with perceptual data in search of quality parameters has always been the lack of convincing and objective perceptual data. Moreover, while many studies actually tried to explain the presumed tonal superiority of the Old Italian violins, and many factors had been proposed and/or investigated to account for it (including properties of the varnish [33, 34]; effects of the Little Ice Age on violin wood [35]; differences in the relative densities of early and late growth layers in wood [36]; chemical treatments of the wood [37, 38]; plate tuning methods [39]), no one actually investigated the fundamental premise of tonal superiority!

Based on historical considerations for the violin, the paper is divided in two parts, corresponding to the two main classes of perceptual experiments: listening and playing tests. Illustrations are provided for other instruments, and general limitations and issues are discussed.

2. FROM LISTENING TESTS ...

2.1 For comparing modern violins to Old Italians

In the violin world, many informal blind tests have been conducted since 1817 [40] to the present and the results all showed that new instruments stand up very well, and often outscore their older, more expensive counterparts. Some tests were particularly famous: a BBC programme in 1977 with violinists Isaac Stern and Pinchas Zukerman and the violin expert and dealer Charles Beare [41]; in Sweden in 2006 with an audience mostly comprising members of the European String Teachers Association [42]; and in 2009 with the British violinist Matthew Trusler, who played his 1711 Stradivarius and four modern violins made by the Swiss violin-maker Michael Rhonheimer. One of the new violins was made with wood that had been treated with fungi [43]. However, each test has been discredited or dismissed as meaningless by the experts as they were unscientific and relied on flawed methodology. In particular, they were rarely conducted in the double-blind format, where neither the panel of judges nor the player knows the identity of the violins being evaluated.

The first scientific test to address not exactly this specific issue of ‘Old Italians versus new violins’ but the more general issue of the effect on violins of ageing and playing was conducted by Inta *et al.* in 2005 [44]. A pair of violins that were as similar as possible was commissioned. One instrument was then kept for three years under environmentally controlled conditions in a museum, whilst the other was played regularly by a professional musician. Listening tests were conducted ‘live’ in a concert hall when the violins were new, and then three years later. The violins were presented a certain number of times in random order, and listeners rated them on a 1 (poor) to 10 (excellent) scale in five categories: *evenness*, *clarity*, *projection*, *distinctive character* and *warmth*. Instruments were rated on the sound alone (as opposed to performance quality). Results showed no significant differences at the 98% confidence level, for any criteria, between the two violins.

2.2 For searching perceptual correlates of acoustical parameters

Almost simultaneously, Fritz *et al.* [45–48] started to establish quantitative links between acoustical parameters of the instrument body and the perceptions of a listener, using the methodology of “virtual violins”. Representative force waveforms are recorded using normal playing on a violin whose bridge is equipped with a piezoelectric force sensor under each string. These pre-recorded force functions can then be applied to different violins, so that sound differences can be compared with no complications arising from variations in playing. The mechanical frequency response function of these different violins was mimicked using a digital filter, and the output signal for listening tests was generated by applying this filter to the recorded bridge force signal. Once the violin response is represented in digital filter form, it becomes very easy to make controlled variations of a kind which would be virtually impossible to achieve by physical changes to a violin. This methodology is similar to the one used by Langhoff *et al.* [13] except that the violin bridge was mounted with a force sensor (and not a velocity sensor, from which it was difficult to derive the force). The goals were also different. For instance the aim of Fritz *et al.* [46] was to report the results of psychoacoustic measures of the ability of musically and non-musically trained subjects to discriminate changes in the frequency and amplitude of single and multiple resonances. This initial study explored two aspects of violin acoustics which received great prominence in the earlier literature as possible indicators of ‘quality’: (1) the three individual low-frequency modes of vibration (below 700 Hz), which dominate the low-frequency output of a violin and are usually labeled A0 (a modified Helmholtz resonance), B1- and B1+ (two strong ‘wood modes’); and (2) a set of four frequency bands proposed by Dünwald [12] (190-650 Hz, 650-1300 Hz, 1300-4200 Hz and 4200-6400 Hz). For modifications of amplitude, the lowest thresholds were in the range 3-5 dB for individual modes and 1-3 dB for the Dünwald bands. For modifications in frequency, the best listeners had thresholds around 3-5% for individual modes, 1-3% for the first three Dünwald bands, and around 1% when all frequencies were varied simultaneously. Frequency changes in the 4th Dünwald band were not detectable.

In [47], Fritz *et al.* explored how the perception of violin notes is influenced by the magnitude of the applied vibrato and by the level of damping of the violin resonance modes. Damping influences the “peakiness” of the frequency response, and vibrato interacts with this peakiness to produce fluctuations in spectral content as well as in frequency and amplitude. Initially, it was shown that thresholds for detecting a change in vibrato amplitude were independent of body damping, and thresholds for detecting a change in body damping were independent of vibrato amplitude. A study of perceptual similarity using triadic comparison showed that vibrato amplitude and damping were largely perceived as independent dimensions. A series of listening tests was conducted employing synthesized and recorded performance to probe perceptual re-

sponses in terms of *liveliness* and preference. The results do not support the conclusion that *liveliness* results from the combination of the use of vibrato and a “peaky” violin response. Judgements based on listening to single notes showed inconsistent patterns for liveliness, while preferences were highest for damping that was slightly less than for a reference (real) violin.

Finally, Fritz *et al.* [48] investigated the perceptual effect of acoustical modifications of violin sounds produced by a roving of the levels in five one-octave wide bands, 190-380, 380-760, 760-1520, 1520-3040, and 3040-6080 Hz. Pairs of sounds were presented, and each participant was asked to indicate which of the sounds was more *bright*, *clear*, *harsh*, *nasal*, or *good* (in separate runs for each descriptor). Increased *brightness* and *clarity* were associated with moderately increased levels in bands 4 and 5, whereas increased *harshness* was associated with a strongly increased level in band 4. Judgements differed among participants for the qualities *nasal* and *good*. The criteria were chosen after an exploration of the verbal descriptions, made by performers, of the distinctive timbres of different violins. Sixty-one common descriptors were collected and then arranged by violinists on a map, so that words with similar meanings lay close together, and those with different meanings lay far apart. The results of multidimensional scaling demonstrated consistent use among violinists of many words, and highlighted which words are used for similar purposes.

2.3 For other goals like testing the realism of a sound synthesis model or of an artificial player

Listening tests have been designed to address a wide range of goals and this paper does not intend to present all of them. Two examples are given below as they offer interesting illustrations and expand the scope of this section to other instruments.

Poirson *et al.* [19] investigated the brightness of trumpet tones which were generated on a trumpet mounted with a mouthpiece of variable depth in three ways: by a musician, by an artificial mouth and by physical modeling simulations. This study allowed the authors to find that the magnitude of the impedance peak corresponding to the second harmonic of the tone was highly correlated with brightness, and seemed to be the cause, which falls under the previous section. It also allowed them to compare the three ways of generation and thus to check the realism of their artificial mouth as well as their numerical model. Sounds generated by the artificial player or simulated by the harmonic balance technique were found to be perceived in a similar way to the natural sounds in terms of brightness. This augurs well for the use of the artificial player for studying the quality of wind instruments (as it is highly reproducible compared to a real player) or virtual acoustics techniques in the conception of new instruments.

Bensa *et al.* [49] aimed at determining the perceptual influence of two control parameters of a piano sound synthesis model: inharmonicity and “phantom” partials. 17 piano sounds were synthesised so that they varied in terms of these two control parameters, in order to cover a wide

range - from a sound with very weak inharmonicity and no “phantom” partials to a sound with exaggerated inharmonicity and a high level of “phantom” partials. These 17 sounds were used in a free categorisation task, and the results put into perspective to derive general conclusions about timbre cognitive processing. In particular, the study shows major differences between the physical and the cognitive descriptions, the first one having a dimensional character, and the second being categorical. Different categorical structures can correspond to the unique description of the stimuli in the physical space and they depend on the strategies of the subjects, based on their expertise and their experience.

2.4 Discussion

About the difference between old and new violins, Coggin [42] writes: “Perhaps the real answer, though, lies not so much in the actual sound that is produced, but more in some intangible interaction between the player and the instrument”.

Regarding the experimental investigation of the perceptual correlates of violin acoustics, the authors acknowledge in [46] that their results are only part of the story of violin discrimination, as higher-level perceptual processes are brought into play when a trained violinist compares instruments in a musical setting - for example during the process of choosing a new instrument. It was thus not surprising that differences were obtained in [47] between the judgements by violinists made with synthetic or live performances. Live performance was achieved by playing on an electric violin whose output was filtered in real time by the same filters used to synthesize the sounds for the listening tests. Even if the evaluation is restricted to sound quality, four causes can explain the differences between a playing evaluation and a listening one. First, playing tests allow many more possibilities for evaluating the timbre of an instrument than listening tests, which are usually reduced to short musical excerpts of even single notes because of time constraints (to reduce listeners’ fatigue). Second, the same instrument may sound very differently when played by somebody else: a violin does not have a sound per se! Third, the sound quality of an instrument is evaluated differently when listening while being “passive” with respect to the sound production - i.e. listening to the sound produced by someone else - compared to listening while being “active”, i.e. while generating the sound. The evaluation during a listening test is indeed made by relying on the sound only and is thus mainly based on the resultant sound without any possible comparison nor control on the nature of the sound and the manner by which it was produced. However such processes of comparison and control of the instrument when producing the sound are essential for the evaluation of the quality in general and the sound quality in particular of a violin by violinists, as proved by the agency given to the violin in the assessments given during the playing task and the players’ statements regarding what is a “good” or a “bad” violin [50]. And finally, in playing tests, sound quality may be entangled with playing properties (such as playability, response, ...). Listening tests are

thus of limited use in the general issue of “instrument quality” and this was already discussed by Pratt and Bowsher in 1978 [16]. Having conducted a preliminary listening test followed by a large scale one, they concluded: “In view of the difficulties experienced by listeners in discriminating between instruments and players, and also since listeners can rate only the timbre of the instruments, it was decided to concentrate on the use of players as subjects for the remaining experiment.”

3. ...TO PLAYING TESTS

3.1 For searching relationships between “subjective” and “objective” assessments of instruments

With the exception of [16] (though it was followed by [51]) and very recent papers on the violin (see section 3.2), papers containing a perceptual study aimed at establishing correlations, quantitative and systematic relationships and, even better, causal links between perceptual properties and mechanical/acoustical characteristics in order to: 1. search for quality parameters, i.e. the determinants of playing quality, in the case of the didgeridoo [29], saxophone or clarinet reeds [24–26], violin bow [52]; 2. check the influence of a single construction parameter, like the geometry of the mouthpiece of the French horn [17], the crook profile of the bassoon [23] or the neck-to-body junction of the electric guitar [32]; 3. check manufacturing consistency, in particular defects like a leak in the bore of a trumpet [21] or differences in bore profiles of oboes [31].

However, there has been little interest until quite recently in better understanding how players evaluate instruments per se, with the search for quality parameters only as a long term goal, after the musicians’ evaluation had been properly addressed.

3.2 For studying how violinists evaluate violins

Weinreich wrote in 1993 [53] “*no [objectively measurable] specification which successfully defines even coarse divisions in instrument quality is known* (author’s italics)” and this still remains a challenge. But finding such a specification would be easier if there were a general consensus among violin players, makers, and dealers on how to rate violin quality, and on which particular violins are better than others, and in what particular ways. In addition, Bissinger remarked, at the end of his largely inconclusive study [14]: “What truly defines violin excellence? If the answer is truly excellent violinists, then the reliability/reproducibility of their psychoacoustic judgements must draw more attention.” This has been the starting point of some of the research conducted recently by the author and her collaborators.

Saitis *et al.* addressed the question of self consistency of experienced violinists as well as between-individual agreement in a series of experiments [54, 55]. Only what is considered by the author as methodologically important will be reported here, and the reader is invited to refer to the corresponding papers for further details. The first playing test [54] involved 20 skilled violinists who had to rank in order of preference 8 violins of different make, age and

price. Players were asked to use their own bow, which was considered an extension of their arm; we thus avoided the potential problems of using a common bow (e.g., participants being uncomfortable with a bow they are not familiar with) which would potentially trigger a similar quality debate [52]. Moreover, having the participants use the bow that they are most familiar with is more representative of how violinists assess instruments while in the process of purchasing one. The experiment took place in a relatively dry (acoustically) environment to help minimize the effects of room reflections on the direct sound from the violins [56]. To circumvent the potential impact of visual information on judgement while ensuring a certain level of comfort for the musicians, we used low light conditions and asked participants to wear dark sunglasses. Finally, in order to maximize ecological validity and emulate as closely as possible a real situation of instruments’ evaluation that could, for instance, happen in the process of purchasing a new instrument, no playing constraint was imposed on the evaluation process (e.g., specific repertoire). The experiment was divided into two identical sessions, at least three days apart. In each session, after a familiarisation phase, participants had to do the ranking five times (the violins being placed on a table in random order each time).

In the second playing test [54], 13 skilled players had to assess 10 violins (of different make, age and price), by five criteria. To ensure common interpretation of the rating scales across all participants as much as possible, each criterion was presented in the form of a descriptive phrase alongside a short explanatory text (not provided here): *the violin is easy to play, the violin responds well, the violin has a rich sound, the violin is well balanced across all strings, the violin has a broad dynamic range* and overall preference. The right end of each scale was labeled as “strongly agree” while the left end was labeled as “strongly disagree”. Violins were presented one at a time and the experiment was conducted under the same practical conditions as the previous one.

The results of the first experiment showed that players are self-consistent when assessing different violins. However, a large amount of inter-individual variability was present in the preference rankings. Overall, known characteristics of the participants (e.g., years of violin training) did not appear to explain self-consistency. The results of the second experiment showed that the perception of the same violin attributes widely varied between individual players and corroborated the large inter-individual differences in the preference for the violins observed in the first experiment. Despite the variability in the evaluation of both preference and violin attributes, violinists appeared to agree on their preference for violins with a rich sound and, to a lesser extent, a large dynamic range.

One of many hypotheses about the origin of the large inter-individual differences in violin preference is that players may take varying playing approaches to assess different attributes of the instrument. A third experiment [55] was thus designed to investigate the perceptual evaluation of *richness* and *dynamic range* in playing tasks based on pre-

scribed musical material and techniques in order to compare intra-individual consistency and inter-individual agreement in constrained (i.e., playing only certain notes on certain registers) versus unconstrained (i.e., playing a certain excerpt from the violin repertoire) tasks for the cases of these two criteria. 16 skilled players were asked to rank/rate five violins (of different makes, ages and prices), presented simultaneously, in terms of *richness* and then *dynamic range*, first in a constrained task, and then in an unconstrained task, for which they had to indicate their overall preference as well. For each trial, the assessment was done on five discrete scales (one for each violin) presented simultaneously on a computer screen (using on-screen sliders) and participants were instructed to always rate their top choice as 1 and their lowest as 0. The results of this third study show a higher inter-individual agreement in the playing tasks relative to the previous studies. On the one hand, this observation seems to support the hypothesis that different violin players may take varying approaches to assess different attributes of the instrument and hence designing focused evaluative tasks may trigger more agreement between individuals. On the other hand, it is possible that participants were able to agree more with each other because they had to evaluate only five violins, a smaller number than in the previous studies. In terms of self-consistency, participants appeared slightly more self-consistent in this study than in the second experiment. To a certain extent, these observations may suggest that when evaluating a set of violins, comparing all instruments at a time is more meaningful and thus more reliable than assessing each violin individually.

In parallel, Wollman *et al.* [57] have been investigating the role of auditory and tactile modalities when evaluating the quality of a violin. In particular, they designed an experiment employing a blind violin evaluation task under different conditions: i) by holding the instruments without producing sound ii) under normal playing conditions, iii) with auditory masking, and iv) with vibrotactile masking. Under each playing condition, 20 violinists evaluated five violins according to criteria related to violin playing and sound characteristics, rated their overall quality and relative preference. Both auditory and tactile modalities appeared important in the violinists' evaluations, but their relative importance was found to depend on the violinist, the violin and the type of evaluation (criteria or preference). In particular, the importance of the sound of a violin to its preference depends on the violin. The investigation intended to establish as well a correspondence between the different attributes of a violin and the sensory modality they appear to be associated with. Three separate groups of criteria were suggested. One group consists of criteria mainly related to violin sound, namely *sound richness* and *sound palette*, though about a third of the players could still judge these criteria with auditory masking! A second group consists of four criteria that relate to both auditory and tactile modalities, namely *liveliness*, *dynamics*, *loudness/power* and *evenness*. The third group consisting of *responsiveness* and *ease of playing* includes criteria that depend to a large extent on tactile cues. Finally, the over-

all quality ratings were accurately predicted by the rating criteria, which also proved to be perceptually relevant to violinists, but were poorly correlated with the preference ratings, suggesting that the two types of ratings may stem from different strategic decision processes.

A first attempt by Fritz *et al.* [58] at using verbal data to better understand how violinists evaluate and conceptualise violin quality focused on the differences between playing and listening. The study showed that there are clearly two different objects under consideration for the musician: the violin and the sound. As far as the psychological evaluation is concerned, musicians mainly focus on their relationship with the instrument while playing (in all the polysemy of the word) with it, the produced sound leading to an eventually different evaluation while listening. The conceptualisation of violin quality evaluated when playing has then been more thoroughly investigated by Saitis *et al.* [59] using spontaneous preference descriptions by experienced performers collected in a playing-based perceptual evaluation experiment. Upon ordering a set of different violins in terms of preference, players were asked to explain their choices via an open questionnaire. The constant comparison technique from grounded theory was employed to develop a classification scheme of concepts and the attributes that embody them. A quantitative analysis, based on the number of occurrences for each attribute and concept, provided a hierarchy of violin preference criteria/quality concepts: the response of the violin to the various techniques and musical intentions in direct association with the quantity and quality of the produced sound as well as the emotions and values of the player influence how the "good" violin is conceptualized. Depending on the musical aims of the violinist, priority between playability and sound quality varies. Besides, violin players use different verbalisations to describe instrument qualities they prefer from those they do not.

3.3 For investigating whether there is a superiority of the Old Italians

The effect on violins of ageing and playing was investigated as well in [44] with playing tests in addition to the listening tests described above (section 2.1). Each member of the listening panel in turn became a player and had to rate the instruments along the same six criteria used in the listening test, to which were added four other criteria more related to playing qualities. The results are similar to those obtained with the listening test: three years of regular playing has not made any statistically significant difference to the performance of one of the pair of violins. However, no statistically significant difference does not mean no difference: the difference may have been too small to be detected at a statistical level with a limited panel. As the number of players/listeners is actually not provided, no firm conclusions can be stated.

These preliminary results along Coggin's comment (see section 2.4) led the author and her collaborators to investigate whether there is actually a "true" superiority of the Old Italians among very experienced players: player's judgments about a Stradivari's sound may well be biased by the

violin's extraordinary monetary value and historical importance! In a first experiment conducted during the 2010 International Violin Competition of Indianapolis (IVCI), 21 experienced violinists were asked to compare three violins by Stradivari and Guarneri 'del Gesu' with three high quality new instruments [60]. The methodology was, for identical reasons, similar to the one presented in section 3.2: room with relatively dry acoustics, reduced lighting, dark goggles, participants' own bow when possible. In addition, to preserve double-blind conditions, the room was divided into two areas by a cloth screen and violins were passed from behind the screen to a researcher wearing goggles, who laid them on a bed in the order received. This study explored player preferences under two sets of conditions. Only the second set, the most interesting one, will be presented here. Designed to maximize ecological validity, it emulated the way players choose instruments at a violin shop, where they typically try a selection of instruments before selecting one to take home for further testing. All six test instruments were laid out in random order on the bed. Participants were then given 20 minutes to choose (i) the single instrument *they would most like to take home with them*, and (ii) the instruments they considered *best* and *worst* in each of four categories: *range of tone colors*, *estimated projection*, *playability*, and *response*. They were invited as well, at the end of the session, to guess the "making-school" of their take-home instruments - an indirect way of assessing their ability to distinguish new instruments from old. Just 8 of 21 subjects (38%) chose an old violin to take home. Given the small sample size, this disinclination toward the old cannot be confidently inferred to experienced violinists in general (CI [18%; 62%]). Still, the upper limit for the CI is not high; moreover, the fact that a new violin was chosen over examples by Stradivari and Guarneri 'del Gesu' by 13 experienced violinists (including two jury members of the IVCI, who compared two of the new violins favorably with their own Stradivari and Guarneri 'del Gesu' violins) stands as a bracing counterexample to conventional wisdom. Subjects seemed not to distinguish between new violins and old, but rather to choose instruments whose playing qualities best fit their individual tastes. Differences in taste among individual players, along with differences in playing qualities among individual instruments, appear more important than any general differences between new and old violins.

It is worth noting that these preferences were based solely on the experience of playing the instruments in a rather dry room. Though this raised numerous criticisms after the publication of the study, it was a deliberate choice (as violinists consider an acoustically dry room best for initial try-outs, so the direct sound from the instrument is not so much colored by room reflections) and none of the players gave any indication they felt the room was unsuitable for trying violins. However, the question of how well judgements made in one room carry over into another (in particular a concert hall) is an interesting one. In addition, though players do routinely estimate projection, they typically acknowledge (as many of the participants did) the need to re-test in a hall with trusted colleagues listening.

The fact that the three distinguished instrumentalists who publicly replied to the paper all raised the issue of projection suggests it is an important one to investigate. Therefore, a new experiment, which took place in Paris in 2012, was conducted to address both of these issues, among others. The data are still under analysis.

3.4 Discussion about the number of participants

With the exception of [32] and the studies presented in sections 3.2 and 3.3, the playing tests were generally conducted with only a few players (less than 10) if not only one or two. The reasons are quite obvious: the tests are thus much easier to run and the analysis of the data is facilitated as the inter-individual variabilities are smoothed (if not completely removed). However, the representativity of such a small number of players can be argued, especially if the player is one of the authors who knows the goals of the study, and one may wonder about the generalisation of the results to a larger sample of players. However, there is not necessarily a need for such generalisation. For instance, to know that a few players can distinguish between two factory instruments is enough to investigate the mechanical reason(s) for such a perceptual difference. Besides, one may want to check first on a very few players whether some relationships (between "subjective" and "objective" assessments) can be found before conducting a large scale experiment. And finally, one may not be interested in studying the perceptual assessments at an inter-individual level but rather at an intra-individual level: knowing the expertise and the experience of the player, his/her judgements made on different instruments in different conditions do not need to be potentially generalisable to be informative.

4. CONCLUSION

While research in musical acoustics originally focused primarily on instruments and their mechanical behavior, the musician has been more and more included in the last couple of decades, from a mechanical as well as perceptual point of view. Regarding the latter (the scope of this paper) it has indeed appeared important to understand and study objectively how musicians evaluate the quality of an instrument in order, for instance, to search for objective quality parameters and potentially offer improvements and new directions for the design and manufacturing of instruments. To this end, different examples were provided in this paper not just to summarize the latest developments in violin research but to illustrate the whole range of methodologies (beyond the crude classification between listening and playing tests) which are available and can apply to any instrument, as well as discuss their limitations and issues.

While it is natural for engineers and acousticians to describe with great care the physical measurements and the experimental devices used to carry them out, their perceptual (psychological) studies are often described much more succinctly. More specifically, there is often a lack of details on the experimental settings and procedures including the selection of the stimulations, the instructions given to

subjects, the exact wording of the questions or scales, the language used in verbal data, the type of records and of data analysis, the number of subjects, ... This lack of details applies as well to the characteristics of the subjects: age, gender, abilities and practices related to the object of concern (e.g. expertise, familiarity to the objects of concern and/or the tasks), inter-individual variability, ... Moreover, the problem of ecological validity is often not addressed properly, sometimes not even considered at all. It is however essential that the experimental protocol be designed so the participants react, to some extent, as if they were in their "natural", i.e. ordinary, situation of practice, as it is a necessary condition for the results collected in laboratory conditions to be generalised (at least to some extent) to an everyday situation. Researchers should be aware that ecological validity of the experimental setting is only ensured by first identifying the specific properties of mental cognitive representations and then defining the variables or parameters according to their relevance in psychological processing, as shown by Guastavino *et al.* [61] in the first study addressing ecological validity in the acoustics domain. There is of course no general recipe, as it depends on the purpose of the study and the type of people involved, except to investigate precisely the psychological processing as such from the subjects' points of view and not from an a priori conception, even if validated by physical knowledge. A collaboration of engineers with makers and/or musicians as well as psychologists is one of the keys to success.

Acknowledgments

The author would like to thank Joseph Curtin for fruitful discussions.

5. REFERENCES

- [1] C. Hutchins and V. Benade, *Research Papers in Violin Acoustics 1975-1993 with an introduction essay 350 years of violin research - Volume 1*. Acoustical Society of America, 1997.
- [2] F. Ablitzer, J.-P. Dalmont, and N. Dauchez, "Static model of a violin bow: Influence of camber and hair tension on mechanical behavior," *J. Acoust. Soc. Am.*, vol. 131, 2012.
- [3] K. Guettler and H. Thelin, "Bowed-string multiphonics analyzed by use of impulse response and the Poisson summation formula," *J. Acoust. Soc. Am.*, vol. 131, pp. 2903-2913, 2012.
- [4] G. E. Gough, "A finite element approach towards understanding violin structural modes," *J. Acoust. Soc. Am.*, vol. 127, p. 1791, 2010.
- [5] M. Demoucron, "Measuring bow force in bowed string performance: Theory and implementation of a bow force sensor," *Acustica - Acta Acustica*, vol. 95, pp. 718-732, 2009.
- [6] E. Schoonderwaldt, "The violinist's sound palette: Spectral centroid, pitch flattening and anomalous low frequencies," *Acustica - Acta Acustica*, vol. 95, pp. 901-914, 2009.
- [7] G. Bissinger, "Structural acoustics model of the violin radiation profile," *J. Acoust. Soc. Am.*, vol. 124, 2008.
- [8] J. Woodhouse, "On the "bridge hill" of the violin," *Acustica - Acta Acustica*, vol. 91, pp. 155-165, 2005.
- [9] J. Woodhouse and P. Galluzzo, "The bowed string as we know it today," *Acustica - Acta Acustica*, vol. 90, pp. 579-589, 2004.
- [10] J. Alonso Moral and E. Jansson, "Input admittance, eigenmodes and quality of violins," Department for Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden, Tech. Rep. STL-QPSR 2-3, 1982.
- [11] C. Hutchins, "A measurable controlling factor in the tone and playing qualities of violins," *Catgut Acoust. Soc. J. (Series II)*, vol. 1, pp. 10-15, 1989.
- [12] H. Dünwald, "Deduction of objective quality parameters on old and new violins," *Catgut Acoust. Soc. J. series II*, vol. 1, no. 7, pp. 1-5, 1991.
- [13] A. Langhoff, A. Farina, and L. Tronchin, "Comparison of violin impulse responses by listening to convoluted signals," in *Proc. of International Symposium on Musical Acoustics*, Paris, 1995.
- [14] G. Bissinger, "Structural acoustics of good and bad violins," *J. Acoust. Soc. Am.*, vol. 124, pp. 1764-1773, 2008.
- [15] H. Fletcher, E. Donnel Blackham, and R. Stratton, "Quality of piano tones," *J. Acoust. Soc. Am.*, vol. 34, pp. 749-761, 1962.
- [16] R. Pratt and J. Bowsher, "The subjective assessment of trombone quality," *J. Sound and Vibration*, vol. 57, pp. 425-435, 1978.
- [17] G. Plitnik and B. Lawson, "An investigation of correlation between geometry, acoustic variables, and psychoacoustic parameters for French horn mouthpieces," *J. Acoust. Soc. Am.*, vol. 106, pp. 1111-1125, 1999.
- [18] R. Caussé, N. Misdariis, and S. Thoréton, "Caractérisation objective de la qualité d'un instrument de musique," in *Proc. of Congrès Français d'acoustique*, Marseille, France, 1997.
- [19] E. Poirson, J.-F. Petiot, and J. Gilbert, "Study of the brightness of trumpet tones," *J. Acoust. Soc. Am.*, vol. 118, pp. 2656-2666, 2005.
- [20] —, "Integration of users perceptions in the design process: Application to musical instrument optimization," *J. Mechanical Design*, vol. 129, pp. 1-9, 2007.
- [21] L. Placido, A. Mamou-Mani, and D. Sharp, "Investigating perceptual differences between two trumpets of the same model type," *Applied Acoustics*, vol. 72, pp. 907-914, 2011.
- [22] G. Widholm, R. Linortner, W. Kausel, and M. Bertsch, "Silver, gold, platinum - and the sound of the flute," in *Proc. of International Symposium on Music Acoustics*, Perugia, Italy, 2001.
- [23] D. Sharp, H. Wright, and W. Ring, "An acoustical investigation into the effect of the crook profile on the sound produced by the bassoon," *Acustica - Acta Acustica*, vol. 89, pp. 137-144, 2003.
- [24] F. Pinard, B. Laine, and H. Vach, "Musical quality assessment of clarinet reeds using optical holography," *J. Acoust. Soc. Am.*, vol. 113, pp. 1736-1742, 2003.
- [25] B. Gazengel, J.-F. Petiot, and E. Brasseur, "Vers la définition d'indicateurs de qualités d'anches de saxophone," in *Proc. of Congrès Français d'acoustique*, Lyon, France, 2010.
- [26] B. Gazengel, J.-F. Petiot, and M. Šoltés, "Objective and subjective characterization of saxophone reeds," in *Proc. of Acoustics 2012*, Nantes, France, 2012.
- [27] J.-F. Petiot, P. Kersaudy, G. Scavone, S. McAdams, and B. Gazengel, "Objective and subjective characterization of saxophone reeds," in *Proc. of International Congress of Acoustics*, Montreal, Canada, 2013.
- [28] R. Caussé, B. Goepp, and B. Sluchin, "An investigation on "tonal" and "playability" of eight digeridoos, perceived by players," in *Proc. of International Symposium on Music Acoustics*, Nara, Japan, 2004.

- [29] J. Smith, G. Rey, P. Dickens, N. Fletcher, L. Hollenberg, and J. Wolfe, "Vocal tract resonances and the sound of the Australian didjeridu (yidaki). III. Determinants of playing quality," *J. Acoust. Soc. Am.*, vol. 121, pp. 547–558, 2007.
- [30] J.-F. Petiot and R. Caussé, "Perceptual differences between cellos: a subjective/objective study," in *Proc. of International Symposium on Music Acoustics*, Barcelona, Spain, 2007.
- [31] A. Mamou-Mani, D. Sharp, T. Meurisse, and W. Ring, "Investigating the consistency of woodwind instrument manufacturing by comparing five nominally identical oboes," *J. Acoust. Soc. Am.*, vol. 131, pp. 728–736, 2012.
- [32] A. Paté, J.-L. L. Carrou, B. Navarret, D. Dubois, and B. Fabre, "A vibro-acoustical and perceptive study of the neck-to-body junction of a solid-body electric guitar," in *Proc. of Acoustics 2012*, Nantes, France, 2012.
- [33] W.H. Hill, A.F. Hill, and A.E. Hill, *Antonio Stradivari, His Life and Work*. New York: Dover Publications, 1902.
- [34] J. Schelleng, "Acoustical effects of violin varnish," *J. Acoust. Soc. Am.*, vol. 44, pp. 1175–1183, 1968.
- [35] L. Burckle and H. Grissino-Mayer, "Stradivari, violins, tree rings, and the Maunders Minimum: a hypothesis," *Dendrochronologia*, vol. 21, pp. 41–45, 2003.
- [36] B. Stoel and T. Borman, "A comparison of wood density between classical Cremonese and modern violins," *PLoS ONE*, vol. 3, p. e2554, 2008.
- [37] C. Barlow, P. Edwards, G. Millward, R. Raphael, and D. Rubio, "Wood treatment used in Cremonese instruments," *Nature*, vol. 332, p. 313, 1988.
- [38] J. Nagyvary, J. DiVerdi, N. Owen, and H. D. Tolley, "Wood used by Stradivari and Guarneri," *Nature*, vol. 444, p. 565, 2006.
- [39] C. Hutchins, A. Hopping, and F. Saunders, "Subharmonics and plate tap tones in violin acoustics," *J. Acoust. Soc. Am.*, vol. 32, pp. 1443–1449, 1960.
- [40] G. Dubourg, *The Violin: Some Account of that Leading Instrument and Its Most Eminent Professors, from Its Earliest Date to the Present Time : with Hints to Amateurs, Anecdotes, Etc.* Robert Cox and Co., London, England, 1852.
- [41] J. Marchese, *The Violin Maker: A Search for the Secrets of Craftsmanship, Sound, and Stradivari*. HarperCollins Publishers, New York, USA, 2007.
- [42] A. Coggins, "Blind testing," *The Strad Magazine*, pp. 52–55, February 2007.
- [43] F. Schwarze, M. Spycher, and S. Fink, "Superior wood for violins wood decay fungi as a substitute for cold climate," *New Phytologist*, vol. 179, pp. 1095–1104, 2008.
- [44] R. Inta, J. Smith, and J. Wolfe, "Measurement of the effect on violins of ageing and playing," *Acoustics Australia*, vol. 1, pp. 25–29, 2005.
- [45] C. Fritz, I. Cross, E. Smith, K. Weaver, U. Petersen, J. Woodhouse, and B. Moore, "Perceptual correlates of violin acoustics," in *Proc. of 9th Int. Conf. on Music Perception and Cognition*, Bologna, Italy, 2006.
- [46] C. Fritz, I. Cross, J. Woodhouse, and B. Moore, "Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin," *J. Acoust. Soc. Am.*, vol. 122, pp. 3640–3650, 2007.
- [47] C. Fritz, J. Woodhouse, F.-H. Cheng, I. Cross, B. Moore, and A. Blackwell, "Perceptual studies of violin body damping and vibrato," *J. Acoust. Soc. Am.*, vol. 127, pp. 513–524, 2010.
- [48] C. Fritz, A. Blackwell, I. Cross, J. Woodhouse, and B. Moore, "Exploring violin sound quality: Investigating english timbre descriptors and correlating resynthesized acoustical modifications with perceptual properties," *J. Acoust. Soc. Am.*, vol. 131, pp. 783–794, 2012.
- [49] J. Bensa, D. Dubois, R. Kronland-Martinet, and S. Ystad, *Computer Music Modeling and Retrieval, Lecture Notes in Computer Science, Volume 3310*. Springer Berlin Heidelberg, 2005, ch. Perceptive and Cognitive Evaluation of a Piano Synthesis Model.
- [50] C. Fritz, A. Muslewski, and D. Dubois, "A situated and cognitive approach of violin quality," in *20th International Symposium on Music Acoustics*, Sydney, Australia, 2010.
- [51] R. Pratt and J. Bowsher, "The objective assessment of trombone quality," *J. Sound and Vibration*, vol. 65, pp. 521–547, 1979.
- [52] R. Caussé, J. Maignet, C. Dichtel, and J. Bensoam, "Study of violin bow quality," in *Proc. of International Symposium on Music Acoustics*, Perugia, Italy, 2001.
- [53] G. Weinreich, "What science knows about violins and what it does not know," *Am. J. Phys.*, vol. 61, p. 1067, 1993.
- [54] C. Saitis, B. Giordano, C. Fritz, and G. P. Scavone, "Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players," *J. Acoust. Soc. Am.*, vol. 132, pp. 4002–4012, 2012.
- [55] C. Saitis, G. Scavone, C. Fritz, and B. Giordano, "Reliability and reproducibility in violin preference judgements by experienced players," in *Proc. of Stockholm Music Acoustics Conference*, Stockholm, Sweden, 2013.
- [56] G. Bissinger and F. Gearhart, "A standardized qualitative violin evaluation procedure," *J. Catgut Acoust. Soc. Series II*, vol. 3, no. 6, pp. 44–45, 1998.
- [57] I. Wollman, J. P. C. Fritz, and S. McAdams, "Investigating the role of auditory and tactile modalities in violin quality evaluation," *Music Perception*, Soumis.
- [58] C. Fritz, "The violin: Perceptual studies and acoustical correlates," *J. Acoust. Soc. Am.*, vol. 127, p. 1791, 2010.
- [59] C. Saitis, C. Fritz, C. Guastavino, and G. Scavone, "Conceptualizing the "good" violin in preference descriptions by experienced players," in *Proc. of Stockholm Music Acoustics Conference*, Stockholm, Sweden, 2013.
- [60] C. Fritz, J. Curtin, J. Poitevineau, P. Morrel-Samuels, and F.-C. Tao, "Player preferences among new and old violins," *Proc. Nat. Ac. Sci.*, vol. 109, pp. 760–763, 2012.
- [61] C. Guastavino, B. Katz, J.-D. Polack, D. Levitin, and D. Dubois, "Ecological validity of soundscape reproduction," *Acustica - Acta Acustica*, vol. 91, pp. 333–341, 2005.

EFFECT OF HOLDING THE VIOLIN AND SOUND-POST REMOVAL ON VIOLIN RADIATION VIA THE DYNAMIC FILTER MODEL

George Bissinger

Professor of Physics emeritus
East Carolina University
bissinger@ecu.edu

ABSTRACT

The dynamic filter model, a deterministic-statistical, network-structural acoustics parameterization of violin radiativity, was developed to simulate violin radiativity *trends* due to plate and bridge tuning for a violin with a soundpost, suspended free-free. The model formalism is quite general however and this generality was tested by successful application to two decades-old experiments: 1) vibration measurements on a violin held (as in playing) vs. free-free and 2) a soundpost removal experiment with pre- and post- removal: a) modal analyses, b) boundary element method (no *f*-hole) radiation efficiency calculations) and c) room-averaged acoustic measurements.

1. INTRODUCTION

The dynamic filter model (DFM) was developed to provide a simplified parameterization of the violin's radiativity "profile", the averaged-over-sphere, mean-square radiativity $\langle R_{\text{exp}}^2(\omega) \rangle$ [1]. The DFM splits the violin's radiativity profile into two regions: 1) *deterministic* in the open string region for the three always strongly radiating low frequency signature modes, including excitation of the two lowest cavity modes and 2) *statistical* for the higher frequency region where a generalized statistical energy analysis (SEA) framework was adopted to include such structural acoustics parameters as critical frequency, radiation efficiency, radiation and total damping and modal density.

NOMENCLATURE

$R_{\text{eff}}(\omega)$ – radiation efficiency
 $R(\omega)$ – radiativity = pressure/force (Pa/N)
 $Y(\omega)$ – mobility = velocity/force (m/s/N)
 $n(f)$ – modal density (number of modes per 250 Hz band)
 $\Phi(f)$ – shape function, radiativity profile parameterization
 $S(f)$ – scaling function for bridge rocking frequency changes
 ζ_{tot} – total damping (percent critical), sum of ζ_{rad} (radiation), ζ_{int} (internal), ζ_{fix} (support fixture) damping
 F_{RAD} – fraction of vibrational energy r adiated
 $Y_{\text{stat}}(f)$ – distributed-excitation statistical mobility

Copyright: © 2013 George Bissinger. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Although the DFM was developed for the specific case of a violin in free-free suspension (using thin elastics) in an anechoic chamber to simulate the general effects of plate and bridge tuning from 0.2-6 kHz [1], its formalism is rather general and thus testable for violins in other states *if* sufficient additional information is available. Here the formalism is applied to two important earlier experiments that were not part of the DFM foundational database to provide some assessment of its overall generality and reliability, viz., radiativity *changes* associated with: a) Marshall's 1986 experiment on the effect of the violinist's holding the violin [2] and b) Bissinger's 1995 soundpost removal experiment [3].

2. DYNAMIC FILTER MODEL

To give some necessary basic physical insights a very condensed exposition of the dynamic filter model is presented here (cf. ref. 1 for complete account). $\langle R_{\text{exp}}^2(\omega) \rangle$ breaks naturally into two independent regions [4]: 1) from ~0.2 to ~0.6 kHz a "deterministic" region where individual modes are easily traced across violins that includes the signature modes A0 (the Helmholtz-resonator-like cavity resonance), A1 (the first longitudinal cavity resonance and a *sometimes* strong radiator), CBR (a strong mechanical, weakly radiating mode neglected in the model), and the 1st corpus bending modes B1⁻ and B1⁺, and 2) from ~0.8 to ~6 kHz a "statistical" region where violin characterizations are best done using statistical measures such as averaged-over-band mode properties. A crucial splice point for the deterministic-statistical regions occurred near 630 Hz in a broad radiativity trough where all violin quality classes shared a common experimental value $\langle R_{\text{exp}} \rangle \approx 0.19$ Pa/N at ~630 Hz ($r = 1.2$ m in anechoic chamber, free-free suspension).

In the deterministic region Weinreich, Holmes, and Mellody created a theoretical coupling model for A0 and B1 modes that agreed well with their experimental mode "veering" results [5]. Fahy and Gardonio have noted that for a thin shell bounding a fluid region (a closed system) the action of the fluid is generally to couple the *in-vacuo* structural modes, although they also note that such a closed system can have orthogonal natural modes [6]. The violin however is not a closed system, it is a ported system, and the large volume changes associated with the B1 modes combined with their relatively large damping

(prior experiment provided reliable total damping as a function of mode frequency) leads to significant f -hole volume flows even at the A0 frequency, *and* at the anti-nodes for A0.

Considerable additional experimental evidence supports the hypothesis that the B1 modes are chiefly responsible for A0-A1 excitation [1], hence in the DFM B1 corpus modes were treated as the “drivers” for A0 and A1 in a wall-driven, dual-Helmholtz resonator network model to compute radiativity for the strongly radiating signature modes A0, B1⁻ and B1⁺ [1]. This dual-Helmholtz resonator model requires signature mode frequencies and reliable total dampings as well as B1 mobility amplitudes as basic parameters.

In the statistical region the fundamental DFM premise is that sound radiation can be expressed as the product of: 1) how vigorously the violin vibrates, parameterized by $\langle Y_{\text{exp}}^2(\omega) \rangle$, the mean-square, surface-normal mobility averaged over the corpus, *and* 2) how effectively vibrational motions are converted to radiation, parameterized by the experimental radiation efficiency $R_{\text{eff}}(\omega)$, computed from $\langle R_{\text{exp}}^2(\omega) \rangle / \langle Y_{\text{exp}}^2(\omega) \rangle$ and various system constants. Expressed first formally as an identity and then with the $R_{\text{eff}}(\omega)$ substitution,

$$\begin{aligned} \langle R_{\text{exp}}^2(\omega) \rangle &= \langle Y_{\text{exp}}^2(\omega) \rangle \cdot \{ \langle R_{\text{exp}}^2(\omega) \rangle / \langle Y_{\text{exp}}^2(\omega) \rangle \} \\ &\propto \langle Y_{\text{exp}}^2(\omega) \rangle \cdot R_{\text{eff}}(\omega) \end{aligned} \quad (1)$$

Because the DFM assumes excitation through the bridge, a “tuned” bridge filter effect is implicit in the model.

To incorporate variable bridge tuning a statistical, *distributed-over-corpus*-excitation mobility [6], $Y_{\text{stat}}^2 \propto n(f)/(f \zeta_{\text{tot}}(f) M^2)$ – where $n(f)$ is the modal density (represented by a 6th order polynomial trendline equation), the total damping $\zeta_{\text{tot}}(f)$ {sum of radiation, internal (heat) and support fixture dampings: $\zeta_{\text{tot}}(f) = \zeta_{\text{rad}}(f) + \zeta_{\text{int}}(f) + \zeta_{\text{fix}}(f)$ }, well represented up to 4 kHz by a power law $\zeta_{\text{tot}}(f) \propto f^{-0.34}$ trendline equation (for free-free supports $\zeta_{\text{fix}}(f) \approx 0$), and M the total mass ≈ 0.4 kg – was used with $\langle Y_{\text{exp}}^2 \rangle$ to define the “shape” function $\Phi = \langle Y_{\text{exp}}^2 \rangle / Y_{\text{stat}}^2$. Φ displays three broad structures simply characterized in the DFM by three resonance functions: two fixed frequency-damping (~ 1 kHz near the expected “ring” frequency and where $n(f)$ is a maximum and ~ 2.3 kHz, the BH structure) as they did not appear to change as plate and bridge tunings were varied and, most importantly, a peak that followed bridge tuning changes in f_{rock} , the rocking frequency of the bridge (measured with feet clamped in a vise).

The experimental radiation damping $\zeta_{\text{rad}}(f)$ computed from R_{eff} led to the fraction-of-vibrational-energy-radiated parameter, $F_{\text{RAD}} = \zeta_{\text{rad}}(f)/\zeta_{\text{tot}}(f)$. After collecting terms the violin radiativity “profile” in the statistical region for *free-free* suspension becomes simply

$$\langle R_{\text{DF}}^2(f) \rangle \propto \frac{\Phi(f) \cdot n(f) \cdot S(f)}{M} F_{\text{RAD}}(f) \quad (2)$$

where the scaling function $S(f)$ was used to incorporate *changing* bridge f_{rock} . $S(f)$ would be expected to be ap-

proximately constant in the statistical region for either holding the violin or soundpost removal.

3. DFM IMPLEMENTATION

3.1 Holding the violin

In 1985 Marshall published the landmark first modal analysis of a violin in free-free suspension [7], followed in 1986 with vibration measurements on a violin held by a violinist [2]. However no acoustic measurements were made and thus only vibration data is available. DFM estimates of *held*-violin radiativity, while thus limited to the *relative* comparison with the free-free case, are quite sufficient in this instance since the free-free state is well parameterized [1]. When the violinist picks up the violin to play, he/she becomes the support fixture and the hand and chin-shoulder contact significantly increases total damping, i.e., free-free $\zeta_{\text{fix}}(f) \approx 0 \rightarrow \text{held } \zeta_{\text{fix}}(f) \neq 0$, more so at lower frequencies (cf. Marshall’s Figs. 2 and 3). Holding the violin, while it significantly increased $\zeta_{\text{fix}}(f)$, hardly affected mode frequencies. The DFM analysis that follows used only Marshall’s summary frequency and total damping results for free vs. held violins tabulated in his Table II.

Marshall’s experimental results lead immediately to some simplifying assumptions: 1) changing mode damping but not frequency implies that mode shapes and thus mode R_{eff} *and* ζ_{rad} do **not** change significantly, 2) the A0-B1⁻ frequency difference remains constant, 3) the total damping power law exponent should increase due to the lower modes having relatively larger damping, and 4) in eq. 2 Φ , $n(f)$, $S(f)$ and M remain approximately constant. (Note however that there appears to be a slight drop off in Φ as frequency decreases from f_{rock} to BH to “ring” regions that is neglected here.)

The approximate constancy of many DFM parameters for this violin argue for taking the ratio of *held*/DF radiativity profiles to cancel these parameters out, leaving (for simplicity all frequency dependences now understood)

$$\begin{aligned} \langle R_{\text{held}}^2 \rangle / \langle R_{\text{DF}}^2 \rangle &\approx F_{\text{RADheld}} / F_{\text{RADfree}} \\ &\approx \zeta_{\text{totfree}} / \zeta_{\text{totheld}} \end{aligned} \quad (3)$$

The statistical region DFM now becomes simply the ratio of total dampings $\zeta_{\text{totfree}}/\zeta_{\text{totheld}}$. Power law fits, $\zeta_{\text{tot}} \propto A f^x$, to the Marshall free-free data returned $x = -0.36$, very close to the $x = -0.34$ exponent from the 14-violin ensemble data set fit used in the original DFM parameterization [1]. For the *held* data $x = -0.88$. The power law trendlines are shown in Fig. 1. Thus using only eq. 3 the DFM predicts that holding the violin diminishes low frequency response relative to high and that near 4 kHz the responses reach very similar values a conclusion that agrees well with Marshall’s response data in his Figs. 2 and especially 3 (which extends to 4 kHz).

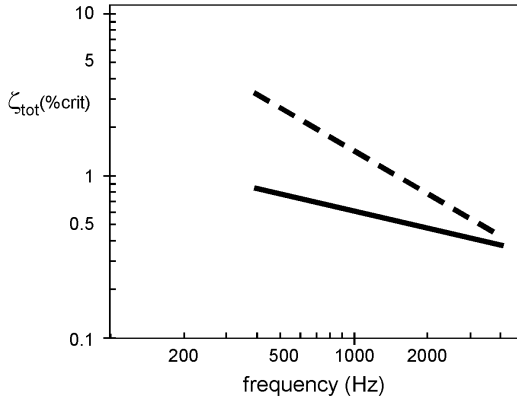


Figure 1 - Power-law total damping trendlines for Marshall's held (---) and free (—) violin. (Ratio of these trendlines is ratio of DFM radiativity profiles, i.e., eq. 3.)

Combining the statistical region simulation – based on the *ratio* of total damping trendlines, viz., $\langle R_{\text{held}}^2 \rangle / \langle R_{\text{free}}^2 \rangle \approx 85.9 \cdot f^{-0.52}$ – with the deterministic region DFM analysis that follows provides a complete simulation from 0.2-4 kHz that can be compared to Marshall's top plate response data in his Fig. 3.

3.1.1 A0 radiativity

In the deterministic region A0 radiativity $\langle R_{A0} \rangle$ (averaged over a sphere) must be accounted for by the wall-driven, dual-Helmholtz resonator model. For comparison purposes $\langle R_{A0} \rangle$ is here assumed to follow the *relative magnitude of induced corpus motion* at A0 frequency in Marshall's Fig. 2 as the damping fit results are similar for both. The DFM for the free-free violins of ref. 1 stitched the two independent regions together at 630 Hz where all violin quality classes had a common radiativity. Such a procedure is not employed here as it is not known whether the *held* data share the same empirical fortune. Hence an alternative approach – relying on the DFM hypothesis that the B1 modes are the A0 “pumps” – extends the statistical region corpus-mode-only simulation down to the B1 region to estimate relative *held* violin B1 strength and then scales A0 from this B1 strength estimate.

In the DFM $\langle R_{A0} \rangle$ depends on: 1) the *strengths of* $B1^-$ and $B1^+$ ($B1^-$ however being dominant), with larger values leading to increased $\langle R_{A0} \rangle$, 2) *total damping*, with larger damping for a fixed B1 peak magnitude leading to increased $\langle R_{A0} \rangle$ and 3) the *frequency separation* Δf between A0 and each B1 mode, with smaller values leading to larger $\langle R_{A0} \rangle$. Experimental values taken from Marshall's Figure 2 plot of free vs. *held* accelerances show $B1^-$ at ~456 Hz drops by a factor of ~0.3, $B1^+$ at ~543 Hz drops by a factor of ~0.2 while A0 drops only by a factor of ~0.5.

Marshall's free \rightarrow held signature mode total damping for: $B1^-$ ζ_{tot} goes from 1.28 \rightarrow 1.99 %crit (Q from 39 to 25), $B1^+$ goes from 0.79 \rightarrow 3.99 %crit (Q from 63 to 12), and A0 goes from 2.16 \rightarrow 2.39 %crit (Q from 23 to 21), a ~10% difference for the latter which is considered negligible for such fits. (Marshall's ζ_{tot} fit values had unreported fit errors; fit values depend on fit band limits, the number of peaks in a band, the number of lines in the

peak, i.e. the frequency resolution, and the presence of nearby out-of-band peaks.)

In the dual-Helmholtz resonator model maintaining B1 mode magnitudes while *changing only damping* increases *held* A0 radiativity by ~50% or ~1.5x original. However $B1^-$, the major influence on A0 radiativity in this model, has dropped to ~30% of its original strength, hence A0 radiativity should nominally end up at $1.5 \times 0.3 \approx 0.45$ or 45% of original A0 strength, close to the observed drop of ~50% noted above.

Since $\langle R_{A0} \rangle$ in the dual-Helmholtz resonator model requires B1 radiativities, which we did not have, an effective alternative was to extend the statistical region predictions (under our assumption of constant radiation efficiency for each mode) down into the B1 region. Hence deterministic DFM predictions of B1 average radiativity to compute A0 were based on statistical model estimates, eq. 3's $\zeta_{\text{totfree}}/\zeta_{\text{totheld}}$ extrapolated down to 500 Hz.

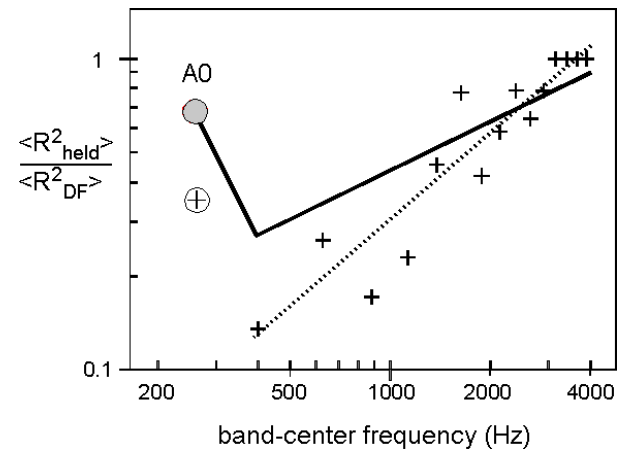


Figure 2 – Dynamic filter model predictions for held violin radiativity relative to free (—) including A0 (○; computed *relative* to DFM B1) compared to band-average experimental ratios extracted from Marshall's Fig. 3 for corpus modes (+, with empirical power law trendline (•••) and A0 (⊕).

For comparison with experiment a rough estimate of experimental radiativity ratios to 4 kHz was taken from the ratio of Marshall's accelerances - squared - at each peak in the vibration spectrum (read from his Figure 3 for top plate response in the upper bout). These experimental estimates were then averaged over 250 Hz intervals for comparison with the statistical DFM predictions extended from 0.4 to 4 kHz. (Important note on all these band averages: the lowest band was always for a relatively narrow region around A0 by itself, the next higher band ran from 300-499 Hz (A0 excluded, band-center at 400 Hz); all successive bands 250 Hz wide.)

Joining the statistical and deterministic regions together leads to the full range DFM predictions shown in Figure 2 compared with the estimated experimental radiativity ratios. The DFM predictions in Figure 2, generated entirely from the ratio of total damping trendlines and the wall-driven, dual-Helmholtz resonator model, reliably follow the relative trends of Marshall's vibration data as seen in his Figs. 2 and 3. Incorporation of the

aforementioned slight drop off in the gatekeeper’s shape parameter Φ as frequency decreased would improve agreement with experiment, but the general trends are well handled by just the egress filter.

Of course one could have predicted the observed radiativity trends directly from Marshall’s figures under the assumption of unchanged R_{eff} values, with one important exception, $\langle R_{A0} \rangle$. While holding the violin diminishes low frequency radiation relative to high due to the relatively larger “support fixture” damping at the lower frequencies, *held* $\langle R_{A0} \rangle$ lies relatively higher than the B1 strengths would suggest because it benefits from the significant increase in B1 total damping.

3.2 Soundpost removal

The removal of the small soundpost (mass $\sim 0.5\%$ of violin total mass) from a region where boundary conditions are crucial greatly debilitates the sound and also substantially diminishes radiation at lower frequencies [3], similar to the effect of holding, but again A0 behaves differently. The DFM simulation of soundpost removal effects on the radiativity profile however is definitely more problematical as it involves a *change in the actual physical system*, whereas holding the violin just increased a heretofore negligible frequency-dependent energy loss path. Thus soundpost removal cannot be approached in the same straightforward manner used in the holding analysis.

The transition from the traditional violin to the violin without a soundpost in the context of the DFM requires somewhat different assumptions than the holding case, and more extensive experimental guidance. Fortunately a great deal of additional experimental information from this 1995 experiment is available to help in the analysis, viz., pre- and post- soundpost removal: 1) *modal analyses* of the violin up to 2 kHz, 2) *boundary element method (BEM) R_{eff} calculations* (cautionary note – these R_{eff} calculations were done for a violin *without f-holes*, a significant matter for the B1 modes as a major part of their radiation is due to large volume flows through the *f-holes* [8]), and 3) *room-averaged, slide-tone acoustic measurements* [3]. (The foam supports used in this experiment added appreciable – but not dominating – support fixture damping, hence $\zeta_{\text{fix}}(f) \neq 0$, but with an unknown frequency dependence assumed common to both measurement states).

3.2.1 Room-averaged, slide-tone analysis

Although the single-microphone, room-averaged, multi-slide-tone (glissandi) acoustic measurements presented in Figure 3 do not share the clarity, precision and quality of the averaged-over-sphere anechoic chamber radiativity measurements, the general radiation trends associated with soundpost removal are apparent: 1) in the deterministic region the violin with a soundpost shows considerably stronger radiation overall, 2) in the statistical region quite similar radiation levels are observed, and 3) A0 drops in frequency and strength.

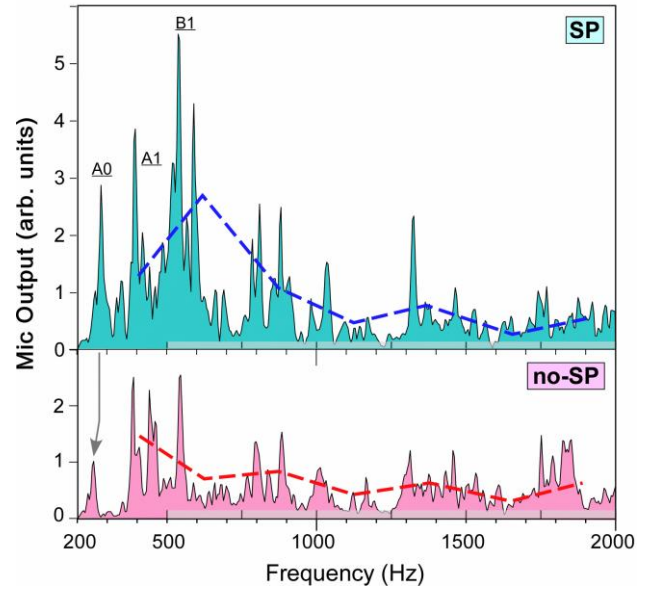


Figure 3 – Room-averaged slide-tone microphone output for a violin with soundpost (top) and without (bottom). (Superimposed 250 Hz band averages (— —) also shown)

3.2.2 Deterministic region

The deterministic region requires additional interpretation due to the significant mode shape changes accompanying soundpost removal, most notably for the B1 modes, the A0 “drivers” in the DFM. (Mode changes in the deterministic region can also possibly affect the observed A0-B1 coupling [5].) The pre-/post- removal modal analyses revealed only one B1⁺-like no-SP mode near 497 Hz (near the frequency of the SP-state B1⁺) that showed the large volume changes *and* rigid-body, bridge-pivot action about an approx. fixed “soundpost foot”, common to both B1 modes in the SP state. Thus the no-SP state has: 1) lost one of the large volume change A0 “driver” modes, and 2) suffered an ~ 30 Hz increase in the frequency difference Δf between A0 and the B1⁺-like mode. Surprisingly, observed A0-induced wall motions are now *contra*-indicative of A0 radiative strength.

These experimental findings have significant consequences in the dual-Helmholtz resonator model for $\langle R_{A0} \rangle$, which depends on: 1) B1 magnitudes, 2) having two (SP) rather than one (no-SP B1⁺-like”) large volume change corpus modes, 3) B1 and A0 total damping values, and 4) the A0-B1 Δf .

Prior DFM simulations for the deterministic region that incorporated all these SP \rightarrow no-SP changes predicted that no-SP $\langle R_{A0} \rangle$ was significantly weaker than SP $\langle R_{A0} \rangle$ [1], quite consistent with earlier single microphone measurements by Meinel [9] of this near isotropic radiation and with the room-average slide-tone results shown in Figure 3.

3.2.3 Statistical region

Vibrational data for the SP/no-SP states show a similar number of peaks, clumped similarly, similar in frequency, and amplitude [3] so that the modal density $n(f)$ effectively remains unchanged, as does total mass M . Total damping trendlines appeared to fall off slightly more slowly

for the no-SP state with $\zeta_{\text{tot}} = 19.2 \cdot f^{-0.50}$ than the SP $\zeta_{\text{tot}} = 32.9 \cdot f^{-0.58}$, although this is hardly an important difference given the data scatter, and the fact that total damping now includes added foam support fixture damping.

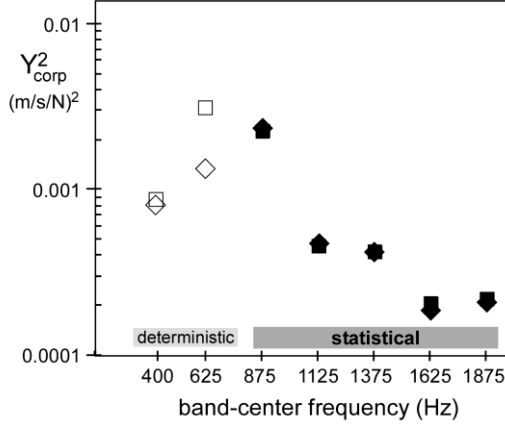


Figure 4 – Corpus band-average $Y_{\text{corp}2}$ for no-SP (■) and SP (◆).

Again subsuming individual mode properties into 250 Hz band-averages, the corpus-mode-only, mean-squared band-average mobilities $\langle Y_{\text{corp}}^2 \rangle$ plotted in Figure 4 are very close in value, hence $Y_{\text{no-SP}}^2/Y_{\text{SP}}^2 \approx 1$ in the statistical region. This similarity in statistical region structure implies that the statistical region shape function Φ in eq. 2, along with $n(f)$ and M , are approximately constant and thus should drop out in the ratio. Finally, modal analysis results in the statistical region up to 2 kHz – unlike in the deterministic region – show quite similar bridge motions for both states, hence $S(f)$ will also be considered approximately constant, to be cancelled out again in the ratio.

With these assumptions the DFM equation for soundpost removal in the statistical region thus becomes

$$\langle R_{\text{no-SP}}^2 \rangle / \langle R_{\text{SP}}^2 \rangle \approx F_{\text{RADno-SP}} / F_{\text{RADSP}} \quad (4)$$

where the total dampings for each state are represented with each state's simple trendline equations. BEM R_{eff} values for modes up to 2 kHz were used to compute $\zeta_{\text{rad}}(f)$ for each mode and these values were then assigned appropriately to respective modes across the entire 2 kHz range for each state. $F_{\text{RAD}}(f)$ was then computed for each mode, followed by 250 Hz band $F_{\text{RAD}}(f)$ averages computed over the statistical region and then the ratio was taken for the DFM predictions.

Note that the no f -hole BEM R_{eff} calculations used to compute radiation damping in the F_{RAD} ratio completely disregard the dominant f -hole radiativity path for the B1 modes. However in the statistical region they retain substantially greater validity as higher surface/ f -hole radiation ratios, generally higher R_{eff} due to proximity to the critical frequency, mobility falling off with frequency all conspire in the no-SP/SP ratio to minimize the effect of neglected f -hole radiation [8].

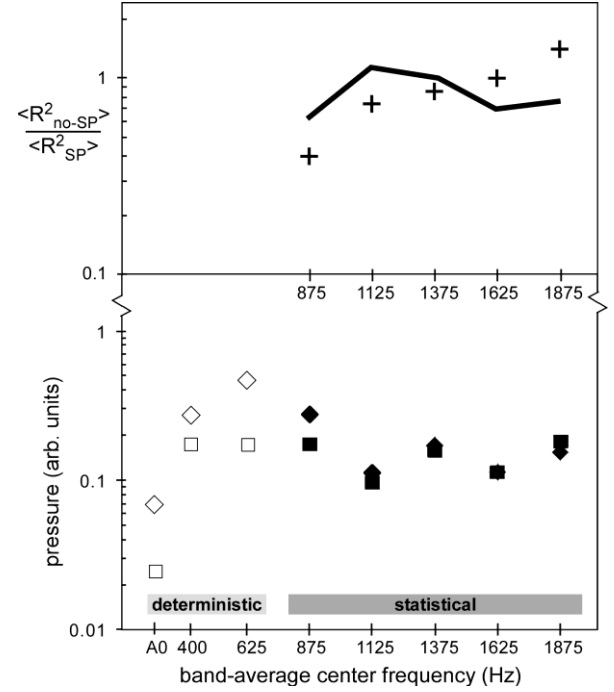


Figure 5 – 250 Hz-band (except A0 and 400 Hz): (bottom) slide-tone, room-average pressures $\langle P \rangle$ for violin with (◆) and without (■) soundpost; (top) statistical region dynamic filter model predictions (—) from F_{RAD} ratio vs. experimental no-SP/SP ratio of $\langle P \rangle^2$ (+).

Band-averages of the room-average, slide-tone acoustic pressure measurements shown in Figure 3 for the SP and no-SP states are shown in Figure 5 (bottom). In the statistical region the SP and no-SP pressures are generally quite similar, especially so in the four highest bands, whereas the deterministic region uniformly shows considerably stronger SP radiation. Squaring these experimental pressures and taking the no-SP/SP ratio gives a reasonable approximation in the statistical region to an experimental radiativity profile for comparison with the DFM predictions as shown in Figure 5 (top).

In the signature mode region peak frequency and mode character changes along with differing SP/no-SP no- f -hole BEM R_{eff} values precludes a reliable estimate of band-average B1 radiativity and thus $\langle R_{\text{A0}} \rangle$ radiativity specific to each case.

A simple interpretation of these DFM predictions is that soundpost removal has little effect on the radiativity profile of the violin in the statistical region, in strong contrast to what the model predicts in the deterministic region for A0, where both A0 and B1(region) radiativities decline significantly.

4. CONCLUSIONS

Overall the remarkably simple dynamic filter model expressions for the statistical region's relative free-free \rightarrow held transition, $\zeta_{\text{tofree}}/\zeta_{\text{toheld}}$, and the SP \rightarrow no-SP expression, $F_{\text{RADno-SP}}/F_{\text{RADSP}}$, give a very good idea of how the violin radiativity profile changes when it is held or undergoes soundpost removal, and its expression as a ratio allows it to be incorporated as an additional “filter” in

computing the effect of holding on the sound of a violin with certain plate-bridge tunings.

In the deterministic region the prediction of enhanced A0 radiation *relative to the B1 modes* upon holding is quite consistent with Marshall's observed increased A0-induced vibration levels *relative to the B1s*. The prediction of significantly decreased A0 radiativity upon soundpost removal on the other hand is fully in agreement with experimental observations as well.

The successful application of the dynamic filter model conceptual framework to these two disparate, decades-old experiments provides a validation of the generality of this model for simulating violin radiativity. This simplified model of violin radiation can now be used to simulate the individual, unrelated effects of plate tuning, bridge tuning, holding and soundpost removal on violin radiativity while providing a suitable framework for combining all these into one overall simulation.

5. REFERENCES

- [1] G. Bissinger, "Parametric plate-bridge dynamic filter model of violin radiativity", J. Acoust. Soc. Am. **132**, 465–476 (2012).
- [2] K.D. Marshall, "The musician and the vibrational behavior of the violin", Catgut Acoust. Soc. J., **45**, 28-33(1986).
- [3] G. Bissinger, "Some mechanical and acoustical consequences of the violin soundpost," J. Acoust. Soc. Am. **97**, 3154-3164(1995).
- [4] L. Cremer, *The Physics of the Violin*, (MIT Press, Cambridge, 1984), ch. 10,12.
- [5] G. Weinreich, C. Holmes, M. Mellody, "Air-wood coupling and the swiss-cheese violin", J. Acoust. Soc. Am. **108**, 2389-2402(2000).
- [6] F. Fahy and P. Gardonio, *Sound and Structural Vibration: Radiation, Transmission and Response*, 2nd ed. (Academic Press, 2007, New York).
- [7] K.D. Marshall, "Modal analysis of a violin", J. Acoust. Soc. Am. **77**, 695-709 (1985).
- [8] G. Bissinger, E.G. Williams and N. Valdivia, "Violin *f*-hole contribution to far-field radiation via patch near-field acoustical holography", J. Acoust. Soc. Am. **121**, 3899–3906(2007)
- [9] H. Meinel "On the frequency curves of violins," Akust. Z. **2**, 22-33 (1937).

THE ACOUSTICS OF THE HARDANGER FIDDLE

Anders Buen

Brekke & Strand Akustikk AS, Pb 1024

Hoff, NO-0218 Oslo, Norway

anb@brekkestrand.no

ABSTRACT

The Hardanger fiddle (HF) is a highly decorated, baroque-like Norwegian folk music instrument with four or five sympathetic strings. Compared to a violin, it has shorter and lighter gut, G, D and A, and a E steel string, a flatter bridge and fingerboard, longer f-holes and the top has a flatter cross arch, mainly between and above the f-holes. Acoustically it is closely related to the violin. Its “ringing qualities” relate it to other bowed string instruments with free vibrating string designs, like the Swedish folk *drone fiddle* and the *nyckelharpa*, and also to the more “distant relative” the *Viola d’amore*.

Acoustic properties and the construction of the HF are compared to the violin, showing close similarities in the lower frequency range. The bridge design and the tonal ideals are different, especially for the higher frequencies. A Hardanger fiddle will in general be a little quieter than a violin, but may sound more “intense”. Scordatura is often used and the HF is traditionally played solo. So the tuning of the A-string can be any frequency between A₄-D₅ (440 Hz - 587 Hz), with B₄ (494 Hz) being the most typical pitch.

1. INTRODUCTION

The investigations discussed in this paper were undertaken in the hope of gaining information on the acoustics of the Hardanger fiddle (HF). The design of the HF, with its sympathetic strings, differs from the violin.

The rise- and decay time effects from these on the bowed notes were investigated by Michaelsen [1]. Trueman later investigated the radiation properties as part of a sound synthetisation project [2].

1.1 Literature on the HF

In the last 20 years a registration project on HF’s made prior to 1900 has taken place. A database over registered old instruments was built up containing pictures and vital data. A webpage and a book have been published on that material, Aksdal [3], Blom et al [4]. We do have some research articles, informal books and biographies on the HF and a few of its makers [5]-[10]. However, most of it is written in Norwegian. The Hardanger Fiddle Association of America publishes *The Sound Post* quarterly [11]. Recently a growing interest in the older instruments has taken place e.g. by concerts performed on historical fiddles as well as recordings, Hauk and Knut Buen [12], Hamre and Maurseth [13].

1.2 My own work

The author has been collecting acoustic and construction data from HF’s and violins for about 20 years [14], [15]. However, results on the HF have this far been published in articles mainly written in Norwegian, as well as in discussion threads, e.g. on www.maestronet.com. Some results have been shared in talks over the years, but for a rather limited audience.



Figure 1: Hardanger fiddles made by Olav K. Venaas from Gransherad in Telemark in the early 1930ties. The left one is an exhibition instrument. The right one is one of the main instruments in the Buen family. Photos: Asmund Buen.

I have earlier published several analyses based on Long Time Average Spectra (LTAS) extracted from recordings of fine violins [16]. LTAS is simply the time averaged SPL measured over a musical piece or a scale. It ignores the short term details in the signal and emphasizes the long term spectral qualities of an instrument or a voice, a much used method in musical acoustics [17]. LTAS from played scales on a larger set of instruments have now, for some time, been collected systematically using calibrated sound level meters under identical measurement conditions. Impulse responses and modal analysis have also been conducted on a set of violins and HF’s parallel with detailed registration of material properties and geometrical data from these instruments.

2. CONSTRUCTION AND TUNING

The violin is usually tuned G_3 (196 Hz), D_4 (294 Hz), A_4 (440 Hz) and E_5 (659 Hz). On HF's, the A-string can be tuned to any frequency between A_4 and D_5 (587 Hz). The most typical tuning is with the A-string tuned to B_4 . The strings are then tuned: B_3 (247 Hz), E_4 (330 Hz), B_4 (494 Hz), and $F\sharp_5$ (740 Hz), and the five sympathetic strings: $C\sharp_3$ (277 Hz), E_4 (330 Hz), $F\sharp_4$ (370 Hz), G_4 (392 Hz) and B_4 (494 Hz). The HF is usually played only with the hand in the “ground position” and more than twenty different tuning schemes exist, offering a variety of timbres and tonalities. Of these tuning schemes “a handful” is more often used.



Figure 2: A violin bridge (not trimmed) and a Viken type HF-bridge.

The sympathetic strings run under the fingerboard and rest on a “bridge” below the playing strings, typically 18–20 mm above the top plate, see Fig 2.

The string length is typically 25–30 mm shorter than on a violin (328 mm). The strings are thinner and lighter so the higher pitch is, to some extent, compensated for regarding the string tension. A HF tuned in B_4 (using gauge 10.5) will have a somewhat higher total string tension than a violin with normal Dominant, G, D and A strings and a 0.26 mm steel E. The string tension in HF's should, of course, be measured.

3. EXPERIMENTS

Instrument response spectra are collected with the fiddles (both violins and HF's) mounted vibration-insulated, on textile covered rubber bands, in a “Curtin rig”, [18]. The excitation is impulses from a miniature PCB 086E80 impact hammer, mounted as a pendulum. The head weighs 1.38 g. A PCB model 480E09 signal conditioner, an Earthworks M30 1/4” omnidirectional microphone, or a PCB 352A73 accelerometer, connected to a t.c. electronic konnekt 8 external soundcard and a laptop running SpectraPlus SC or George Stoppani's analysis software is used for the analyses. A B&K mini accelerometer type 4374, a type 2647A preconditioner as well as a DeltaTron Power Supply type WB 1372 is also used in some of the analyses.

LTAS of played scales are collected using a calibrated Norsonic N140 sound pressure level (SPL) meter mounted about 1 m from the player seated in a corner of a “hard room”. The room dimensions are: $L \times W \times H = 2.7 \text{ m} \times 1.6 \text{ m} \times 2.2 \text{ m}$, and the average reverberation time over the $1/3^{\text{rd}}$ octave bands from 200 Hz to 20 kHz is: $T_{30} = 0.35 \text{ s}$. The reverberation radius is 0.3m so the mic @ 1 m is well inside the reverberant field. The Schroeder frequency, indicating the start of the statistical region of

the room modes, is a bit high at 374 Hz, but the reverberation time curve is almost flat, with no ringing modes.

3.1 Signature modes and frequency “clusters”

The resonances of fiddles appear in separate identifiable *signature modes* up to about 1 kHz, where the resonances start to overlap and form *clusters*. Figure 3 show a spectrum of the volume change of a fiddle body calculated from modal analysis of the top and back plates up to 1 kHz using G. Stoppani's software. The accompanying vibration shapes are shown with their names.

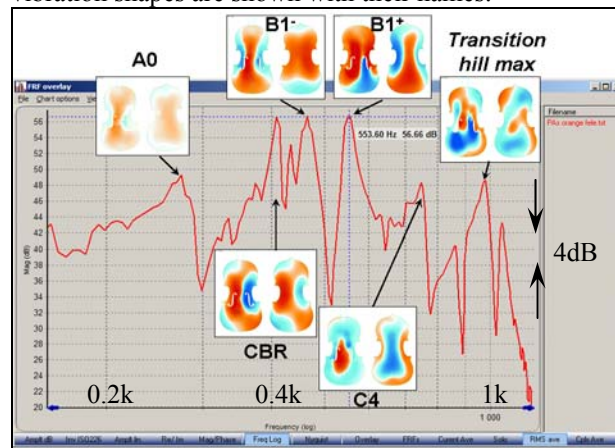


Figure 3: Volume change per Newton input force (mm^3/N) of a violin body with a HF neck with names and vibration shapes for the signature modes. The top and back plates are seen as if they were the outside of an opened book. Red and blue regions move in opposite directions. Same color in large areas of the top and back plates indicates a significant *breathing* motion of the body.

The *signature modes* have been given names related to their shape and function:

A0: is the *main air resonance* formed by the f-holes, the volume inside and the flexibility of the plates, like the Helmholtz air resonance of a bottle. A *breathing mode* (Most of the top and back plate move like a “breathing body”)

CBR: The *C Bout Rhomboid* with a plate-like bending motion of the body. The name is related to the vibration shape of section through the C-bout region. The symmetric form makes it a weak sound radiator, but it can radiate significantly if it has a *breathing component*.

B1- and B1+ Baseball modes, from how the modal lines look somewhat like the stitching lines of a baseball. The nodal lines cross the back plate upper and lower bouts and along the top on each side in B1- and reversed for the B1+. Both are “breathing modes”, usually with strong sound radiation.

C4: A *ring mode* in the back plate. Both bridge feet move in phase, so it is more easily driven by up and down movement of the bridge than a sideways drive. The C4 also can show significant *breathing* of the body.

The next modes are a *cluster* named the **Transition Hill** (TH) which is the frequency region where the sound-post (right) bridge foot starts to move more than the bass bar (left) bridge foot. (Name courtesy: Evan C. Davis). The inner f-hole wings usually start to move a lot in *free edge motions* in this region. The highest level peak of the TH is called the TH_{max}

Figure 4 show a wider normalised SPL frequency spectrum of the fiddle in Fig 3 tapped on the G string side of the bridge with an impulse hammer.

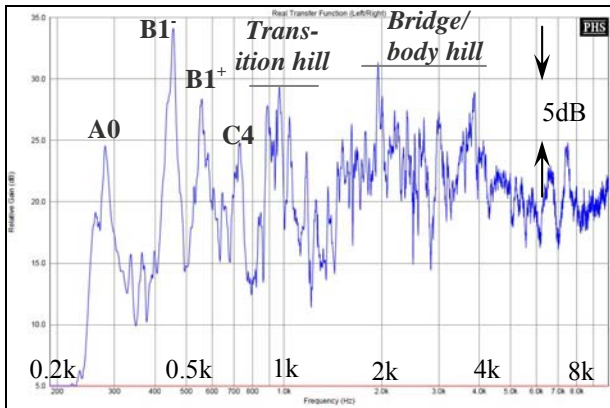


Figure 4: An impulse response spectrum (SPL/Force) averaged over 360 degrees around the fiddle with descriptors for the signature modes and names for the *clusters*. (This fiddle has a HF neck, understrings and a violin bridge. The heart of the bridge is modified as a bridge for the understrings).

The bridge/body hill is a cluster with details not yet disclosed in detail. It is believed to be related to the first bridge in plane resonance and properties of the top plate in the central region, including the f-hole wings.

Figure 5 shows average signature mode frequencies and levels for a set of 42 violins (orange) and 50 HF's (blue). The data overlap, with a tendency for the C4 mode to lie lower in frequency for the HF's. Figure 5 also shows data from twelve Old Italian violins, six contemporary master-built violins as well as a Vuillaume in green, data courtesy of Joseph Curtin. The level for the green data points in Fig 5 is adjusted to fit the other data, so absolute level comparisons are not possible between these.

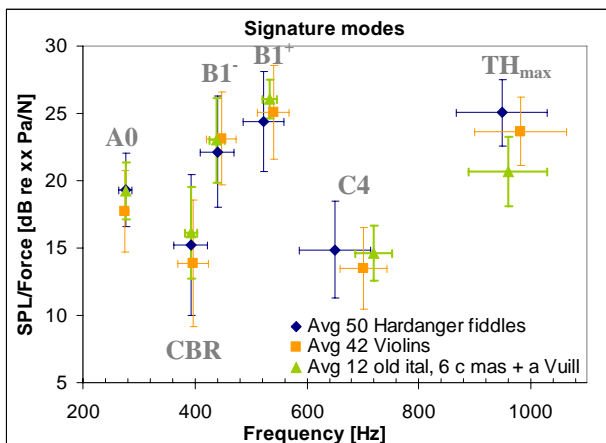


Figure 5: Average signature mode frequencies and levels of 12 Old Italian, 6 contemporary masters + a Vuillaume (green triangles) compared to the data from 50 HF's (blue diamonds) and 43 violins (orange squares). Error bars show one standard deviation to each side of the data point.

The transition hill maximum (TH_{max}) is relatively a little weaker in the green data in comparison to the other two

instrument groups. The signature mode frequencies are otherwise quite similar, but with less spread in the data.

Figure 6 shows LTAS of played scales on the violins and HF's. (Whole and half notes, as well as open strings, are played). In general the SPL is in average 3dB lower for the HF's. The levels tend to overlap for all frequency regions, except for the 3.15-4 kHz bands where the violins have a significantly stronger response. The reason for this is likely to be the bridge design and the higher first in plane resonance of the violin bridges: 3270 ± 303 Hz versus 2720 ± 408 Hz for the HF-bridges, all measured with the bridge feet fixed in a vise.

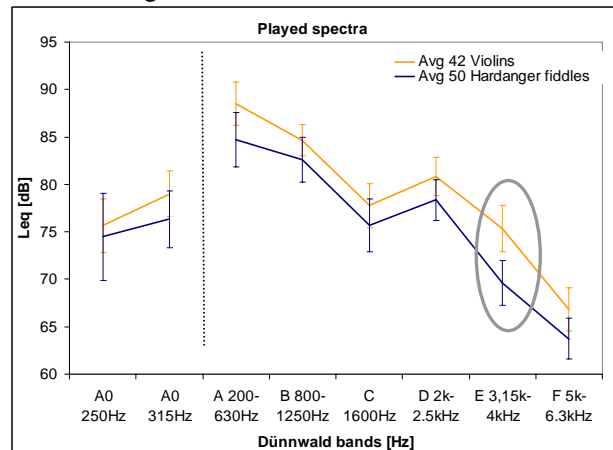


Figure 6: Long Time Average Spectra of played scales on 42 violins (orange) and 50 HF's (blue). The frequency regions are calculated from $1/3^{rd}$ octave band values. The 250 Hz and the 315 Hz bands are shown because of the A0. The other bands are adaptations to those used by Dünwald. The error bars show one standard deviation variation to each side of the line.

The average overall SPL are: $L_{eq} = 87.9 \pm 2.2$ dB for the HF's and $L_{eq} = 91.1 \pm 1.7$ dB for the violins (\pm one σ) both linear values (no weighting of the spectrum is used).

3.2 Test of bridge designs

Most makers have their own bridge designs. E.g. the Valdres makers Knut Ø. Rudi (1878-1972) and Olav Viken (1921-2005) have very different bridge designs: Figure 7 shows Rudis model and Fig 2 show a copy of Vikens model. Among makers and players with some experience, the design difference between these is known to have an influence on the instrument timbre.



Figure 7: A Rudi HF bridge type (not trimmed) left and a fitted one right. [Left: From Wulffnstejns Hardanger fiddle and Mandolin works <http://www.hardingfele.com/>]

I tested this hypothesis on a fiddle using LTAS of played scales with a Rudi and a Viken type bridge on the same

instrument. The first in plane resonance of the bridges was tested with the feet fixed in a vise: Rudi bridge: 4.1 kHz/1.71 g and the Viken bridge: 2.9 kHz/2.01 g. (The Viken bridge was 0,4 mm higher than the Rudi bridge, an insignificant difference). Fig 8 shows the resulting spectra with error bars showing one σ variation to each side of the curve, calculated from three separately measured spectra.

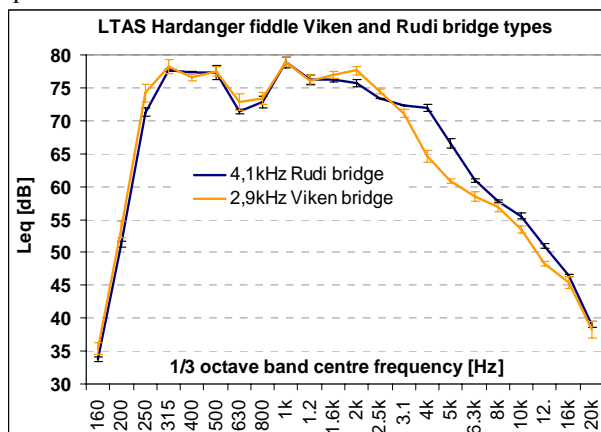


Figure 8: Long Time Average Spectra of played scales on a HF with a Rudi (blue) and Viken (orange) type bridges measured @ 1m. The first in plane resonances of the Rudi-bridge: 4.1 kHz and the Viken-bridge: 2.9 kHz. The error bars show one standard deviation. The fiddles do have understrings in both cases.

We see that the response in the 2 kHz band has increased and the 4-12 kHz bands have decreased using the lower tuned and heavier Viken type bridge. The difference in the 4 and 5 kHz bands are 7 and 6 dB respectively. The author prefers the instrument with the Viken style bridge.

3.3 Influence on the signature modes from the HF neck design

The HF has four or five extra pegs and thus a longer and heavier pegbox and neck. The string length is shorter and the grip region will be shorter too. As the HF usually is played in the ground position, the length of the fingerboard (FB) is shorter by about 40-50 mm (~220-230 mm versus 270 mm for the violin). A violin FB weighs 60-70 g while the shorter and hollowed out HF fingerboard can weigh less than half of that, a bit dependant on design and material.

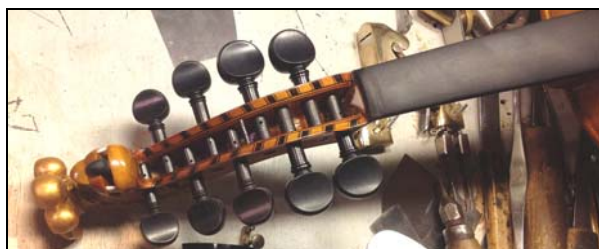


Figure 9: A HF pegbox and neck.

How do the longer and heavier neck and its lighter and shorter fingerboard influence the frequency response of a fiddle body?

A solid built Guarneri model violin was picked for a “makeover” from violin to HF. LTAS of played scales, SPL/Force and Acceleration/Force spectra was recorded before and after the experiment. The sound post was not moved in the process.



Figure 10: The Guarneri model violin fitted with a HF neck. The violin fingerboard (FB) was carved flatter and hollowed out for the sympathetic strings on the HF. Response curves were measured for each step in the working process.

The violin weighed 428 g without chinrest and 456 g as HF, also without the chinrest, so the HF version weighed 28 g more.

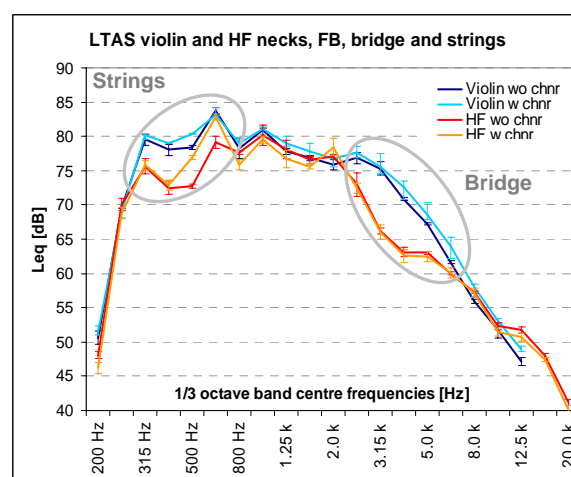


Figure 11: LTAS of scales played on the same fiddle body with a violin and a HF neck, FB, bridge and strings, see Fig 10. Separate runs with and without mounted chinrests are shown.

From Fig 11 we see that the frequency response from 2.5 kHz up to 5 kHz is much stronger in the violin setup. This is likely to be due to the different bridges. The in-plane resonance frequencies and weights of the violin and HF bridges are: 3.4 kHz/2.35 g and 2.7 kHz/1.96 g, respectively.

Also the 250 Hz and 315 Hz octave bands are weaker, probably due to the thinner, shorter and lighter HF strings. The HF strings are 10.5 gauge and the violin strings are higher tension of Chinese origin with a vibration sensitivity about 4-6dB higher than the HF strings for the G, D and A. The E is only about 2 dB higher. The difference seen in the low frequencies are likely to be related to the heavier and higher tension violin strings. We also see the chinrest allows for extra sound output, especially in the 500 Hz and 630 Hz bands for the HF.

The strings were put on for the experiment and to experience how the fiddle played and sounded. The fiddle was surprisingly interesting. The E-string is even sounding. It

sounds a bit shallow compared to many HF's, but may have some similarities to some thick plated HF designs. As a HF, the instrument is “ringy”, suggesting that a stiff body (see the signature data in Table 1) might make the sympathetic string decay a bit longer than normal. We will return to that issue in sec 3.4.2.

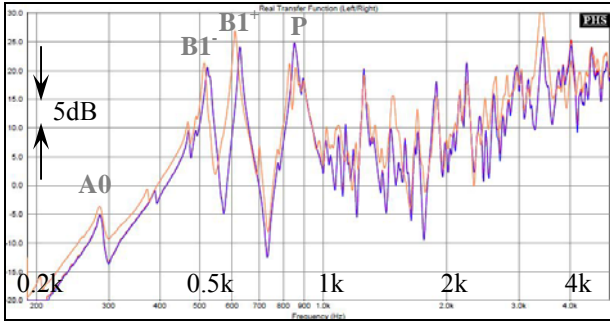


Figure 12: Accelerance $a(\omega)/F(\omega)$ for the instrument shown in Fig 10 without strings fitted with a violin neck, pegs and a FB (blue curve) and a HF neck, pegs and FB (orange curve).

From Fig 12 and Table 1 we see that the modes have moved somewhat lower with the heavier HF neck and 9 pegs. The “P resonance” is influenced by both the pegs and the FB. Without FB, but with pegs, the “P top” is shown as two tops for both the violin and the HF. The same happens when there is no neck. Without FB and pegs the “P resonance” was divided into three tops.

Unfortunately we do not have a modal analysis of the fiddle, so we do not know in detail how the mode vibrates, including the “P”. A modal analysis probably would have helped to explain why the pegs, the neck and the FB had such an influence.

	Frequencies [Hz]			
	B1 ⁻	B1 ⁺	C4	TH _{max}
Violin	523	619	730	1083
HF	504	604	722	1089
Change [Hz]	19	15	8	6
Change [%]	4 %	2 %	1 %	0,6 %

Table 1: Signature mode frequencies and changes measured for the violin and HF, both without a chinrest.

3.4 How do sympathetic strings influence the sound?

Saunders found that playing the same note as one of the open strings on a violin, with the open string damped gave a 2 dB increase in the sound output [19]. A sympathetic string makes the played note sound “fatter”, and the note decays longer as the open string is not damped by the finger tips. The effect is clearly audible, especially for the HF where the understrings are always free to vibrate.

The effects are experienced while stringing up a HF playing the instrument with only the G, D, A, and E strings in place and later with the sympathetic strings in place. It is also easy to just dampen the strings below the fingerboard e.g. with a piece of rubber foam for playing tests.

The playing experience and the sound are quite different in the two cases. What happens to the sound and the playability?

3.4.1 Objective and subjective loudness

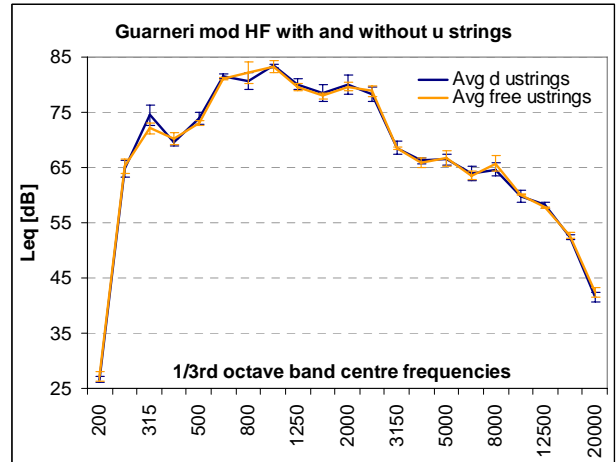


Figure 13: LTAS of scales played on the same fiddle with a HF neck, see Fig 10, with the understrings damped using a rubber foam piece under the fingerboard (blue) and no damping (orange). Errorbars show one σ variation to each side of the curve.

Figure 13 shows a comparison of LTAS of played scales taken from the same HF with the understrings damped and free to vibrate. The curves show no significant difference in the sound level using a t-test on the data sets of three separate scale runs. Spectra taken with and without the small foam rubber piece between the fingerboard and the understrings shows that the foam rubber piece does not dampen the instrument modes significantly.

Psychoacoustically the ears and listening system integrate the nerve activity from a perceived sound over 100-200 ms duration, the so called *Temporal integration* [20]. A note with a string singing along and decaying parallel with it, and with a slower decay, could give the impression of a louder sounding note, even if we may get a lower reading from the SPL-meter. Repetitions increase the sensitivity to details in the sources of sound spectra [21]. This should be investigated further.

3.4.2 Reverberation times of body, strings and room

The “ring” is a characteristic of the HF. Originally the HF was played in rather small and “dry” sounding rooms. Following the tradition after Torgeir Agundssons (Myl-larguten’s) concerts with Ole Bull, the HF did enter the concert hall stage around 1850. It is more often used in public meeting halls for dance, competitions and concerts. It is now fairly common that it is amplified, even in major larger competitions where sports halls often are used to fit an audience similar to the larger concert halls.

When the HF is amplified, the ringing understrings are well heard, also in seats in larger audiences, an experience a quiet and intimate auditorium otherwise would ensure with no amplification of the HF.

Reverberation times of normal old and modern home environments may be 0.4-0.6 seconds, quite moderate values. The hypothesis is that the sympathetic strings give a more reverberant experience, somewhat resembling the reverb experienced in concert halls or churches. But without the envelopment one would experience from a full 3D decaying soundfield.

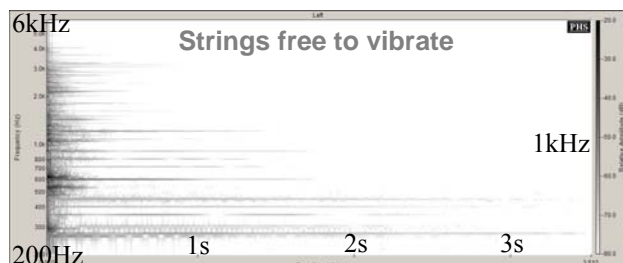


Figure 14: Spectrogram of a bridge tap on a HF with all strings free to decay. The plain sympathetic strings ring longer than the overwound playing strings. The grayscale is black: 0 dB white: -60 dB. Frequency resolution: 10.7 Hz time resolution: 0.09s.

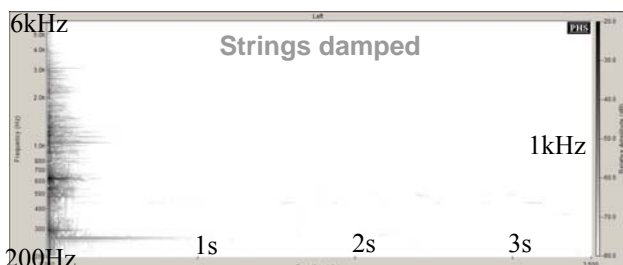


Figure 15: Spectrogram of a bridge tap on a HF with all strings damped. The grayscale is black: 0 dB white: -60 dB. “The ringing thing” below 300 Hz is a tailpiece resonance. Frequency resolution: 10.7 Hz time resolution: 0.09s

In Figs 14 and 15 we show spectrograms of recorded bridge taps of the stiff experimental HF seen in Fig 10. The taps are done with a miniature impulse hammer with a soft tip. A soft tip gives a stronger low frequency response and a somewhat weaker high frequency response. This mimicks to a certain degree the 6 dB per octave decreased force input on the bridge from the bowed string. The spectra have not been normalised. From Fig 15 we see that the body decays to -60 dB within 200-250 ms. (The decaying resonance ringing to about 1 s is from a resonance in the tailpiece).

In Fig 14 the strings are free to vibrate and we see the same initial body decay overlapped by the ringing string fundamentals and harmonics. The lowest sympathetic string are still ringing after 3.5 s and the other four rings for at least 2.5 seconds while the harmonics around 1 kHz ring for about 1.4-1.7 s. The string resonances close to a strong body resonance will decay faster. E.g. the 1st sympathetic string, and the open A string, (here tuned to C₅=523Hz) decays faster than the other sympathetic strings with a weaker coupling to strong body resonances.

Now can the spectrograms show any differences between instruments?

In Fig 16 a spectrogram of the ringing strings from an instrument with a stronger low frequency response. The sound looks, and indeed is, “fatter”. But the string fundamentals are not ringing as long as they do on the stiffer instrument in Fig 14.

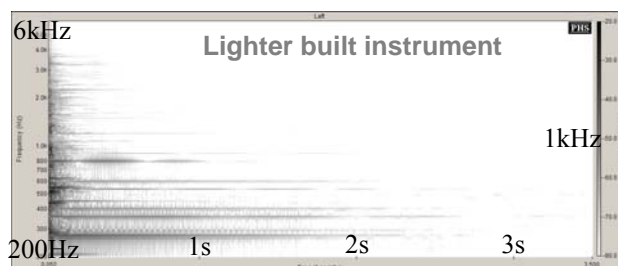


Figure 16: Spectrogram of a bridge tap on a lighter built HF with a stronger low frequency response. The grayscale is black: 0 dB white: -60 dB. Frequency resolution: 10.7 Hz time resolution: 0.09s

Figure 17 shows a spectrogram of a popped balloon in the measurement room used in Figs 14-16. The impulse shows the decaying sound field in the room. The reverberation time is 0.43 s averaged over the 1/3rd octave bands 200 Hz – 20 kHz.

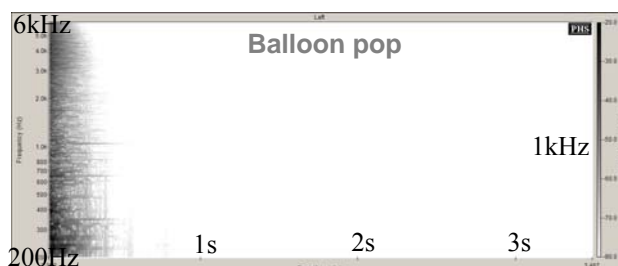


Figure 17: Spectrogram of a balloon pop in the measurement room. The T₃₀ is earlier measured to 0,43s. The grayscale is black: 0 dB white: -60 dB. Frequency resolution: 10.7 Hz time resolution: 0.09s

It is quite clear that the reverberation of the strings dominates the combined reverb of the living room and the decaying instrument, as it is much longer. The measurements in Fig 14-16 are done with the microphone only 20 cm from the top of the fiddle, that is: in the near field. This simulates what a player will hear.

4. DISCUSSION

Not much has been written about the acoustics of the HF, but the violin has been studied acoustically by enthusiasts for more than 200 years. That literature serves as a good reference for comparisons.

4.1 Signature modes

The average signature mode frequencies and levels of HF and violins in average lie quite close. The C₄ (ring mode in the back plate) have lower frequencies, possibly due to somewhat thinner and, sometimes, softer backs (some instruments have black alder backs). The B1+ mode appears at a lower frequency, something we may relate to the heavier neck and more pegs.

The spread in the level and frequency data are larger for my sets than the Old Italian and modern violin set. I have included anything from mass produced to fine or re-graduated instruments, often in “before and after” pairs. A larger variation in the data are thus to be expected.

4.2 HF bridge designs

HF’s sound different with a short legged Rudi bridge as opposed to a long legged Viken style bridge. Spectra from played scales (LTAS) show that there is a cut in the higher frequencies by using a Viken type bridge.

It is assumed that the lower in plane resonance of the longer legged bridge is the cause for this effect. Even though the violin and HF bridges are very different looking this ties in with Woodhouses conclusion in his bridge hill article for violins, [22] A working hypothesis is that the HF bridges have a “side to side” first in plane mode similar to the first in plane mode of the cello bridge, Rodgers [23]. Work is in progress to investigate this further.

4.3 The effect of the HF setup

The pilot study changing the violin neck to a HF neck and FB on a rather stiff violin body shows that the signature modes tend to move downwards 0,5-4%. The bridge design (and the first in plane resonance) influences the high frequency response, with the long legged Viken style HF-bridge cutting some of the higher frequencies. Both the thinner and somewhat lower tension HF strings drive the body with a weaker force, explaining the weaker SPL from a HF than a violin. On average the sound output from the HF in my set is 3 dB weaker. There are no studies known to the author comparing the acoustics of the violin and the HF, but the relation between the violin and the Baroque violin has been studied in a PhD work by McLennan [24].

4.4 The effect of the sympathetic strings

A pilot study of played scales on an instrument with the sympathetic strings damped and not damped, using LTAS, indicates that there is no measurable difference in the sound output in the two cases. This is not supporting Saunders finding for the sympathetic effect of violin strings where he found a 2 dB increase of the sound output when the sympathetic (open playing) string was damped [19]. The reason is likely to be the large level difference between the played and the sympathetic string. Michaelsen found that while playing a single note on the HF, the sound level of the harmonics of the sympathetic string is 10-40 dB below its exciter [1]. Adding the sound level of two sounds with such a difference will give a very low or no effect on the net SPL, which fit well with the measured data in Fig 13.

The sympathetic strings decay more slowly than the playing strings. So at a certain time the sound level of the decaying sympathetic string will be louder than the playing string, if no energy is fed to the playing string. Psychoacoustically a longer lasting sound should be per-

ceived as being louder, due to the temporal integration effect.

4.5 Room acoustics and the decaying understrings

Traditionally the HF has been used in smaller rooms than the violin. Most HF’s are probably played mainly in home environments, or in smaller practice rooms with similarities to “home acoustics”. But the HF does not have the same history on concert hall stages, churches and “homes with high ceilings” as the violin. The extra reverb from the strings might be a compensation for not having “lively rooms” to play in. Or maybe it just turned out to be more fun to play the fiddle with sympathetic strings.

Preliminary tests indicate that the reverberation time (-60 dB decay) of the understrings may be in the order of 1.5s to over 3.5s. And it seems to be related to where their frequencies lie in relation to the HF body resonances, as indicated in Woodhouses work on the guitar [25]. Sonograms seem to give a good qualitative indication of the decaying strings, and may be utilized in quantitative analyses, as suggested by Woodhouse [26].

Preliminary results and experiences indicate that stiffer built instruments might ring a bit longer than lighter built instruments. This opens up for a variation in the construction of the fiddles for their ringing qualities. Lighter built instruments, on the other hand, tend to sound more “bassy” and somewhat louder. The “ringy quality” of the HF is best heard in slower played music and by the player. But in intimate quiet rooms, the ring is well heard, also by the audience. When the HF is amplified, the ring of the understrings, as it sounds for the player, is presented to a wider audience.

The decaying strings are, at least to a first approximation, following the simple physics of a decaying harmonic oscillator. The bowed strings, however, is a highly non linear physical phenomenon of “slip stick motion”, Guetler [27] and Woodhouse [28]. The effect the freely vibrating strings has on the playability, and the rise time of played notes, is something to investigate further. Michaelsens experimental study indicates that the rise time increases due to the sympathetic strings, but they have very narrow resonances to “hit”, so the rate of hitting them is not guaranteed to be 100 % [1].

This is also a matter of how well the instrument is tuned, how “lossy” the string nuts and bridge are, as well as the intonation skills of the player. It may also depend on the detailed intonation style in the played tradition. Even if the transients might become longer, sympathetic strings do not seem to prevent a fast playing style, as can be heard in some of e.g. Olov Johanssons *nyckelharpa* music (Väsen) or some of the HF music after Ola Mosafoinn.

5. CONCLUSIONS

The HF show close similarities to the violin regarding its body resonances. The bridge design, strings and the tonal ideal is, however, different. The HF is somewhat weaker sounding, mainly due to the lighter and shorter strings. The first in plane resonance of the bridge should, ideally, lie lower in frequency than for a violin, at least in the author's opinion.

A more detailed study of HF should be done to reveal some of the acoustic effects of the more variable designs used in HF making historically. The effect of the sympathetic strings should also be investigated further regarding the playability and transients. The nature of the HF-bridge, and the different designs, should also be investigated in more detail.

Acknowledgments

Thanks to Brekke & Strand Akustikk AS for borrowing the Norsonic N140 units and to Tor Halmrast, Sigmund Olafsen, Jens Jørgen Dammerud and Cahterine Springer for help with the manuscript. Thanks to Joe Curtin for fine violin data and to the rewiewers for helpful suggestions.

6. REFERENCES

- [1] P. Michaelsen: "The Hardingfiddle - An investigation of its sympathetic strings" *Catgut Acoust. Soc. Newsl.*, #34, November 1980, 4p.
- [2] P. Cook and D. Trueman: "A Database of Measured Musical Instrument Body Radiation Impulse Responses, and Computer Applications for Exploring and Utilizing the Measured Filter Functions", *Proceedings from ISMA*, 1998, 6p
- [3] B. Aksdal: *Hardingfela*, tapir akademisk forlag, Trondheim, 2008, 246 p.
- [4] <http://hardingfeler.no/>
- [5] Sverre Sandvik: *Vi bygger Hardingfele*, Tiden Norsk Forlag, 1983, 169 p.
- [6] Å. Nisi: "Kunstnaren Olav K. Venås", Buen Kulturverkstad, 1985, 39 p.
- [7] Versto, et al: *Hardingfela Det norske nasjonalinstrumentet*, Grøndahl og Dreyers Forlag AS, 1997, 128 p. Contains a CD with HF music.
- [8] Kjell Midtgaard,: *Hardingfelemakaren Gunnar Røstad*, Published by Kjell Midtgaard, 1997, 118 p
- [9] Vidar Lande: *Frøysaatradisjonen Musikaner og felemakarane frå Iveland, Aust Agder: Torleiv og Olaf Frøysaa*, Department of Norwegian Folk Culture, Telemark University College, Rauland, 2005, 177 p.
- [10] O. S. Krakstad: *Felemakaren Sveinung Gjøvland 1892-1977*, Åmli Boknemnd, 2012, 192 p.
- [11] www.hfaa.org
- [12] Hauk and Knut Buen: *Myllarfela - Tunes played on the Myllarfela*, MC, Buen Kulturverkstad, 1991.
- [13] K. Hamre and B. Maurseth: *Rosa i Botnen*, CD and DVD, Grappa Musikkforlag AS, 2005.
- [14] A. Buen: *Svingemoder i fioliner*, MSc thesis, NTNU, Trondheim 1994, 135p.
- [15] A. Buen: "Operating Deflection Shapes of Five Conventional and two Unconventional Violins", in *Proc. of the Stockholm Music Acoustics Conference*, 2003 (SMAC 03), 4p.
- [16] A. Buen: "On Timbre Parameters and and Sound Levels of Recorded Old Violins", *J. Violin Soc. Am. VSA Papers*, Summer 2007, Vol XXI, No 1, pp 57-68.
- [17] Sten Ternstrøm: "Long-Time Average Spectrum characteristics of different choirs in different rooms" *Voice, No 2, 1993*, British Voice Association.
- [18] J. Curtin and T Rossing: "Chapter 13 Violin", in *The Science of String Instruments*, ed. T Rossing, Springer, 2010, 470 p.
- [19] F. A. Saunders: "Recent work on violins", *J. Acoust. Soc. Am.* Vol 25, No 3, May 1953, pp 491-498.
- [20] B.J. Moore: *An introduction to Psychology of hearing*. Fifth ed. Emerald Group Publishing Ltd, 2008, 413 p.
- [21] F. Toole: *Sound reproduction Loudspeakers and rooms*, Focal press, 2008, 550 p. Ch 19.2. Hearing tilts, peaks, dips bumps and wiggles.
- [22] J. Woodhouse: "On the "Bridge Hill" of the Violin", *Acta Acoustica*, Vol 91, 2005, pp155-165.
- [23] O. Rodgers and T. R. Masimo: "The effect of wood removal on bridge frequencies", *Catgut Acoust. Soc. J.* Vol 1, No 6 (Series II), November 1990, pp 6-10.
- [24] J.E. McLennan: *The violin music acoustics from Baroque to Romantic*, PhD thesis, School of Physics, The University of New South Wales, Sydney, Australia. 236 p.
- [25] J. Woodhouse: "Plucked Guitar Transients, Comparison of measurements and synthesis" *Acta Acoustica*, Vol 90, 2004, pp 945-965.
- [26] J. Woodhouse and R. S. Langley: Interpreting the input admittance of the Guitar and the Violin, *Acta Acoustica*, Vol 98, 2008, pp 611-628.
- [27] K. Guettler: *The Bowed String On the development of Helmholtz Motion and On the Creation of Anomalous Low Frequencies*, PhD thesis, KTH, Speech, Music and Hearing, Stockholm, 2002, 26p + 5 articles.
- [28] J. Woodhouse: "On the Playability of Violins: Part II Minimum Bow Force and Transients", *Acta Acoustica*, Vol 78, 1993, pp 137-153.

Acoustical features of complex bowing patterns in violin performance

Matthias Demoucron

Institute for Psychoacoustic and
Electronic Music (IPEM),
Ghent University,
Belgium

matthias.demoucron@ugent.be

Erwin Schoonderwaldt

Institute for Music Physiology
and Musicians' Medicine (IMMM),
Hanover University of
Music, Drama and Media
Germany

schoondw@gmail.com

Marc Leman

Institute for Psychoacoustic and
Electronic Music (IPEM),
Ghent University,
Belgium

marc.leman@ugent.be

ABSTRACT

Music performance relies on fine motor skills and subtle judgements of the resulting sound in order to adjust body movements involved in playing. In this respect, fast repetitive bowing patterns in which notes are alternatively played on two adjacent strings offer an interesting case study of subtle coordination of bow movements. In recent experiments, it was shown that a particular timing relation between bow velocity and string crossing was preferred in both real performance and perceptual tests, the bow change occurring consistently after the string crossing. The aim of the present study was to identify and quantify acoustical features on which violinists may base this preference. For that purpose, we performed simulations over a wide range of coordination parameters. The results of perceptual tests were examined under the light of the simulations in order to determine significant acoustical features and the acceptance limits associated with them.

1. INTRODUCTION

Recent motion-capture measurements of violin bowing revealed an interesting coordination behavior of bow changes and string crossings. This behavior could be clearly observed in fast repetitive bowing patterns, in which notes are alternately played on two adjacent strings, leading to typical circular spatial trajectories of the bowing hand. It was shown that in this type of bowing patterns bow changes were consistently lagging behind string crossings in all observed performances by several performers [1, 2]. This timing relation was achieved by a phase delay of the bow-velocity signal of about 10 to 30 deg. relative to the bow-inclination signal, both signals being approximately sinusoidal. In a subsequent study, we investigated the origin of this coordination behavior through perceptual tests with a virtual violin, in which some parameters of the repetitive string crossings could be controlled by the listener [3]. The result was that the same shift between bow changes and string crossings was preferred by the subjects, implying a

perceptual origin for the development of this coordination behavior.

The purpose of the present study is to analyze the acoustical features that may explain the observed perceptual preferences. As a matter of fact, complex bowing patterns involve different features for the bowed string that are somehow concurrent: the player has to play alternatively on two different strings, but the coordination between bow changes (or reversal of bowing direction) and string changes may imply extraneous sounds that have to be minimized for a clean performance. The analysis presented here is based on simulations inspired by the experimental conditions of the perceptual tests. Simulations representing a wide range of coordination parameters between bow velocity and string crossings were analyzed in order to identify influential acoustical features in the subjects' preferences. The paper is structured as follows. First, we will present the theoretical background of the study and the perceptual experiment that was performed before (Sect. 2). Then, in Sect. 3, the simulations will be shortly described. In Sect. 4, important acoustical features will be illustrated and discussed. Finally, the results of the perceptual tests will be compared with the simulations in order to deduce some empirical limits for the acoustical features described before (Sect. 5).

2. THEORETICAL AND EXPERIMENTAL BACKGROUND

2.1 Coordination of string crossings and bow changes

This study deals with complex bowing patterns involving fast repetitive string changes combined with bow changes, for example playing alternatively down-bows on one string, and up-bows on another adjacent string. Beside motor issues related to the technical performance of these bowing patterns, which will not be tackled here, there are some acoustical constraints that may explain why some coordination behaviors are favored by players.

One might first assume that these bowing patterns are achieved by synchronizing precisely bow changes with string crossings. However, string crossings cannot be instantaneous and there is a finite angular range in which the bow touches the two strings, while going from one string to another. On one hand, when the bow starts to touch the next string before the bow change, a stroke is initiated on the

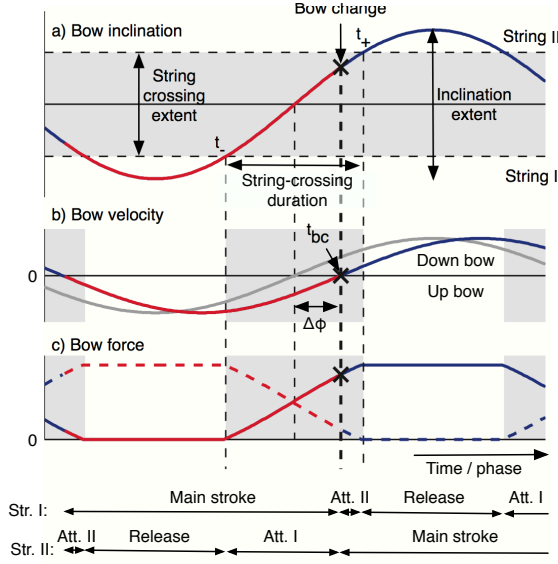


Figure 1. Coordination model and phases of the bowing patterns for fast repetitive string crossings. From the top: bow inclination, bow velocity and bow force. The gray area represents the string-crossing extent where the bow is in contact with the two strings at the same time. Bow inclination and bow velocity are modeled as sinusoidal over time, with a phase difference $\Delta\phi$ (see text for details).

next string with finite bow velocity and zero bow force, producing a preliminary attack that we will call false attack of type I (also called false attack I in the following). On the other hand, when the bow is still in contact with the previous string after the bow change, a new stroke is initiated on the previous string, which will be called false attack of type II (or false attack II). As a consequence, the player has to develop a subtle coordination in order to avoid producing extraneous sounds that may hinder a clean performance of the bowing patterns considered here.

2.2 Coordination model

The coordination between bow changes and string crossings can be formalized with a simple model presented in this section. This model is a simplification of actual performances observed with measurements and its limits are discussed in [3].

In the model, string crossings are controlled by bow inclination, shown in Fig. 1a. The gray area in the figure represents the string-crossing range where the bow is in contact with the two strings at the same time. The string-crossing angular range r is normalized with respect to the peak-to-peak extent of the bow-inclination movement. Typically, the normalized range r takes values between zero and one, with zero representing an instantaneous string crossing, and one a configuration in which the bow is always in contact with the two strings. The string crossing takes place between time point t_- at which the bow inclination enters the string-crossing range from string I, and time point t_+ at which it leaves the string-crossing range to string II. During this time interval, bow force is distributed

across the two strings in proportion to bow inclination, as shown in Fig. 1c.

Bow inclination and bow velocity are both modeled as sinusoidal over time, which is a good approximation of actual performance, with bow changes corresponding to zero crossings of bow velocity (Fig. 1b). The main coordination parameter in the model is the relative phase $\Delta\phi$ between bow velocity and bow inclination, which controls the timing of the bow change relative to the string crossing. The phase $\Delta\phi$ (and thus the position of the bow change t_{bc} in- or outside the string-crossing area) determines the conditions for starting and stopping of the string vibrations.

The space of range-phase combinations basically divides into three areas. For values of $\Delta\phi$ below $\Delta\phi_- = -\arcsin(r)$, the bow change occurs before bow inclination enters the string-crossing area. In that case, the main stroke is started “from the air” (with zero bow force and a finite bow velocity), stopped with full bow force, and only false attacks II occur. For values of $\Delta\phi$ above $\Delta\phi_+ = \arcsin(r)$, the main stroke is started with full bow force, released before the next bow change and only false attacks I occur. Finally, between the two, the main stroke is started and stopped with a bow force below the maximal bow force, and false attacks of both type I and II occur.

2.3 Perceptual experiment

For the perceptual experiment, the above coordination model was implemented in Max/MSP to control a virtual violin. Note patterns implying repetitive string crossings and bow changes with various string combinations and dynamics were presented to the subjects. They were played at a tempo of 100 bpm (note rate of 6.67 notes per second). Participants could adjust in real-time the coordination parameters via a simple graphical user interface in order to achieve the best sounding note transitions according to their judgment. A total of 16 participants was recruited. They were all experienced string players, considered to be able to judge the subtle influence of the coordination parameters on the quality of the transitions. A detailed description of the experiment can be found in [3].

In the two first experimental conditions, string-crossing range r was set to fixed values of 0.54 and 0.28, and the subjects were asked to move a horizontal slider controlling phase in order to “find the clearest/best sounding note transition”. In the two other conditions the participants were asked to adjust the coordination parameters $\Delta\phi$ and r via a 2D slider in order to obtain optimal sounding note transitions. These two conditions were slightly different in their instruction. In the first condition, participants were entirely free to choose the combination of coordination parameters (“optimal condition”). In the other condition (“minimal condition”), they had to find the best sounding note transition while striving to minimize the amplitude of the inclination (i.e. maximizing the range r). For each stimulus the participants were given as much time as they needed to find the optimal combination of coordination parameters. The main result of the perceptual tests was that a positive phase was preferred, similar to the phase difference observed in actual performances. This result showed that

movement features observed in quick repetitive bowing patterns emerge for a great part from acoustical features perceived by the player. It was also found that the phase was not strictly following the arcsine curve of $\Delta\phi_+(r)$ where the bow change coincides with the moment of leaving the string-crossing range. Second, the increasing standard deviation of phase at range values close to one indicated an increased ambiguity in the phase responses. A decrease in the number of responses showed that the participants avoided this region, which could be indicative for a deterioration of perceived sound quality at large values of range. This latter observation was supported by the fact that the “minimal” condition gave mean ranges around 0.6. The “optimal” condition gave range values well below, with a mean range around 0.3.

3. SCOPE OF THE STUDY

The aim of this study is to identify, discuss and illustrate important acoustical features that may explain the preferred phase difference observed in both measurements and perceptual tests. For that purpose, we performed a large amount of simulations inspired by the experimental conditions of the perceptual tests, over a wide range of coordination parameters (range r and phase $\Delta\phi$). This section describes the simulations and illustrate them with an example. The next section (Section 4) will discuss important acoustical features that we considered in the analysis. Then, in Section 5, we will use the results of the perceptual tests in order to deduce some empirical acceptance limits for the effective duration of these acoustical features.

3.1 Simulations

The coordination model was used to control a virtual violin based on physical modeling in order to simulate various coordination scenarios and to investigate the related acoustical features. The virtual violin models the bowed string with a string equation coupled to a slip-stick friction force representing the action of the bow. The motion is simulated by solving the equation on its modal basis and is controlled with the three main bowing parameters: bow force, bow velocity and bow-bridge distance. Detailed descriptions of the model and its numerical implementation can be found in [4, 5].

In the perceptual study [3] a four strings model was used, allowing to examine preferred coordination parameters for string crossings in different adjacent strings combinations. However, in the simulations presented here, only one string combination was examined in order to restrain the analysis to a limited number of cases. The string combination A-D with both strings stopped one tone above the fundamental frequency of the open string was considered a median version of all configurations examined during the perceptual study, representative of the effects we were focusing on. The coordination model used the following parameters: 0.8 N for the maximal bow force, 0.88 m/s for the velocity amplitude, 0.13 for the normalized bow-bridge distance, and 0.15 s for the main stroke duration T .

Simulations were performed with combinations of coordi-

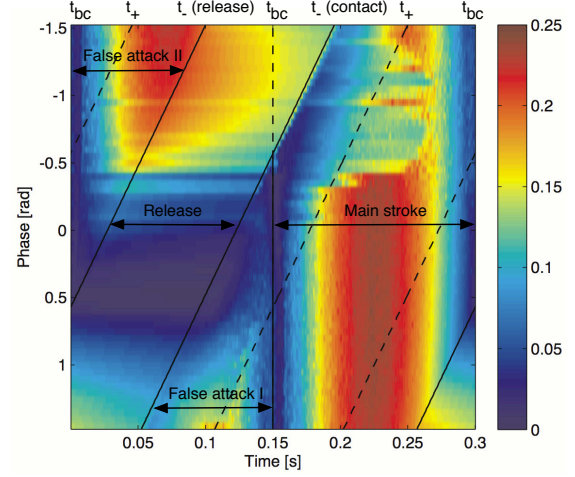


Figure 2. Energy of the bridge force signal obtained for one of the strings in the simulations on one bow stroke period for string-crossing range $r = 0.54$ and phases $\Delta\phi$ varying from -1.5 to 1.5 rad. The main stroke is situated in the right part of the figure. The diagram has to be read in a cyclical way, with $t = 0.3$ corresponding to $t = 0$. The solid black oblique lines show the time of release and contact. In the contact-release interval, the time interval situated between two successive dashed black oblique lines (labelled t_+) shows the region where the bow force is maximal. The bow changes are situated at $t = 0, 0.15$ and 0.3 .

nation parameters covering an extensive area of the coordination space. Phases were varied from -1.5 to 1.5 rad with four string-crossing ranges: 0.05, 0.28, 0.54 and 0.95. Two string-crossing ranges (0.28 and 0.54) corresponded roughly to the results of the perceptual tests in the “optimal” and “minimal” conditions, respectively (see Sect. 2.3). In addition, two ranges represented extreme cases in which the string crossing is almost instantaneous ($r = 0.05$), or the bow is almost always in contact with both strings ($r = 0.95$). Each combination of coordination parameters was used to simulate two seconds of sound, representing about seven samples for each stroke.

3.2 Example

As an illustration, Fig. 2 shows the energy of simulations along one bow stroke period, for one of the two strings and for a range $r = 0.54$ and phases varying from -1.5 to 1.5 rad. The diagram has to be read in a cyclical way (for instance, the black line corresponding to the time of release continues on the right part of the figure, when it reaches the left axes). The successive steps of the stroke can be identified: 1) false attacks II in the upper part of the diagram, between the first bow change and the release, 2) release, between the time of release and the time of contact, 3) false attacks I in the lower part of the diagram, between the time of contact and the second bow change, and 4) the main stroke, situated in the right part of the diagram, between the second bow change (or the time of contact, when the time of contact is after the bow change) and the third

bow change (or the time of release, when it is before the third bow change).

In this diagram, we can make three main observations. First, the main stroke (on the right part of the figure) is highly degraded below $\Delta\phi = 0$. This is due to the fact that, in this region, the stroke is initiated with zero or low bow force. For this reason, it is reasonable to think that players will naturally tend to avoid this region. Second, in the left part of the figure, it can be seen that there is a clear region, for values of $\Delta\phi$ between 0.1 and 0.6, where the release of the string is minimized, which means that the previous main stroke is rather well stopped. Third, there is no symmetry between false attacks I and false attacks II: for a similar duration, false attacks II tend to develop quicker and with higher energy than false attacks I. This can be explained by different attack conditions that are more favorable for false attacks II (attack “from the string”, with high bow force) than for false attacks I (attack “from the air”, with zero bow force). This aspect will be discussed further in the next sections.

4. ACOUSTICAL FEATURES

In this section, we will discuss more precisely the acoustical features related to repetitive bowing patterns involving both bow changes and string crossings. The most obvious features are the two types of false attacks shortly described before. The quality of the main stroke, especially the attacks, should also deserve a special attention. However, with the bowing parameters used in our simulations, it did not have a drastic influence in the range of subjects’ preferences (for values of $\Delta\phi$ above 0). For this reason, the examination of attacks in the main stroke was let aside in the current study. In contrast, the quality of the stopping of the main stroke could be of interest and will be discussed later in this section.

4.1 False attacks of type I

False attacks I are produced when the bow touches the next string before the bow change. This happens for phases $\Delta\phi$ greater than $-\arcsin(r)$. They are started “from the air”, with a finite bow velocity and zero bow force. During the time interval preceding the bow change, the force on the new string increases, while bow velocity decreases.

When initially at rest, the string is not captured by the bow at the moment when they contact each other: the string slides constantly under the bow until the conditions for adherence are met, i.e. when the string velocity under the bow and the bow velocity are equal. However, in the bowing patterns examined here, the string is not initially at rest but may already have a given velocity due to the ringing of the previous note. Therefore, the initial capture of the string by the bow depends on the bow velocity and string velocity amplitude at the moment of contact. After the capture, the subsequent development of the vibration is governed by the evolution of bow force and velocity.

Different conditions for the development of false attacks I are illustrated in Fig. 3. The figure shows the string velocity under the bow for increasing phases (from the top:

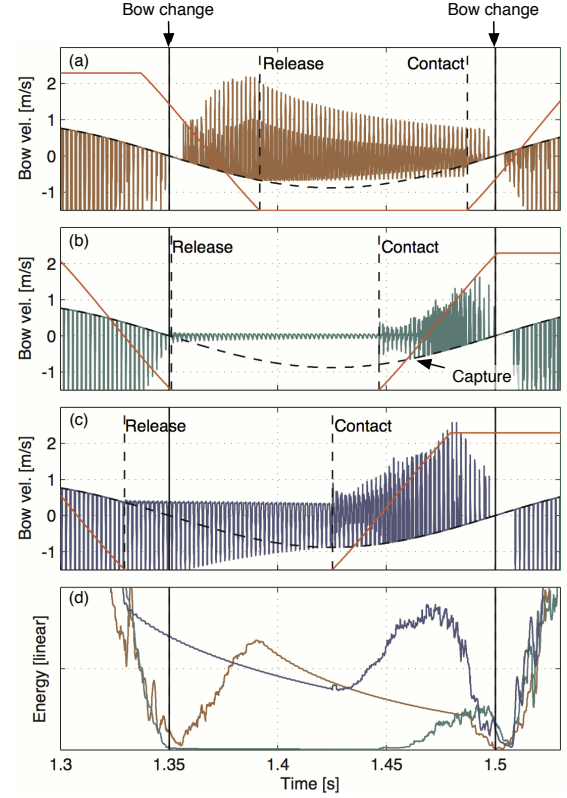


Figure 3. Illustration of different conditions for the development of false attacks I and II with a string-crossing range $r = 0.54$. (a-c) string velocity under the bow for increasing phases $\Delta\phi$ (from the top: $\Delta\phi = -0.3$ (red), 0.55 (green) and 1 (blue)). Bow velocity is indicated by a dashed black line, and the shape of bow force by a red line. Plain and dashed vertical black lines indicate the position of bow changes and release / contact, respectively. (d) energy of the force at the bridge with colors corresponding to the different cases.

-0.3, 0.55 and 1), and the corresponding energies (Fig. 3d). False attacks I can be observed before $t = 1.5$, between the dashed vertical line indicating the contact with the string, and the bow change. In Fig. 3a, the initial velocity of the string is greater than the bow velocity and the string is therefore immediately caught by the bow. However, the duration of contact before bow change is too small and the false attack has no time to develop. In this case, the string is consequently simply damped before the bow change, as can be seen in the corresponding energy (Fig. 3d). In contrast, Fig. 3c shows also a condition in which the string is immediately captured, but with a much longer duration for the false attack. The false attack has time to develop and produces an extra note before the bow change. In Fig. 3b, the initial string velocity is very low and the string is not immediately captured by the bow. During a short time interval, the string is constantly sliding with a slowly increasing amplitude until the amplitude reaches a value that allows it to be captured by the bow. Then, during the remaining time interval, the false attack can develop, but however less than if the string was immediately captured by the bow. The phase value may therefore imply three

different cases for false attacks of type I: simple damping of the string vibration, long sliding with delayed attack development, or immediate capture with full development of the vibration.

This last feature is an essential aspect of the current analysis: the effective duration of the false attack may be actually shorter than the theoretical duration given by $\frac{T}{\pi}(\Delta\phi + \arcsin(r))$. In the following, the duration of false attacks of type I will be therefore described by this effective duration between the capture (given by the moment when the string starts to stick to the bow) and the subsequent bow change.

4.2 False attacks of type II

False attacks II are produced when the bow is still in contact with the previous string after the bow change, which happens for phases lower than $\arcsin(r)$. They are started “from the string”, with a finite bow force and zero bow velocity, and the onset depends mainly on the initial bow force and bow acceleration, as described by Guettler [6]. However, there are some major differences with the situation examined by Guettler. First, the bowing parameters vary during the attack. For example, bow acceleration is maximal at the bow change and decreases until the maximal velocity is reached, at the middle of the stroke. Another difference with Guettler’s analysis is that the string is not initially at rest, which implies different initial conditions possibly influencing the development of the motion. Different false attacks of type II can be seen in Fig. 3, between the first bow change (at $t = 1.35$ s) and the release. In Fig. 3a, the false attack is initiated with a rather high force decreasing quickly. The attack has time to develop, then the string is released about 40 ms after the bow change. As a result, the amplitude of the vibration is rather high in the time interval between the release and the next contact with the bow. This situation clearly shows two flaws that have to be avoided in the performance of the bowing pattern we are interested in: first, the false attack may be clearly heard, and second, the ringing of the old string may be too high and hinder the perception of the main stroke on the new string. In contrast, in Fig. 3b, the duration of the false attack is very short (less than 1 ms), and the string vibration is well damped at the bow change. This may be the ideal situation for the minimization of false attacks II. Finally, in Fig. 3c, the release takes place before the bow change and there is consequently no false attack of type II. However, the ringing of the old string is quite high because the vibration of the main stroke is not adequately damped, which is not an ideal configuration as well.

With false attacks of type II, there is no difference between the effective duration of the attack and its theoretical expression $\frac{T}{\pi}(\arcsin(r) - \Delta\phi)$. However, we can foresee a lower acceptance for this type of false attacks for two reasons that were illustrated here. First, the conditions for the vibration development are optimal, compared to false attacks I. Second, the subsequent ringing after release will tend to make it even more noticeable, implying almost an additional stroke.

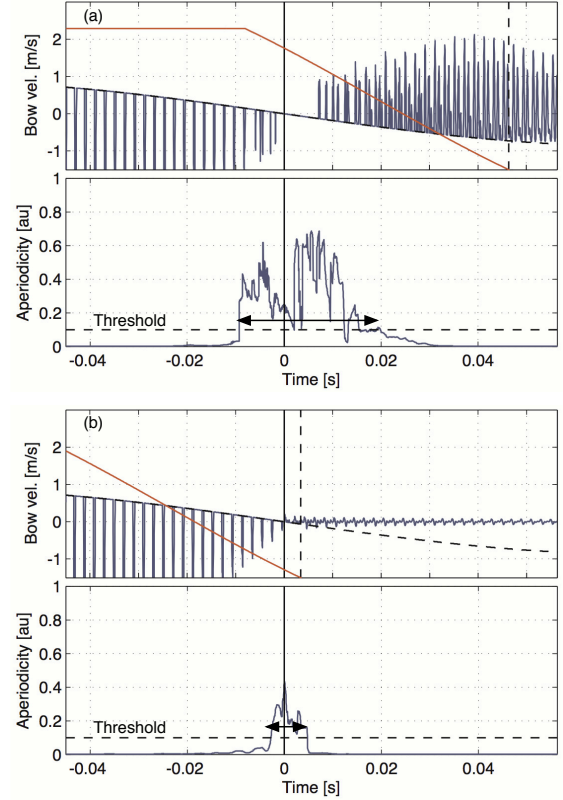


Figure 4. String velocity under the bow and related aperiodicity signal for two bow changes. (a) $r = 0.54$ and $\Delta\phi = -0.3$. (b) $r = 0.54$ and $\Delta\phi = 0.55$. A measure of the aperiodicity duration during the bow change is given by the time interval during which the aperiodicity signal is above a threshold empirically set to 0.1.

4.3 Stopping of the stroke

Beside false attacks produced in the finite string-crossing range, there is another acoustical feature that could be of interest to explain the perceptual preferences of the subjects. When the stroke is stopped with a too high bow force, the Helmholtz motion can not be maintained anymore, which results in aperiodic slips triggering with high amplitude producing a clear raucous sound at the end of the note. The high bow force may also tend to increase the time interval during which the string sticks to the bow around the bow change, between the last slip before and the first slip after the bow change, interrupting the regular slip triggering. Then, the subsequent false attack of type II may also be aperiodic. These three consequences of high bow force at the end of the main stroke suggest to take into account an acoustical feature measuring the time interval during which the vibration is aperiodic around the bow change.

In the present analysis, simulations were analyzed with YIN [7] in order to compute a measure of the aperiodicity given by the ratio between the energy of the non-periodic part of the signal and the total energy. Two examples of aperiodicity are illustrated in Fig. 4, showing the string velocity under the bow and the related aperiodicity signal for two bow changes ($\Delta\phi = -0.3$ (top) and $\Delta\phi = 0.55$ (bot-

tom)). In Fig. 4b, the Helmholtz motion is maintained until the bow change with a very short duration between the last slip before and the first slip after the bow change. In contrast, in Fig. 4a, the last slips before the bow change are clearly aperiodic, the constant sticking state at the bow change is rather long (almost 10 ms), and the first slips after the bow change are aperiodic, resulting in a rather high aperiodicity signal. An aperiodicity duration is given by the time interval during which the aperiodicity signal is above a threshold empirically set to 0.1.

4.4 Trade-off between acoustical features

In this section, we have identified three main features that could explain the perceptual preferences of the subjects when they had to choose the best combination of coordination parameters (phase $\Delta\phi$ and range r). These acoustical features are the duration of false attacks I (on the new string, before the bow change) and II (on the old string, after the bow change), and the aperiodic time interval at the end of the main stroke, representing the stopping of the main note.

There may be other features at play in the preferences of players and subjects of the perceptual tests, such as the ringing of the string when it is not in contact with the bow. This ringing can be due to the release of the main stroke before the bow change, or the release of false attacks II when the time of release is after the bow change. However, in the current study, we considered ringing as a secondary feature for two reasons. First, it is expected that limits related to over ringing are already covered by acceptance limits of other features. For example, a intense ringing after false attacks II is due to a long duration of false attacks II. Similarly, an intense ringing of the main stroke is linked to longer durations of false attacks I. Second, ringing of the string may also be a matter of taste and musical context, for example in order to achieve more or less dry note transitions. For these reasons, this acoustical feature was considered as secondary and let aside in the current analysis.

It is expected that subjects try to minimize each of the acoustical features identified before. However, they can not be all minimized at the same time, and it is more probable that the final choice results from a trade-off in which all the features are in acceptable proportion. The purpose of the next section is to discuss the results of the perceptual tests under the light of these acoustical features and obtain empirically some acceptance limits.

5. EMPIRICAL ACCEPTANCE LIMITS

First, we will quantify the acoustical features discussed before for various coordination parameters. For that purpose, durations of the false attacks and stopping of the stroke will be computed from the simulations presented in Section 3. Then, variations of these features with ranges and phases will be used to discuss the results of the perceptual tests and to determine some empirical acceptance limits for their durations.

5.1 Variations of acoustical features with coordination parameters

Fig. 5 shows the variation of some acoustical features in function of the phase difference $\Delta\phi$ for two string-crossing ranges $r = 0.05$ (left) and $r = 0.54$ (right). The panels show, from the top, the effective duration of false attacks I and II, the duration of the aperiodic time interval around the end of the main stroke, the cumulative energy during the time interval corresponding to false attacks I and II, and during the time interval from bow change to bow change, as well as the distribution of subjects' preferences during the perceptual tests described in Section 2.3.

A first remark is that there is a minimal cumulative energy around $\Delta\phi_+$ in both cases. However, if the preferred phases correspond well to this minimum for $r = 0.54$, it is not the case for $r = 0.05$, where preferred phases are greater than the phase at minimal energy. This means that subjects did not only try to minimize energy of the vibration on the unplayed string (due to false attacks and ringing), but that they took other constraints into account for judging the quality of the bowing patterns.

In this respect, the case $r = 0.05$ is quite informative. In this case, preferred phases are mostly above $\Delta\phi = 0$, where false attacks II are very short or inexistent. As a consequence, there must be another constraint explaining preferred positive values of $\Delta\phi$. The duration of the aperiodic interval around the bow change is a good candidate for this additional constraint: around $\Delta\phi = 0$, the variation of aperiodic duration is very sharp, representing a strong obstacle when one tries to decrease phase values. This may explain why subjects tended to choose greater phases.

For $r = 0.54$, the aperiodic duration is lower and shows less sharp variations around the lower limit of preferred phases. In this case, the durations of false attacks I and II may consequently play a more significant role in subjects' preferences. In particular, there is a difference between the effective and theoretical duration of false attacks I for phase differences $\Delta\phi$ between 0 and 1, which is due to the initial conditions with zero force for these attacks. This difference with the theoretical duration may also explain why there is a better tolerance for attacks of this type.

To conclude this short illustration, it is interesting to look back at the left figure, for $r = 0.05$. As mentioned before, subjects had no other choice than choosing phases above $\Delta\phi_+$, with a strong ringing of the string because the main stroke is released before being stopped. In this range, we can consequently obtain a very strong ringing of the string while minimizing false attacks II and keeping false attacks I in reasonable proportions. In contrast, this is not possible with $r = 0.54$: obtaining ringing strings during the release would imply having long durations for attacks I. As a consequence, if, for musical reasons or personal taste, a player wants to obtain a strong ringing of the strings with acceptable false attacks durations, it is better to go towards low ranges, i.e. a high inclination extent compared to the string crossing extent. This is actually what is done by the players: they tend to have more ample bow inclination movements when intense ringing is looked for.

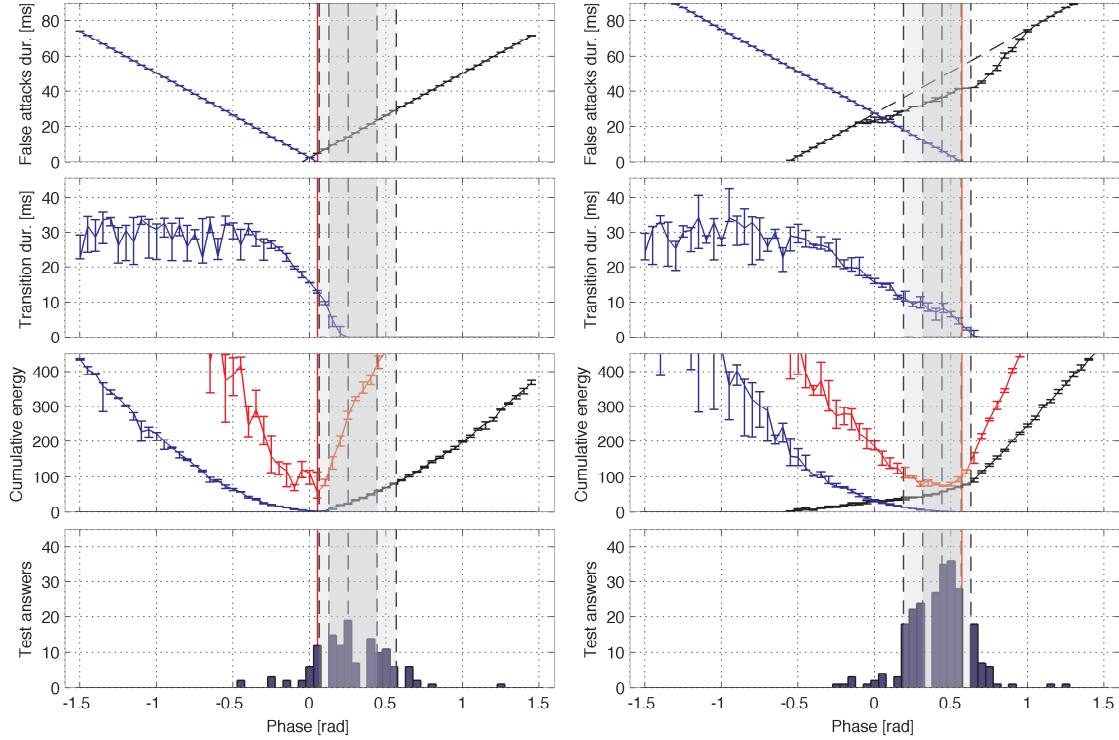


Figure 5. Acoustical features as a function of phase for $r = 0.05$ (left) and $r = 0.54$ (right). From the top: effective duration of false attacks I (black) and false attacks II (blue), with theoretical durations (oblique dashed black lines), duration of aperiodic time interval around the end of the main stroke, cumulative energy during false attacks I (black), false attacks II (blue) and during the bow change to bow change interval (red, including false attacks I and II and release). The bottom panel shows the distribution of subjects' preferences during the perceptual tests. Vertical dashed lines indicate the 10th, 25th, 50th, 75th and 90th percentile, from left to right. The red vertical line indicates the position of $\Delta\phi_+ = \arcsin(r)$, where the bow change would coincide with the moment of leaving the string-crossing range.

5.2 Acceptance limits

In the perceptual tests, the “minimal” condition gave a mean range $r = 0.6$ above which it was difficult to find good sounding note transitions. These difficulties may be related to the fact that, above this threshold, it is impossible to find a combination of phase and range that keep under acceptable limits the duration of both false attacks of type I and type II. As a consequence, the results of simulations with range $r = 0.54$ may be used to deduce some approximative limits for the duration of these two types of attacks. Taking the 10th percentile of answers as a lower limit giving an acceptance limit for false attacks of type II, and the 90th percentile of answers as a higher limit giving an acceptance limit for false attacks of type I, we find acceptance limits of 18 ms and 42 ms for type II and type I, respectively.

As discussed in the previous section, the lower limit of phase preferences for $r = 0.05$ may be explained by the duration of the aperiodic interval around bow change. Taking again the 10th percentile as a lower reference, the acceptance limit for aperiodic duration at the end of the stroke would then be around 12 ms.

Using these acceptance limits for the durations of the aperiodic interval, false attacks I and false attacks II, we can determine some phase limits for each range examined in the simulations. These limits would give an optimal area

that should be compared with the the subjects' answers. This comparison is displayed in Fig. 6. The green area shows the phase-range combinations for which the duration of false attacks I are between $\pm 25\%$ of the acceptance limit determined before (42 ms, shown with a dashed line inside the area). For phases above the dashed line, the duration is higher than 42 ms, and for phases below, it is lower. The same representation is used for the duration of false attacks II (in blue) and for the aperiodic duration (in red). For these two latter features, phases above the dashed line represents durations lower than the acceptable duration determined before.

The figure clearly shows an acceptable area for the combination of ranges and phases situated in the white region below $r = 0.6$, between the green area, on one side, and the blue and red areas, on the other side. For ranges higher than 0.6-0.7, the situation gets more confused with acceptance limits crossing each other. In this area, it is impossible to find combinations of parameters that stay within acceptable limits for all acoustical features. The results of the simulations are in a fairly good agreement with the results of the perceptual tests, shown with blue errorbars in the figure. For each range, the median value of subjects' answers is indicated with a circle, as well as the 10th, 25th, 75th and 90th percentile, on each side of the median value. For

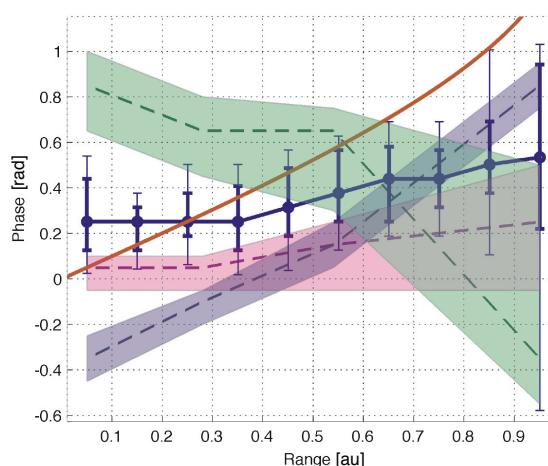


Figure 6. Empirical phase limits based on acceptance limits of acoustical features, versus string-crossing range. The colored areas show the phase values in which the simulations showed $\pm 25\%$ of the acceptance limits for false attacks II (blue), false attacks I (green) and aperiodic duration (red). The green area is an upper limit, while red and blue areas are lower limits. The errorbars display upper and lower quartiles (thick) and 10th and 90th percentiles (thin) of preferred phases during the perceptual tests. The red line shows the arcsine model where the bow change would coincide with the moment of leaving the string-crossing range.

ranges below 0.65, the preferred phases lay in the optimal region defined before, within the acceptance limits of all acoustical features. In particular, the lower limit for ranges below 0.3 is well constrained by the duration of the aperiodic period, which explains why the subjects tended to prefer phases above the theoretical arcsine model (shown with a solid red line in the figure). Between 0.15 and 0.45, it is possible to minimize the ringing by following the theoretical model, which explains why this region was considered optimal by the subjects. Above 0.65, there is no possible solutions within the acceptance limits and the answers start to be more spread, possibly reflecting personal preferences of the subjects in the minimization of acoustical features.

6. CONCLUSION

This study was intended to complement previous perceptual tests in which subjects had to adjust the coordination parameters of complex bowing patterns. The aim of the tests was to obtain good sounding note transitions simulated with a virtual violin. In this experiment, it was found that the subjects clearly preferred positive shifts between bow changes and string crossing, depending on the string-crossing range. These results were also in good agreement with measurements of real performances, in which the phase between bow velocity and bow inclination was consistently lagging between 10 and 30 degrees (0.17 to 0.52 rad). An open question remained after these perceptual tests: what were the acoustical features on which the

subjects based their preferences and was it possible to determine acceptance limits for them? In the present study, we have examined and illustrated some acoustical features that may explain these preferences: duration of false attacks before and after the bow change (due to the finite string crossing duration), and quality of the stopping of the main stroke. Based on the different experimental conditions in the perceptual tests, it was possible to deduce some empirical acceptance limits for these acoustical features. Altogether, these acceptance limits draw a clear optimal region in a phase-range diagram, which corresponds well with the region favored by the subjects.

However, these results are based on simulations for only four different string-crossing ranges, and it would be interesting to make more systematic simulations with ranges varying regularly from 0 to 1 in order to test the consistency of our results and get more accurate acceptance limits. Also, this analysis is based on a limited number of acoustical features, but other features may be at play in subjects' preferences (e.g. ringing of the string during the release, masking effects, or quality of the attack at the onset of the main stroke) and would be interesting to investigate.

7. REFERENCES

- [1] E. Schoonderwaldt, "Control and coordination of complex bowing patterns," in *Proc. of 2nd Vienna Talk*, Vienna, 2010.
- [2] E. Schoonderwaldt and E. Altenmüller, "Mastering the violin: Motor learning in complex bowing skills," in *Proc. of International Symposium on Performance Sciences*, Toronto, 2011.
- [3] E. Schoonderwaldt, M. Demoucron, E. Altenmüller, and M. Leman, "Auditory perception of note transitions in simulated complex bowing patterns," *J. Acoust. Soc. Am.*, vol. 133, no. 6, pp. 4311–4320, 2013.
- [4] M. Demoucron, "On the control of virtual violins: Physical modelling and control of bowed string instruments," Ph.D. dissertation, Université Pierre et Marie Curie (UPMC), Paris & Royal Institute of Technology (KTH), Stockholm, 2008.
- [5] M. Demoucron and N. Rasamimanana, "Score-based real-time performance with a virtual violin," in *Proc. of DAFx09*, Como, 2009.
- [6] K. Guettler, "On the creation of the Helmholtz motion in bowed strings," *Acta Acustica united with Acustica*, vol. 88, no. 6, pp. 970–985, 2002.
- [7] A. De Cheveigné and H. Kawahara, "Yin, a fundamental frequency estimator for speech and music," *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1917–1930, 2002.

Analysis of bridge mobility of violins

Benjamin Elie, François Gautier

Laboratoire d'Acoustique de l'Université du Maine
UMR CNRS 6613
benjamin.elie@univ-lemans.fr

Bertrand David

Institut Mines-Télécom
Télécom ParisTech CNRS LTCI
bertrand.david@telecom-paristech.fr

ABSTRACT

This paper focuses on the bridge mobility of violins. The mobility, or mechanical admittance, quantifies the efficiency of the instrument body to vibrate when a force is applied to the structure. The computation of the mean mobility, after the Skudrzyk's mean-value theorem, enables a global characterization of the bridge mobility. The choices made by the luthier, when he builds, restores, or adjusts an instrument, modify the mobility and the mean mobility: this is the signature of the instrument. This study shows that the bridge mobility measurement may be helpful for luthiers to objectively characterize an instrument. Two experimental applications on the violin are presented in the paper: the first one studies the characterization of the same violin in several configurations, corresponding to different positions of the soundpost. The second application studies the effect of a violin mute on both the bridge mobility and the spectral characteristics of the produced sound. This study is a part of the PAFI project, which aims to develop a set of tools dedicated to instrument makers.

1. INTRODUCTION

The role of a luthier is to propose a new instrument, or adjust settings on an existing instrument, so that it corresponds to particular wishes of a musician. He may be brought to choose materials, to replace a few pieces, to change the configuration, or to move the soundpost, for instance. All of these modifications have multiple consequences, which are, *a priori*, unpredictable, on the produced sound. The sound produced by violins is the result of several interacting subsystems: the musician, the string, the instrument body, the surrounding fluid, and the listener. The instrument body, which is itself made up of different subsystems (bridge, top and back plates, soundpost, ribs, and so on), is the element on which the luthier possesses the most important control.

In the context of lutherie assistance, the acoustician can propose tools designed for luthiers, namely methods enabling the objective characterization of the coupling between the instrument body and the string, and/or the acoustic radiation. It eventually enables to guide the choices

made by the luthier regarding the modifications of the instrument. These characterizations, when they are relevant, define the signature of the instrument. The objective of this study is to determine how the mechanical behavior of the violin can be modified by the luthier.

The coupling between the string and the instrument body is commonly studied by means of bridge mobility measurements [1]. It is known that the bridge mobility acts on the playability of the violin [2, 3], or the acoustic radiation [4, 5]. Studies on the mechanical response of the violin body reveal a salient characteristic of the violin: it exhibits an amplification of both the bridge mobility and the sound spectrum in a frequency range from 2 kHz upto 3 kHz. It is the so-called *Bridge Hill* [6–8]. More recently, the mechanical behavior has been investigated in a broader frequency range [9], using statistical analysis.

Since the bridge mobility of violins may be very complicated, due to the contributions of numerous mechanical eigenmodes, the description of the mechanical behavior, by means of the sole mobility, is not suitable for luthiers. In this study, we propose to observe the modifications of macroparameters, which globally describe this latter. The previously cited studies usually require laboratory setups, which are not suitable for a daily practice in a workshop. Since we aim to develop tools specifically designed for luthiers, the macroparameters should be adapted for them, namely able to be measured in the artisan's workshop, with affordable, robust and easily handleable devices.

The main aspects of our approach are highlighted by the organization of the paper. The features of the violin bridge mobility are studied in Section 2. It consists in computing the mean bridge mobility curves, following Skudrzyk's mean-value theorem [10]. Then, Section 3 presents two applications: the first one is a study of the influence of the position of the soundpost on the mean mobility of the violin. The second application compares the mobility of a violin in two configurations: the nominal configuration, and a configuration where a mute is attached to the bridge. The consequences of the modification in the bridge mobility, due to the mute, in the spectral characteristics of the produced sound is studied in the last section.

2. BRIDGE MOBILITY OF THE VIOLIN

2.1 Bridge mobility : modal description

Bridge mobility measurements, especially for violins, are not straightforward since the number of degrees of freedom that should be considered is large, actually equal to 6 (*cf.*

Ref. [11]). Since we want the experimental frameworks to be suitable for daily instrument maker practice, we only consider the normal motion of the soundboard. The force that is applied by the string to the bridge presents two main components: the F_x component, called the lateral force, and the F_y component, called the transverse force. According to these considerations, the coupling at the bridge can be described by a simplified admittance matrix $\mathbf{Y}_{2 \times 2}$ such that

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} (A) = \mathbf{Y}_{2 \times 2} \begin{bmatrix} F_x \\ F_y \end{bmatrix} (B), \quad (1)$$

where

$$\mathbf{Y}_{2 \times 2} = \begin{bmatrix} Y_{xx} & Y_{xy} \\ Y_{yx} & Y_{yy} \end{bmatrix}. \quad (2)$$

The excitation is given in two different polarizations, leading to two mobility curves, called the transverse mobility (impact given in the normal direction to the soundboard) and the lateral mobility (impact given on the side of the bridge). These two mobilities (denoted $Y_T(\omega) = Y_{yy}$ and $Y_L(\omega) = Y_{xx}$ for respectively the transverse and the lateral mobility) are thus the ratio between the transverse velocity $V_y(\omega)$, and the excitation forces, denoted $F_y(\omega)$ and $F_x(\omega)$. When a force is applied at a point denoted by E and the velocity is measured at point A , the mechanical admittance $Y(A, E, \omega)$ writes:

$$Y(A, E, \omega) = j\omega \sum_{k=1}^{\infty} \frac{\Phi_k(A)\Phi_k(E)}{m_k(\omega_k^2 + j\eta_k\omega_k\omega - \omega^2)}, \quad (3)$$

where Φ_k , m_k , ω_k , and η_k denote respectively the modal shape, the modal mass, the modal pulsation, and the modal loss factor of the k^{th} mode. Note that only Φ_k depends on the location.

2.2 Bridge mobility measurement: typical experimental result

A simple method to measure the mobility is to record the acceleration signal of the structure, using a small accelerometer, when it is submitted to an impulse force, performed by a small impact hammer at the driving-point location. The violin is hanged with wires, to create a free edges condition, and the strings are damped. The acceleration and force signals are obtained simultaneously, using a small accelerometer PCB 352C23 (0.2 g) and a small impact hammer PCB 086E80.

Figure 1 shows typical variations of the modulus of the transverse and the lateral mobility measured on a violin. These plots include numerous modal contributions, leading to a complicated pattern. In this section, we attempt to highlight the underlying tendency of bridge mobility curves of violins. For that purpose, the characteristic admittance, following Skudrzyk's mean-value theorem [10] is computed. We expect this descriptor to describe efficiently the mechanical behavior of violins.

The so-called characteristic admittance Y_C is defined by Skudrzyk [10] as the mobility of the structure with infinite dimensions (and consequently without any resonances). In

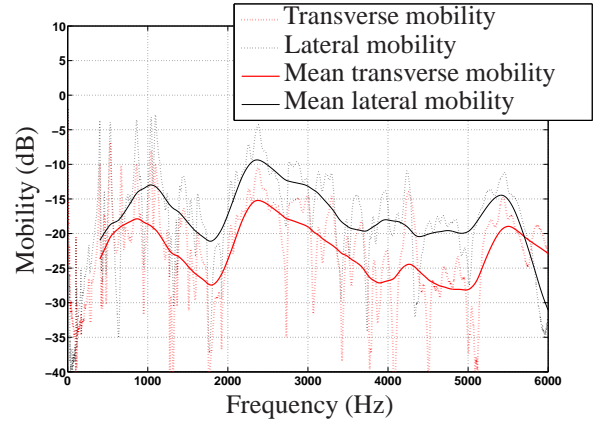


Figure 1. Mobility curve measured at a violin bridge and its corresponding mean mobility, estimated after Equation (4)

practice, its estimation is rather complicated for complex vibratory systems such as violins. Since the characteristic admittance is the mean line of the logarithmically plotted mobility curve, it can be estimated by computing the moving average of the mobility curve, in dB. It consists in computing the mean-value of the mobility, expressed in dB, contained in a sliding window of a certain span, this latter moving from a sample to the next. The obtained mean mobility, denoted by G_{CdB} , writes:

$$G_{CdB}(\omega_c) = \frac{1}{\Delta\omega} \int_{\omega_1}^{\omega_2} Y_{dB} df, \quad (4)$$

where $\Delta\omega = \omega_2 - \omega_1$, $\omega_c = \frac{\omega_1 + \omega_2}{2}$, ω_1 and ω_2 being respectively the lower and upper angular frequency bounds of the sliding window. In this paper, $\Delta\omega = 500\pi$ rad/s. Since the value is expressed in dB, a reference value is required. We chose a reference value for 0 dB corresponding to the characteristic admittance of a 2-mm thick infinite plate having the typical mechanical properties of the spruce [12]. The characteristic admittance Y_C of such rectangular flat panels is computed from the following relationship:

$$Y_{C\infty} = \frac{1}{8\sqrt{\rho h D}}, \quad (5)$$

where ρ is the mass density of the material, h and D are respectively the thickness and the bending stiffness of the structure. Thus:

$$Y_{dB} = 20 \log \left| \frac{Y}{Y_{ref}} \right|, \quad (6)$$

where Y_{Ref} is the characteristic admittance computed from Equation 5, with $\rho = 420 \text{ kg.m}^{-3}$, $h = 2 \text{ mm}$, and $D = 2.1 \text{ N.m}$. Therefore, $Y_{Ref} = 0.094 \text{ m.s}^{-1}.\text{N}^{-1}$.

Figure 1 shows a typical example of mobility curves measured at a violin bridge and their corresponding mean mobility, measured in the lateral and in the transverse directions.

The lateral mobility level is higher than the transverse mobility. At frequencies greater than 1000 Hz, the dif-

ference of magnitude between these mobilities is around 5 dB. Under 1000 Hz, the difference is smaller: the lateral and transverse mobilities are similar. The characteristic mobilities, both lateral and transverse, exhibit a few local maxima. Two of them are more noticeable: the first one is around 1000 Hz, and the other one is around 2500 Hz.

3. INFLUENCE OF MODIFICATIONS OF THE VIOLIN ON THE MEAN BRIDGE MOBILITY: POSITION OF THE SOUNDPOST

The aim of this section is to observe the variation of the mean mobility when the instrument body is modified. The example of structural modification, which is presented in this paper, is the variation of position of the soundpost. The soundpost is a piece of wood, usually a cylinder, located inside the soundbox, between the back plate and the top plate. Its location is between an f-hole and the high-pitch string side of the bridge. The position of the soundpost is considered to be an important setting for the luthier: it enables to adjust the prestress, which is applied to the soundboard. This study is an example of objective characterization of the same instrument in different configurations.

3.1 Position of the soundpost

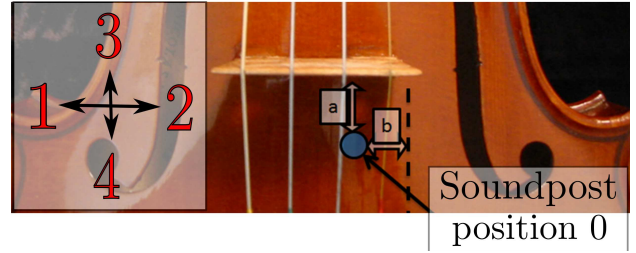
This study has been made with the collaboration of Nicolas Demarais, a violin maker. The principle of the study is to measure, and estimate the mean transverse mobility of the same violin, denoted V_1 , in 5 different configurations, corresponding to 5 different positions of the soundpost. The configurations are labeled position 0 to position 4, the position 0 being the initial configuration, namely the one that is assessed as "optimal" by the luthier. The different positions are displayed in Figure 2. For each soundpost position, the luthier assessed the instrument:

- Position 0 : optimal position
- Position 1 : favors low-frequency components
- Position 2 : favors high-frequency components
- Position 3 : rich sound, low playability
- Position 4 : poor sound, high playability

3.2 Effect on the mean mobility

The mean transverse mobility of the violin in the different configurations are displayed in Figure 3.

The global shape of the mean mobility curve is similar for the 5 configurations: they exhibit a local amplification between 500 and 2000 Hz, centered around 1000 Hz, then the mean mobility is globally increasing with the frequency. However, the mean mobility levels are not similar. For instance, the position 0 presents a mobility level much higher than other positions in the frequency range from 3000 to 4000 Hz. The difference with the mean mobility for other positions is larger than 5 dB in this frequency range.



(a)

Position	a (mm)	b (mm)
0	1.15	8.0
1	1.15	9.0
2	1.5	7.0
3	1.0	8.0
4	2.5	8.0

(b)

Figure 2. a) Initial location of the soundpost. b) Coordinates of the 5 positions of the soundpost. The distance a is the distance from the soundpost to the bridge. The distance b is the distance from the soundpost to the extremity of the bridge, denoted by a dashed line.

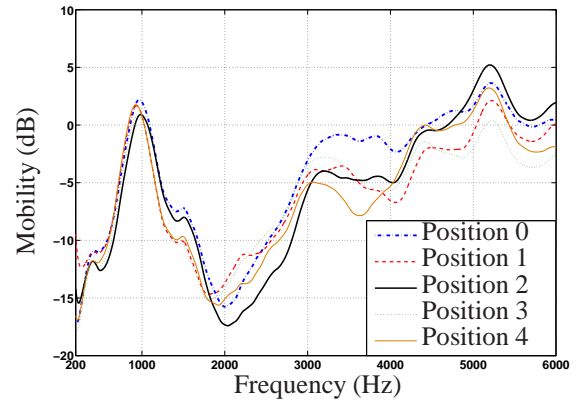


Figure 3. Mean mobility of V_1 , for the 5 different soundpost positions.

3.3 Effect on the mean mobility level

The mobility is arbitrarily sectioned into three different frequency domains:

- from 200 to 2000 Hz
- from 2000 to 4000 Hz
- 4000 to 6000 Hz

The first domain starts at 200 Hz because the lowest fundamental frequency of the violin is around 200 Hz. The upper bound of the first frequency range is set to 2000 Hz, because it basically corresponds to the end of the first amplification in the mean mobility. The choice of 4000 Hz for the other limit corresponds to the beginning of the amplification in the mean mobility, centered around 5000 Hz. The upper limit at 6 kHz corresponds to the bandwidth of the excitation signal. The structure is barely excited for

frequencies larger than 6 kHz. In each section, the mean mobility level is computed. It consists in the mean value of the mean mobility, expressed in dB, included in the frequency range. The three mean values are labeled L_1 , L_2 , and L_3 . They are displayed in Figure 4.

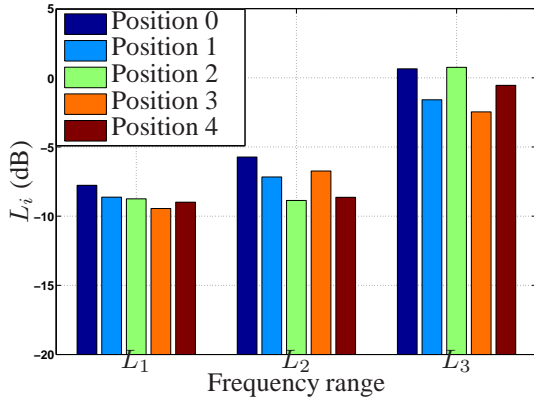


Figure 4. Mean mobility level in three frequency areas ([200 2000] Hz, [2000 4000] Hz, and [4000 6000] Hz) of the violin in the 5 different configurations

Figure 4 shows that the position 0, assessed as optimal by the luthier, corresponds to the position for which the mobility level is constantly higher than the mobility level of other positions. This suggests that the luthier chose the position of the soundpost so that the mobility is as high as possible. This probably leads to a more powerful sound. Position 1 presents a relatively large mobility level, compared with other positions, in the low frequency range, but present a small mobility level in the highest frequency range. In the highest frequency range, position 2 has a large mobility level, whereas it is not the case for L_2 . These observations confirms the remarks made by the luthier in Section 3.1: position 1 favors the low frequency components, unlike position 2, which favors high frequency components. There are no noticeable tendencies for position 3 and 4.

4. RELATIONSHIP BETWEEN MOBILITY AND PRODUCED SOUND: EFFECT OF A VIOLIN MUTE

4.1 Principle

The violin mute is a small device which is usually placed on the bridge. It aims at weakening the bridge vibration, and consequently the sound radiation. The present section deals with the effect of a particular mute (shown in Figure 5) on both the bridge mobility and on the spectral characteristics of the violin sound. The studied mute is a light rubber mute, its weight is approximately 1 g. The methods described in previous sections are applied on the same violin in both configurations: with and without the mute placed on the bridge. Note that the studied violin, labeled V_2 is a different violin than the violin V_1 studied in section 3.



Figure 5. Violin mute used for the study.

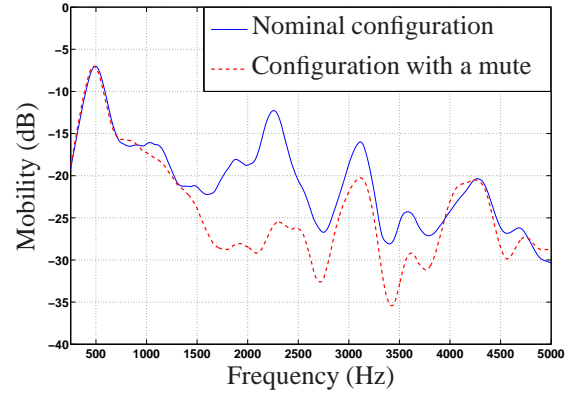


Figure 6. Mean lateral mobility measured at the bridge of the violin V_2 for two configurations. The solid line represents the mobility of the violin on the nominal configuration (without the mute), the dashed line is the mobility measured on the violin with a mute attached to its bridge.

4.1.1 Effect of the mute on the mobility curves

The lateral mobility of the violin without the mute, and then with the mute, have been measured following the protocol described in section 2.2. Figure 6 shows the mean mobility curves for both violin configurations.

The effect of the mute on the bridge mobility can be clearly seen. Although, the low frequency range is hardly modified, namely the mobility curves are fairly similar, in the frequency range from 1000 Hz upto 2500 Hz, the mean mobility curve in the mute configuration is much lower than the one of the normal configuration. The difference spanning values from 5 to 13 dB. Then, in the high frequency range, the difference is no longer significant.

4.2 Effect of the mute on the produced sound

4.2.1 Sound analysis

In this section, we study the effect of the violin mute on the produced sound. Previous section showed that the mute had an important effect on the mean mobility. To study the effect on the produced sound, we assume that the instrument acts like a source-filter system, where the source is composed of the musician and the string, and the filter is the instrument body. In a source-filter model, the output

signal $y(t)$ is seen as the convolution of an input signal $x(t)$ convolved by a filter impulse response of a linear invariant system $h(t)$. Hence:

$$y(t) = x(t) * h(t). \quad (7)$$

If the source $x(t)$ is harmonic, like violin strings, the output signal $y(t)$ is also harmonic, but its spectrum $Y(\omega)$ is then disturbed by the frequency response of the filter $H(\omega)$. Indeed, the low variations of $Y(\omega)$ (i.e. its spectral envelope) follows the shape of $H(\omega)$. The resonances of $H(\omega)$ are therefore detectable in $Y(\omega)$ by looking at its spectral envelope. The broad peaks in the spectral envelope of a source-filter model output signal are commonly named *formants* [13, 14]. In speech acoustics, the source-filter model is mainly used to detect formants [13], responsible of vowel detections [15, 16]. We chose to use the *Linear Predictive Coding* (LPC) [17] method, because of its wide use in speech signal processing. The model order is the number of poles in $h(t)$, i. e. twice the number of resonances or maxima residing in the spectral envelope. In our case, $h(t)$ contains numerous poles, corresponding to each body mode. The aim of this analysis is to emphasize the averaged mobility in the sound spectrum. This latter exhibits a few maxima, basically 3 or 4, in our frequency range of interest (0-4500 Hz). Consequently, the number of poles should be slightly greater than twice the number of maxima. The LPC order is then set to $p = 10$.

4.2.2 Experimental protocol

The analyzed signals are obtained by recording the near-field sound pressure signal, using a microphone. The measurements are done in an anechoic chamber. During experiments, a musician is asked to play glissandi on the lowest string from D (292.5 Hz) to G (195 Hz). Experiments are done on violin V_2 in two different configurations: with and without the mute. The sampling frequency F_s is 11050 Hz, and the recorded signals are 10 seconds long.

4.2.3 Results

Figure 7 shows the long-time averaged spectral envelope of glissandi sound pressure signals recorded on the violin with and without the mute violin. The long-time averaged spectral envelope is computed by averaging the spectral envelope, obtained via LPC, of each temporal frame of the glissando signal. The effect of the violin mute on the spectral envelope is very similar to the one on the mean mobility: the second formant, between 2000 and 3000 Hz, disappears in the mute configuration.

It is worth noting that, in the mute configuration, the weakened part of the bridge mobility, and consequently that of the envelope curves, corresponds to a frequency area where a formant usually occurs in most violins. It is often referred as the "Bridge Hill" in the scientific literature [6–8]. It also corresponds to the frequency range in which the human ear is the most sensitive. The main effect of the mute is to weaken the sound level of components residing in the most sensitive frequency area of the human ear.

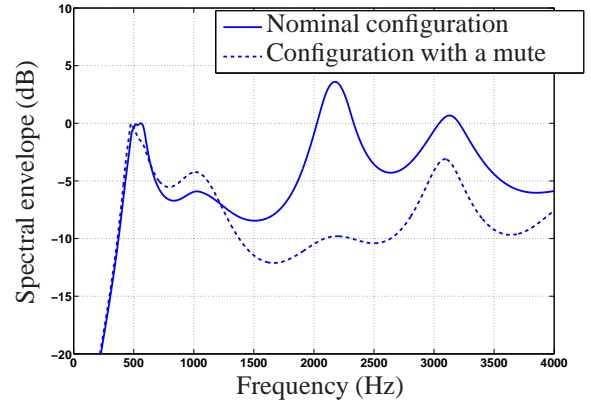


Figure 7. Long-time averaged spectral envelope of the sound pressure glissando signal recorded on the violin V_2 in its nominal configuration (without the mute) and with a muter attached to the bridge.

5. CONCLUSIONS

The study presented in this paper analyzes the bridge mobility of violins. Since the bridge mobility is the sum of numerous modal contributions, leading to a complicated pattern, the analysis is performed via global parameters, which describes its underlying tendency. This global parameter is called the mean mobility.

The mean mobility curves of several violins exhibit a common feature: two local amplifications are constantly present in the mean mobility curve. The first one is located between 500 and 1000 Hz, the second one is located between 2000 and 3000 Hz. The second one is commonly called *Bridge Hill* in the scientific literature.

An example of application showed that the mean mobility level can be adjusted by the luthier. For instance, the position of the soundpost may change the level of the mean mobility upto 5 dB, in a certain frequency area. In the presented example, the luthier seems to choose the position of the soundpost so that the mean mobility level is the largest, probably leading to a more powerful sound.

Another application focused on the effect of a violin mute on the mean mobility, and its consequences on the produced sound. The spectral envelope of the sound produced by a violin, in its normal configuration, exhibits two maxima, called *formants*, corresponding to the local maxima of the mean mobility. The lateral mobility can thus be considered as an acoustic signature of the instrument. When the studied mute is attached to the violin bridge, the mean mobility in the frequency range 2-3 kHz is weakened: the existing formant in this frequency range, in the nominal configuration, disappears. This suggests that the distribution of the amplitude of harmonics in the violin sound spectra is controlled by the lateral mobility of the violin.

To summarize, the lateral bridge mobility of violins is an important quantity to take in account when the luthier makes, or sets an instrument. The PAFI project, to which belongs this study, developed tools and methods specifically designed for luthiers in order to be able to measure it by their own. A peer data system will enable the luthiers to

share information, and by their own experience, will eventually improve their savoir-faire in making instruments.

Acknowledgments

The authors would like to acknowledge the *Agence Nationale pour la Recherche* for the financial support of projet PAFI (*Plateforme d'Aide à la Fabrication Instrumentale*), and Nicolas Demarais for his participation, and the loan of the instrument.

6. REFERENCES

- [1] J. Moral and E. V. Jansson, "Eigenmodes and input admittances and the function of the violin," *Acustica*, vol. 50, pp. 329–337, 1982.
- [2] J. C. Schelleng, "The violin as a circuit," *J. Acoust. Soc. Am.*, vol. 35(3), no. 3, pp. 326–338, 1963.
- [3] J. Woodhouse, "On the playability of violins. part II: Minimum bow force and transients," *Applied Acoustics*, vol. 78, pp. 137–153, 1993.
- [4] G. Bissinger, "Structural acoustics model of the violin radiativity profile," *J. Acoust. Soc. Am.*, vol. 124(6), no. 6, pp. 4013–4023, 2008.
- [5] G. Weinreich, C. Holmes, and M. Mellody, "Air-wood coupling and the swiss-cheese violin," *J. Acoust. Soc. Am.*, vol. 108(5), pp. 2389–2402, 2000.
- [6] F. Durup and E. V. Jansson, "The quest of the violin bridge-hill," *Acta Acustica United with Acustica*, vol. 91, pp. 206–213, 2005.
- [7] G. Bissinger, "Structural acoustics of good and bad violins," *J. Acoust. Soc. Am.*, vol. 124(3), pp. 1764–1773, 2008.
- [8] J. Woodhouse, "On the bridge hill of the violin," *Acta Acustica United with Acustica*, vol. 91, pp. 155–165, 2005.
- [9] J. Woodhouse and R. S. Langley, "Interpreting the input admittance of violins and guitars," *Acta Acustica United with Acustica*, vol. 98(4), pp. 611–628, 2012.
- [10] E. Skudrzyk, "The mean value method of predicting the dynamic response of complex vibrators," *J. Acoust. Soc. Am.*, vol. 67(4), pp. 1105–1135, 1980.
- [11] X. Boutillon and G. Weinreich, "Three-dimensional mechanical admittance : theory and new measurement method applied to the violin bridge," *J. Acoust. Soc. Am.*, vol. 105(6), pp. 3524–3533, 1999.
- [12] B. Elie, F. Gautier, and B. David, "Macro parameters describing the mechanical behavior of classical guitars," *J. Acoust. Soc. Am.*, vol. 132(6).
- [13] G. Fant, *Acoustic theory of speech production*. Mouton, The Hague, 1960.
- [14] A. H. Benade, *Fundamentals of musical acoustics*. Oxford University Press, London, 1976.
- [15] R. K. Potter and J. C. Steinberg, "Toward the specification of speech," *J. Acoust. Soc. Am.*, vol. 22(6), pp. 807–820, 1950.
- [16] P. Ladefoged, *Preliminary of Linguistic Phonetics*. University of Chicago Press, Chicago, 1971.
- [17] S. McCandless, "An algorithm for automatic formant extraction using linear prediction spectra," *IEEE Trans.*, vol. 22, pp. 135–141, 1974.

STRING “AFTER-LENGTH” AND THE CELLO TAILPIECE: ACOUSTICS AND PERCEPTION

Eric Fouilhé

Maker

Laboratoire de Mécanique et Génie Civil de
Montpellier
ericf26@gmail.com

Anne Houssay

Laboratoire de Recherche et de Restauration
Musée de la musique
Cité de la musique / Paris
houssay.a@gmail.com

ABSTRACT

In a long term research on cello tailpieces, we have first identified the vibrating modes of a cello tailpiece mounted on a Dead Rig [1], and have worked on the possible influence of the wood on these modes. Among musicians and violin makers, several empirical theories exist about an ideal “after-length”, i.e. the distance of the tailpiece to the bridge which leaves a small length of vibrating string. Here we describe on the parameters involved when varying the “after-length”, and explore the influence of the position of the tailpiece on the modes and on the sound.

1. INTRODUCTION

On a modern cello, the tailpiece is where the four strings are attached. The tailpiece (C, fig 1) has one attachment at each end, and the setting is more or less standardized in its 3 lengths.: We call the “after-length” the distance B between the bridge (a) and the tailpiece (C fig.1). On the other side, it is held by the tail-cord (D fig.1) which passes around a saddle (d), and fixed by a loop around the cello endpin.

Empirical theories declare that the after-length should be $1/6^{\text{th}}$ of the playing length of the string.

Our question is whether varying the three lengths B, C and D changes the motion of the tailpiece and the sound of the instrument. We attempt to make connections between acoustic measurements and the perception of sound by sound perception experts when varying the position and length of the cello tailpiece.

The three components B, C and D are interdependent in the “tailpiece chain” BCD. In order to isolate the *after-length* question, we used an adjustable tailpiece to characterize the influence of this *after-length* compared with the influence of the tail cord and of the tailpiece’s length on the tailpieces modes on a Dead Rig and then on the vibration modes of a cello.

2. MATERIAL AND METHODS

2.1 Material

Two modern standard tailpieces in ebony and one adjustable tailpiece in African blackwood were used:

- Tailpiece T.1: ebony (*Diospyros*), 62 g, length 235 mm.
- Tailpiece T.2: ebony (*Diospyros*), 62 g, length 250 mm.
- Tailpiece T.3: African blackwood (*Dalbergia*), 76 g, 235 mm, with double system of attachment both baroque and modern, two possible *after-cords*: one of standard length, the other with an extension of 23 mm (“baroque type” attachment) (Figure 2).

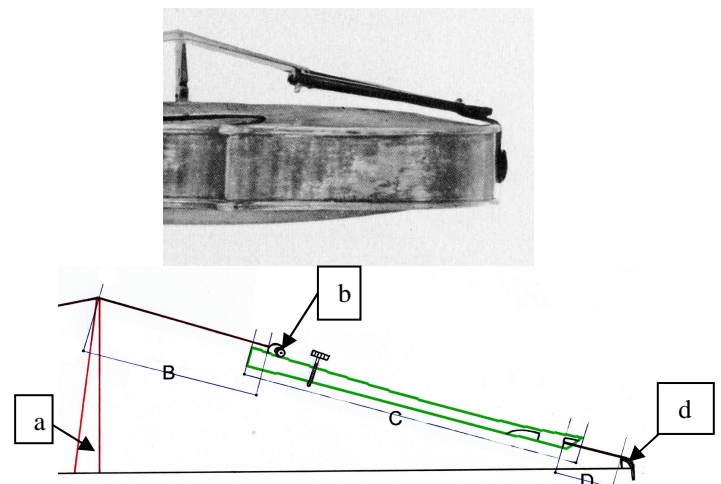


Figure 1. Profile of (from left to right) a: bridge, B: after-length, C tailpiece length with its stopping nut b, D:tail-cord length, saddle d. (Above, profile of the Stradivari violin “l’Aiglon”.

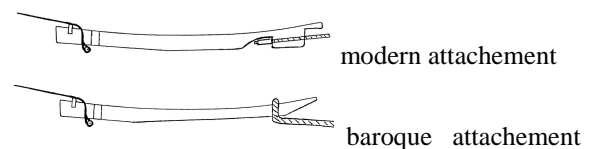


Figure 2. Tailpiece T.3, of length 235 mm, with two possible attachments. The attachments are made in composite fibers and do not stretch significantly with 60 kg tension of the cello strings.

The Dead Rig is made of a strong metallic beam with no resonance at the modal frequencies we are studying, see [1]. It holds a string length and a fixed bridge at the same angle as on the cello, and an end attachment with the same dimensions as on the cello.

The violoncello is a good student cello made in Mirecourt in 1930.

As described in our papers on the modal analysis of the cello tailpiece [1], we used a 2g PCB impact hammer, a PCB uniaxial Accelerometer and George Stoppani's Software. George Stoppani's modal analysis software, gives the FRF to the hammering of specific points on the tailpiece, amplitude, damping, and an animated visualization of the mode shapes of the tailpiece and / or the cello.

2.2 Methods

On both the Dead Rig and on the cello, we have stretched and tuned cello strings with different *after-length* configurations of tailpieces. Three different methods were used: modal analysis of the tailpieces under tension, Bridge Admittance measurements on the cello, and sound perception of the cello played three expert listeners: the player and two violin makers.

The bridge admittance of the cello was measured by hammering at the treble side of the bridge, the response of the body was measured with the accelerometer at the top bass side of the bridge: it gives an RMS where the A0 air mode and the B1- and B1+ torsion modes of the cello are identified with their frequencies and modes shapes.

Repeatability is achieved in the modal analysis of the tailpieces on the Dead Rig with a mean of 300 measurements taken, and in the 10 admittance measurements of the cello's bridge for each tailpiece Set-up.

Perception learning by musicians leads to an expert type of listening different from that of non musician [2]. From the point of view of neurosciences, there is a change in cortex that is linked to learning experiences in both visual arts and music [3] and the training of the ear leads to a change in hearing perception and cognition [4]. Thus, we argue that musicians but also some instrument makers have developed an expert type of perception, and this leads us to chose relevant experts: a professional cellist and two violin makers.

A chromatic scale, extracts of Bach's suite n°1 in G Major and of Brahms Sonata N°1 in E minor was played. Each expert listener expressed their perception of the instrument's power, balance, tonal quality, and dynamic range, precision of attack and wolf note. The latest is generally found on most cellos between E and G, on one or more strings.

Quantitative appreciation of the qualitative judgments, the semantic diversity of the terms employed by the musician and makers, and the influence of the room used for the test have their importance, and our protocol was chosen similar to one used by a violin maker in his workshop: The musician made her comments first in order to avoid the influence of the listeners. The maker who was not involved in the experiment spoke second. The third expert took notes of the comments and asked for explana-

tion to get a more precise idea of what the expert meant to say, each expert being free in which order he reacted about the different parameters, because constraint would have spoiled their first immediate perception. Effectively, perception work has to take into account short and long term sound memory and comparisons two by two with short term perception memory of the experts was privileged.

3. RESULTS

3.1 Modal Analysis of the Tailpieces on the Dead Rig :

We compare seven different settings (Figure 3) on the Dead Rig giving four different *after-lengths*: 95, 115, 116, 128 mm (A standard *after-length* used today for cellos is 115 - 117 mm.)

For this, we use

- 3 different tailpiece lengths : tailpiece T.1: 62 g, length 235 mm. tailpiece T.2: Ebony, 62 g, length 250 mm, *after-length* 116 mm. tailpiece T.3 with two possible *after-lengths*.

- Six different tail-cord lengths : 51 mm, 39 mm, 30 mm, 28 mm, 18 mm, 10 mm.

These Set-ups are summarized in Figure 3.

Type of setting	After-length	Tailpiece	Tail cord
Set-up 1	116 mm Standard	235mm T.1 Standard	30mm Standard
Set-up 2	95 mm Short	250mm T.2 Long	30mm Standard
Set-up 3	116mm Standard	235mm Standard T.3 adjustable	51 mm Long
Set-up 4	128mm Long	235mm Standard T.3 adjustable	39mm Long
Set-up 5	128mm Long	235mm Standard T.3 adjustable	18 mm Short
Set-up 6	116 mm Standard	235mm Standard T.3 adjustable	28mm Standard
Set-up 7	115mm Standard	250mm T.2 Long	10mm Very short

Figure 3. Types of settings for the experiments: colors are used for reading the following curves.

3.1.1 Comparison of Standard and Short After-length

T.1 (standard) and T.2 (long) are compared with two different *after-lengths*: normal (116 mm: Set-up 1) or short (95 mm: Set-up 2) and the same standard tail-cord length.

The first modes of the tailpiece on the Dead Rig, in group I and group II (Figure 4) as we have shown in previous articles [1] are rigid Body Modes. On the Dead Rig, we get the same frequencies for a normal Set-up and for a smaller *after-length*.

In group III (the bending and torsion modes), the RMS shows strong differences between #6 (Bending 1) and #7 (Torsion 1): the longer tailpiece has a lower frequency for the flexion and torsion modes (beam modes) indicating a greater flexibility of the tailpiece itself giving them more amplitude. The shorter tailpiece and longer *after-length* have lower energy in amplitude.

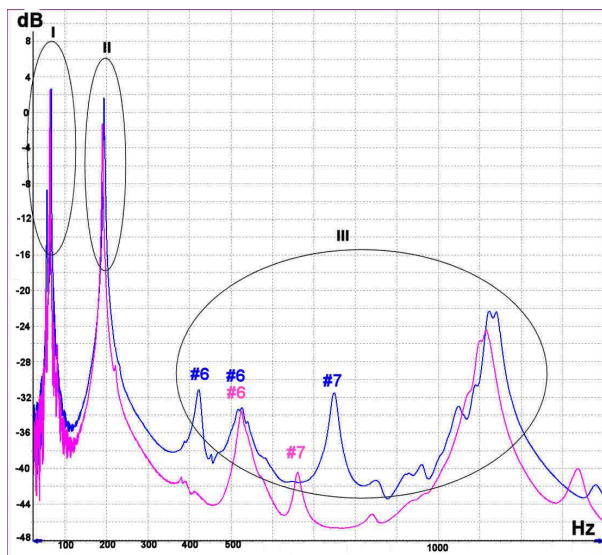


Figure 4. Test 1 : Modal analysis of the tailpieces T1 and T2 on Dead Rig, RMS of Set-up 1 (pink) and Set-up 2 (blue).

Here, we see that the modal differences are due to the length of the tailpiece, which is less flexible in the shorter tailpiece (Set-up 1) in the bending and torsion modes, rather than effects coming from the *after-length* modification.

3.1.2 Comparison of three different tail cord lengths

We compare the Set-up 3, 4 and 5 on the same Adjustable Tailpiece T.3 mounted on the Dead Rig (Figure 5).

- Set-up 3: standard *after-length*, long tail cord.
- Set-up 4: long *after-length*, standard tail cord.
- Set-up 5: long *after-length*, short tail cord.

3.1.2.1 Effects on Swing and Rotating Modes (Group I):

For modes #1, #2, #3 (the swing and rotating mode (see [1]) around 57 Hz) we have shown previously that the tailpiece has rigid body modes and that it swings and rotates on its attachments. The three Set-up show the same frequencies. Amplitudes of Set-up 3 with standard *after-length* and a long tail cord is significantly above the two others (8.2 dB, 9 dB).

3.1.2.2 Effects on See-saw Modes (Group II):

For modes #4, #5, we found previously that the tailpiece on the Dead Rig present see-saw balancing modes around 200 Hz: Here, frequency rises +10% between Set-up 3 and Set-up 4 with a diminution of 23% of the tail cord length. Frequency rises + 25 % for a diminution of 54% of tail cord. The increase in frequency is significant with the diminution of the tail cord.

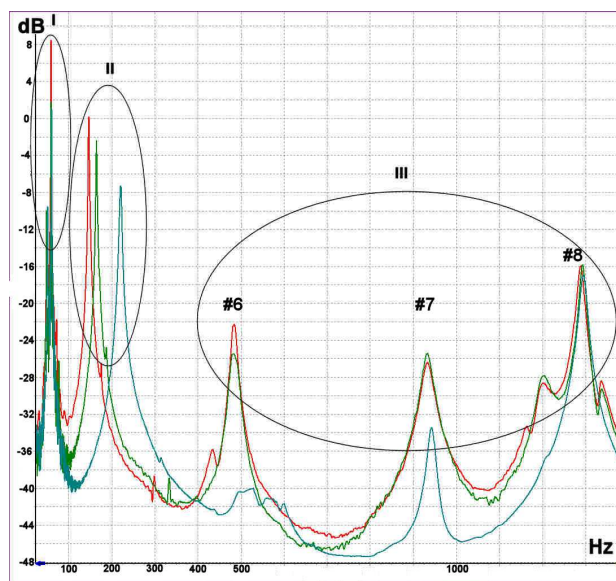


Figure 5. Modal analysis of tailpiece T.3 on Dead-Rig: Long tail cord: Set-up 3 (red), standard tail cord: Set-up 4 (green), short tail cord: Set-up 5 (turquoise).

3.1.2.3 Effects on Bending and Torsion Modes (Group III):

Modes #6, #7, #8, of the tailpiece on the Dead Rig were found to be bending and torsion modes just above 450 Hz: Here, frequencies are very similar for Set-ups 3, 4 and 5 of the adjustable tailpiece.

Amplitudes are similar for Set-up 3 and Set-up 4, while the amplitude of Set-up 5 is lower in the middle range: the tail cord being very short damps the #7 bending and torsion mode.

Thus, it is more the length of the tail cord that affects these modes, and not so much the *after-length*. When the tail cord is shorter, in Set-up 5, the movement is damped in amplitude but the flexibility of the tailpiece itself is not much affected.

3.2 Modal analyses of the tailpieces on cello

Bridge Admittances give us the principal Body Modes of the cello mounted with different Set-up. We compare the effects of two different *after-lengths* and then of three different after cord lengths on the cello to compare them later with the tailpiece modal analysis on the Dead Rig results.

3.2.1 Comparison of standard and short after-lengths

T.1 (standard) and T.2 (long) are compared with two different *after-lengths*: normal (116 mm) or short (95 mm,) a shortening of 21 mm (-15, 8%) between the two; the same standard tail-cord length is used. We explore the coupling of the tailpiece with the cello Body Modes extracted from Bridge Admittance measurements (Figure 6).



Figure 6. Cello Body Modes for standard Set-up 1 (pink) and Set-up 2 (blue) with long tailpiece T2 and shorter after length. Same tail cord length for both.

The difference between the standard Set-up and the shorter *after-length* is very slight:

Near A0, two peaks are visible: the lower in frequency is connected with the first modes of the tailpiece (group I figure 4); the second corresponds to the first air mode of this cello A0 which is around ... There is a deep split which corresponds to the coupling of the tailpiece and the cello, and it is similar in both Set-ups. However, when shortening the *after-length* from Set-up 1 to Set-up 2, i.e. from 116 mm to 95 mm (-15, 8%), the first cello air mode A0 rounded peak on the right is raised only 2% in frequency and decreases (-2 dB) in amplitude, while the tailpiece peak on the left is also raised about the same amount with a slightly higher amplitude.

Body Cello Mode B1- shows even less difference between the two Set-ups: an increase of 1, 9 % in frequency, and similar amplitudes.

Body Cello Mode B1+, as the *after-length* is shortened goes up 2, 5% in frequency and is more separated in three different peaks. The standard Set-up has higher amplitude on B1+. This is the most affected cello Body Mode.

3.2.2 Comparison of three lengths of tail cord

The tail cord lengths seem to be of relatively greater importance as we have seen in preceding comparisons (see 3.1.2). We compared the following set ups on the cello (figure 7):

- Set-up 4 is the adjustable tailpiece with long tail cord and long *after-length*.
- Set-up 5 is the adjustable tailpiece with shorter tail cord and long *after-length*.
- Set-up 2 is a longer tailpiece with standard tail cord and short *after-length*.

We can see that the cavity mode A0, and the two main Body Modes B1- and B1+ of the cello change with the different settings of the tailpieces. The two normal length tailpieces (Set-up 2 green and 5 turquoise) do not react in the same way, which show the importance of their attachments.

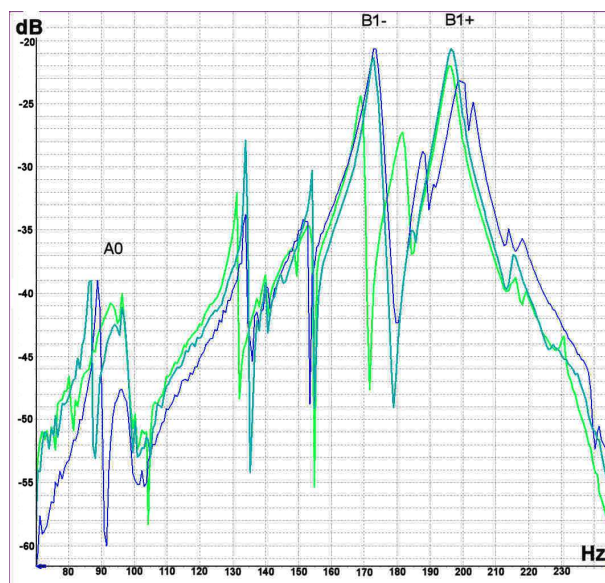


Figure 7. Cello Body Modes for long tail cord: Set-up 4 (green), standard tail cord: Set-up 5 (turquoise), and short tail cord: Set-up 2 (blue) with long tailpiece T2 and shorter after length.

3.2.2.1 Effects of tail cord length on the Air Mode A0: (92-95 Hz)

With Set-up 2, the frequency of mode #3 of the tailpiece itself (around 75 Hz), is distinct from cello A0 frequency (around 92 Hz). The amplitude of A0 lowers very little (-2 dB) when the tail gut is shortened, from Set-up 4 to Set-up, 5. But from a long to a short tail cord, the cello Body Mode A0 peak splits in two, indicating a coupling interaction of the tailpiece mode #3 with the cello mode A0 when shortening the tail cord. The split of A0 is even more striking, and the amplitudes diminish even more (- 8 dB) when the *after-length* and the tailcord both get smaller with a longer tailpiece in Set-up 2.

3.2.2.2 Effects of tail cord length on the cello Body Mode B1-: (169-173 Hz)

The B1- peak is split for Set-up 4 (*after-length* and tail cord longer), a new peak appears at 181 Hz between B1- and B1+ showing a strong effect of the coupling. There, the amplitude is minimum (- 4 dB). Set-up 2 and 5 (*after-length* and tail cord smaller) have a very clear and strong B1- peak.

3.2.2.3 Effects of tail cord length on the cello Body Mode B1+: (195-199 Hz)

B1+ peak is high and clear for Set-up 5 (Turquoise) and Set-up 4 (green) although a little lower in amplitude for the latter.

Mode B1+, with Set-up 2 (blue), is losing a little amplitude, leaving a main peak with less amplitude and one smaller peak on each side. This indicates the coupling of one or two modes of the tailpiece and a consequence of the split of A0 on A1 (which on cello, is just below B1+). In this Set-up, the tail cord is very small, attaching more firmly the bottom of the tail to the body of the instrument.

It seems that while adjusting the Set-up of the tailpiece, the main modes of the cello can be coupled with modes of the tailpiece. This is shown by the split of the peaks into different peaks which lowers the amplitude of the main Body Mode. Set-up 4 with long *after-length* and long tail cord splits dramatically A0 and B1-. Set-up 5 where the tail cord is very small, attaching more firmly the bottom of the tail to the body of the instrument splits B1+ in three.

3.3 Perception results

The cello used for this test is usually qualified as powerful, open, and slightly more powerful and hollow towards the treble, with a clear sound (in a sense of a lack of roundness). It has a strong wolf note on the F# on the G string.

3.3.1 Tailpiece T.3- adjustable tail cord length:

When lengthening the *after-length* from Set-up 6 to Set-up 5, not much change is noticed. The sound gets a little more precise, the tone nicer, with a little unbalanced treble and a more metallic sound; less dynamic for 3, from small to extra long, the balance between bass and treble is better, treble notes are better and have a larger dynamic range. The whole sounds better, with the basses more open and global resonance also, the precision of attack remaining. However, the wolf note is now strong. The wolf note was reduced with a medium tail cord length.

3.3.2 Tailpiece T.2- Short to Standard After-length:

When lengthening the *after-length* while diminishing the tail cord, from Set-up 2 to Set-up 7, the main sound gets better, from a powerful and metallic character with attack difficulties and unbalanced trebles, towards much better basses and trebles, easy playing and a lesser wolf note. The instrument is more difficult to play but globally has a better tone.

4 SYNTHESIS

With three different approaches: modal analysis of tailpieces under tension on a Dead Rig, comparison of the main Body Modes of the cello obtained from the Bridge Admittance measurements, and musical perception of the instrument, we have tried to isolate the effects of the variability of the *after-length*. Even though this dimension is linked to that of the tailpiece and of the tail cord, we have isolated this parameter by using artifacts, such as an alternate use of baroque type or modern type of attachment on same adjustable tailpiece, and the use of tailpieces of different lengths. The analysis is then approximate, and is getting more precise with other complementary tests which are not mentioned in this study, but are included in a PhD in progress. The perceptive analyses remain modest and only qualitative in order to confront dynamic mechanical measurements with perception for each Set-up

4.1 Swing and Rotating Modes (Group I) and A0:

Group I is the group of the three first modes of the tailpiece described by Stough [5] and by Fouilhé and al. [1]. They have a strong amplitude around the frequency of the lowest string of the cello ($C = 65.4$ Hz) and are linked together, however, they do not produce any perceived sound because A0 acts as a filter of lower frequencies.

When both the *after-length* and the tail cord get smaller with a longer tailpiece in Set-up 2, the peak of B1+ is split, which indicate a coupling of some kind.

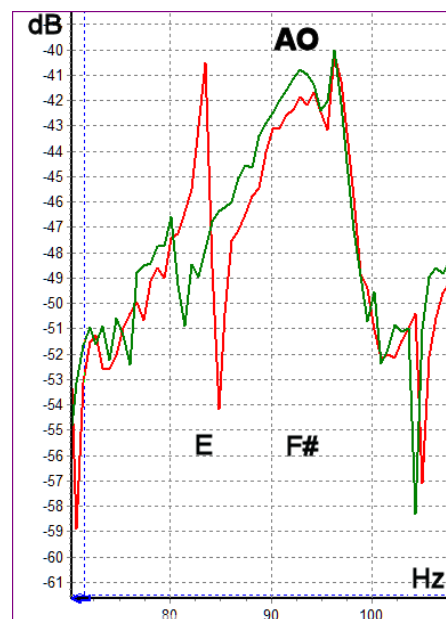


Figure 8: Cello Body Mode A0 (from the Bridge Admittance) for Set-up 3 (red) and Set-up 4 (green).

The large red peak of Set-up 3 seen at 82 Hz on the left of the cello's A0 (93 Hz) (Figure 8) correspond to Tailpiece mode #3 at 75 Hz on the Dead Rig. This can be proved when damping with the hand the vibrations of the head of the tailpiece, then the peak at 83 Hz disappears, leaving a single A0 peak.

A0 and #3 could be coupled in some Set-up, for instance in Set-up 3 but not in Set-up 4.

However, when the *after-length* or tail cord are varied Group I is not affected in frequency but only slightly in amplitude.

As early as 1819, Felix Savart who was working with Vuillaume mentioned the importance of the tuning of A0 with other modes [7]. Carleen Hutchins [8] and Jim Woodhouse [9] confirmed the importance of the tuning of this mode. Recently, Bissinger has shown a correlation between the amplitude of A0 and the tone quality of violins [10].

Thus we have shown that the coupling of the tailpiece with A0 divides the peak of the resonance of the instrument, thus sharing the energy between A0 and that mode. This lessens the sound quality especially in the bass register. It is thus preferable to un-tune the tailpiece from the A0 by means of setting up *after-length* and tail cord.

4.2 Coupling of the Tailpiece's See-saw Modes (Group II) and the Cello's Main Body Modes B1- and B1+

Two rigid Body Modes, #4 and #5, belong to group II and have been described by Stough [5] and Foulhé [1] (Mode #4=Rh), they are see-saw modes. Their frequency is near B1- and B1+ below 200 Hz, in a range where a coupling interaction will influence the tone.

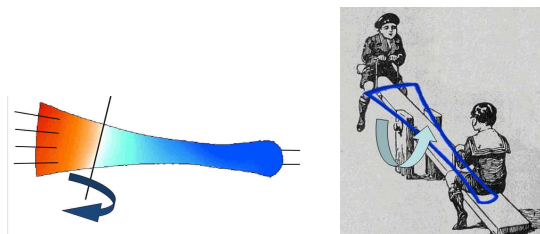


Figure 9: Mode #4 of the tailpiece under tension : strong see-saw motion, see [1].

Because of the asymmetric lever which is shorter towards the head of the tailpiece (in red Figure 9), because the tailcord firmly holds it on the lower end (in blue), Mode #4 is little affected by the *after-length* but more by the length of the tail cord.

In Set-up 3, 4 and 2, the tail cord gets progressively smaller at each Set-up; the end of the tailpiece is maintained progressively stiffer near the saddle. The consequence is an increase in frequency and lowering amplitude of Mode #4 (figure 5). The perception is that the cello's sound is powerful, richer in harmonics, but more demanding in the emission.

In Set-up 4, when the tail cord is at maximum length, if Mode #4 gets below the frequency of B1-, B1+ splits in two peaks (Figure 7). The sound of the instrument is modified, and there is a diminution or even disappearance of the wolf note. The sound is milder, less powerful, less aggressive, lower harmonics are lost, the general tone has less character, and the articulation is less precise under the bow.

The tail cord can thus be adjusted between this two extremes but one can expect that it is dependant also on the weight of the tailpiece.

5 CONCLUSIONS

The importance of the lengths of the "chain" = *after-length* + tailpiece length + after cord has been described with modal analysis and related to tonal adjustment.

Variations in *after-length* from standard to smaller *after-length* do not significantly affect the tailpiece modes frequencies measured on the Dead Rig, nor the Body Modes of the cello on which the Set-ups were tried, except on the B1+ whose frequency was raised 2,5% with a -15, 8% *after-length*. The *after-length* has been found to be more sensitive to diminution than to increase around the standard length.

The changes in the standard tailpiece lengths (of 116 mm \pm 5 mm) did not affect sensibly the frequencies of the Cello Body Modes nor the perception of the tone, except where the flexibility of the tailpiece itself is involved.

It is more in the variations of the tail cord that differences were measured. Frequency rises of + 25 % for a diminution of 54% of tail cord have been noted. The increase in frequency is significant with the diminution of the tail cord, and these changes were related to perception changes.

It is known that the air mode of the cello A0 is important for the quality of lower tones: The higher in frequency and the steeper is the A0 peak, the quicker there is saturation when pushing the string hard with the bow. On the opposite, when the A0 peak is moved and widened towards lower frequencies, the general tone of the instrument is lower, and the bow can be pressed harder.

Here, we have found how tailpiece adjustments can be used to move A0 in order to enhance these effects when desirable, as well as how it acts on the wolf note.

Other factors are to be associated like weight and wood variations, and have as much importance in the tonal adjustments of the cello.

6 REFERENCES

- [1] E. Foulhé, G. Goli, A. Houssay, G. Stoppani, "Vibration modes of the cello tailpiece", in *Archives of Acoustics*, 36, 4, 2011, pp.713-726. <http://acoustics.ippt.gov.pl/> DOI: 10.2478/v10168-011-0048-2.
- [2] Jeremy L. Loebach, Christopher M. Conway, and David B. Pisoni, "Audition: Cognitive influences", in Bruce Goldstein ed., *Encyclopedia of Perception*, SAGE, 2010, Volume 1, p.138-141.
- [3] Andrew J King & Israel Nelken, "Unraveling the principles of auditory cortical processing: can we learn from the visual system?", *Nature Neuroscience* 12, 698 – 701, 2009. doi:10.1038/nn.2308
- [4] Laurel J. Trainor and Andrea Unrau, "Development of pitch and music perception", in *Springer Handbook of Auditory Research: Human Auditory Development*, Werner, Lynne; Fay, Richard R.; Popper, Arthur N. (Eds.), Vol. 42, Chapter 8, January 2012.
- [5] Stough, Bruce. "The lower tailpiece resonances" in *CASJ Catgut acoustical society journal* 3, N°1 series II, pp.17–25, 1996.
- [6] Curtin Joseph, "Scent of a violin, the signature modes of Old Italian violins & violas, part 1", in *The Strad*, June 2009.
- [7] Savart, Félix, "Mémoire sur la construction des instruments à cordes et à archet", Paris, 1819, p.135.
- [8] Hutchins, Carleen M, and Duane Voskuil. "Mode tuning for the violin maker." in *Catgut acoustical society journal* Vol. 2 N°4 serie II, 1993.
- [9] Woodhouse, Jim. 2002. "Body vibration of the violin- What can a maker expect to control ?", in *Catgut acoustical society journal* vol 4 N°5, 2002.
- [10] Bissinger, Bissinger George, "The violadeas project: Structural acoustics of good and bad violin" in *JCAS* vol 124, 2008, p.1764-1773.

VIBRATIONAL MODES OF THE VIOLIN FAMILY

Colin Gough

School of Physics and Astronomy,
University of Birmingham, UK
profough@googlemail.com

ABSTRACT

The generic wave-mechanical properties of violin-shaped instruments are described by considering their bodies as simplified, shallow, thin-walled, guitar-shaped, shell structures with the arched plates connected around their edges by the ribs. COMSOL finite element software is used to illustrate the strong dependence of the shapes and frequencies of the low frequency $A1$, CBR , $B1-$ and $B1+$ *signature* modes on the rib coupling strength, the island area between the f -holes, coupling to the internal cavity pressure fluctuations via the Helmholtz f -hole resonance, and the soundpost position and strength. The model illustrates the relationship between the free plate modes and those of the fully assembled instrument. It also identifies the important $B1-$ and $B1+$ signature modes as normal modes involving the in- and out-of-phase combinations of a bending and breathing mode of the shell, with the breathing component responsible for both the directly and indirectly (via the $A0$ mode) radiated sound. The model describes the vibrational modes over the whole playing range of the violin and can be used to predict both the admittance at the bridge and the radiated sound.

1. INTRODUCTION

Knowledge of the vibrational modes of the violin has advanced rapidly over the last few years from experimental modal analysis of instruments of widely varying quality, including many fine modern and classic Italian instruments by Stradivari and Guarneri instruments and their contemporaries, notably – notably from modal analysis measurements by Marshall [1], Schleske [2], Bissinger [3] and Stoppani [4]. In addition, several finite element computational investigations have successfully reproduced the shapes and frequencies of many observed vibrational states, including those of Knott [5], Roberts [6], Rogers and Anderson [7], and Bretos et al [8].

Nevertheless, the origin and nature of even the most important $A0$, CBR , $B1-$ and $B1+$ signature modes, responsible for almost all the sound radiated below around 800Hz to 1 kHz, has not been clearly understood. This is largely because of the asymmetric and often complicated shapes of the observed and computed modes

This investigation therefore adopts a somewhat different approach to that of previous finite element computa-

tions. It aims to explain and elucidate the origin and nature of the important acoustical modes of vibration, rather than simply to predict them for specifically chosen physical parameters related to a particular or typical violin.

To this end, we assume the simplest possible model having the necessary symmetry and constraints to reproduce the vibrational modes of violin-shaped instruments. This can be achieved by describing the hollow shell of such instruments as shallow, thin-walled, arched plate, guitar-shaped boxes, with f -holes cut into the top plates and offset soundpost and bass bar. Despite the added simplification of assuming uniform, isotropic, elastic constants and plate thicknesses, the model closely reproduces the frequencies and shapes of the free plate modes before assembly, reproduces the frequencies and shapes of the signature modes of the assembled instrument shapes, and correctly predicts modal densities of anisotropic plates at high frequencies as recognized by Cremer [9, §11.2]

As an aid to such physical understanding, the modes of the instrument are first investigated for a symmetric empty shell without soundpost or bass bar, then with a centrally placed sound post and finally with the symmetry-breaking offset soundpost and bass bar. The symmetry-breaking results in relatively complicated asymmetric modes involving the coupling together of previously symmetric and anti-symmetric modes about the central axis.

The model describes how the modes of the initially free plates are transformed into those of the assembled instrument, as the interactions between the initially free plates by the ribs, the internal cavity pressure fluctuations and the offset soundpost are slowly increased from near zero to typical values.

Although the focus of this paper is on the low frequency signature modes, the shapes and frequencies of the shell modes can be determined over the full playing range of an instrument. The absence of damping in the modal computations circumvents the practical problem of the strong overlap of modal resonances from damping at high frequencies leading to increasingly complex deflection shapes (*ODS*) involving combinations of the previous ideally symmetric and anti-symmetric flexural plate modes.

Nevertheless, damping is easily introduced into the computations. This enables the frequency dependence of the input admittance at the bridge and radiated sound to be derived as a continuous function of frequency over the whole playing range.

As described in this paper, the model enables one to investigate the dependence of important mode frequencies and shapes on the rib strength, plate masses, arching

profiles, internal cavity air resonances, and, most importantly, the sound post strength and location. The position of the sound post within the island area relative to the feet of the bridge is shown to have a major influence on both the resonant frequencies of the signature modes and the strength with which they can be excited by the bowed string.

In the following sections, the model is introduced and used to illustrate:

- the transformation of the front and back free plate modes into those of the assembled instrument,
- the identification of the *B1*- and *B1*+ signature modes as the coupled vibrations of the component bending and breathing modes of the shell,
- the influence of the *f*-holes on both the internal cavity air pressure vibrations and the coupling of such fluctuations to the shell vibrations which excite them,.
- the dependence of the signature and higher frequency modes on rib strength, the *f*-holes and island area, coupling to the air inside the cavity via the *f*-hole Helmholtz resonance and, most importantly, the soundpost position.

A longer paper will include and quantify the computed influence of elastic anisotropy, the linings and corner/end blocks, arching profiles and plate thicknesses, in addition to the coupled vibrations of the strings, neck and finger-board and higher-order cavity air modes, as well as a description of the generic properties of the higher frequency modes.

2. FINITE ELEMENT MODEL

The finite element geometry of the violin illustrated in Fig.1 is loosely based on the internal rib outline, arching profiles and other physical dimensions of the Titian Strad (Zygmuntowicz [10]). As we are interested in the vibrational properties of all instruments of the violin family, the exact dimensions and detailed geometry are only of secondary importance. Variations in physical and geometric properties can always be included later as relatively small perturbations, changing specific mode frequencies, but not the symmetry of their underlying shape. The model will be described in detail in a separate publication.

The unmeshed geometric model used for the finite element computations is illustrated in figure 1. The 15 mm high arching profiles of the plates were defined by simple mathematical functions, with identical top and back plate profiles across the width, but slightly different profiles along the length.

For simplicity, plates of uniform thickness have initially been considered, with uniform elastic properties representing the geometric mean of the anisotropic properties along and across the grains. This ensures the correct mode density at high frequencies (Cremer [9, §11.2]). The top and back plate thicknesses (2.5 and 3.5 mm), densities (460 and 660 kg/m³), masses (57 and 118 g) and averaged along/cross-grain elastic constants (2.39 and 2.17 GPa) were chosen to closely reproduce typical

arched front and back plate frequencies, as listed in table 1.

Frequency Hz	#1	#2	#3	#4	#5
Back FEA	93	169	248	252	348
Hutchins	116	167	222	230	349
Front FEA	82	158	218	231	333
Hutchins	80	146	241	251	295
Cremonese		134			314

Table 1. Arched FEA back and front free plate frequencies compared with Hutchins’ “tuned” plates (cited from Roberts [6]) and average values for nine Cremonese violins (Curtin [11]) both with *f*-holes in the front plate

The 2.1:1 ratio of the computed top plate #5 and #2 modes is slightly larger than the octave tuning advocated by Hutchins [12], but lower than the average 2.3:1 value for fine Cremonese violins measured by Curtin [11]. However, the values are well within the experimental scatter amongst fine instrument, which justifies the use of the use of the assumed isotropic plate parameters for our subsequent development of the modes of the assembled instrument. The influence of elastic anisotropy and graduations of plate thicknesses on mode shapes and frequencies will be described in a later paper.

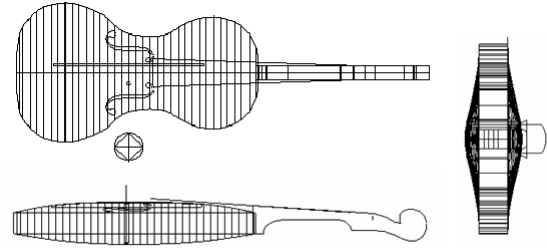


Figure 1. The unmeshed geometric model before meshing used for the finite element computations illustrating the guitar-shaped outline and arching of the plates and schematic representation of the neck. The transverse and longitudinal lines across define the cross-sections along which the arching profiles were defined. The circular disc is used to demonstrate the induced *f*-hole Helmholtz vibrations.

Although the computations presented in this paper are for the violin, the mode shapes and dependencies of modal frequencies on physical and geometric factors are expected to be much the same for instruments of any size, as the symmetry of the mode is largely determined by the symmetry of the shell structure, which is much the same for all instruments of the violin family.

The computations were made using the structural shell module of COMSOL 3.5 Multidisciplinary software. An automated mesh with typically 50,000 degrees of freedom was generated, with the first 20 to 100 or so vibrational modes of the freely supported instrument computed in a few tens of seconds on a desk-top PC.

The influence on the vibrational modes of the individual plates, the ribs, *f*-holes, central and offset soundpost, and internal air cavity are now considered in turn. Space

precludes a description of the influence of bass-bar and neck/fingerboard assembly, which perturb but do not significantly influence the generic properties of the vibrational modes of the instrument of main interest her.

3. FREE PLATE MODES

Figure 2 illustrates the computed free plate modes used to model the assembled instrument, before f -holes have been cut or bass bar added. Throughout this paper a colour scale will be used to illustrate displacements perpendicular to the plates, with dark red and blue representing equal but opposite displacements perpendicular to the plate, with the nodes at the transition between green and yellow.

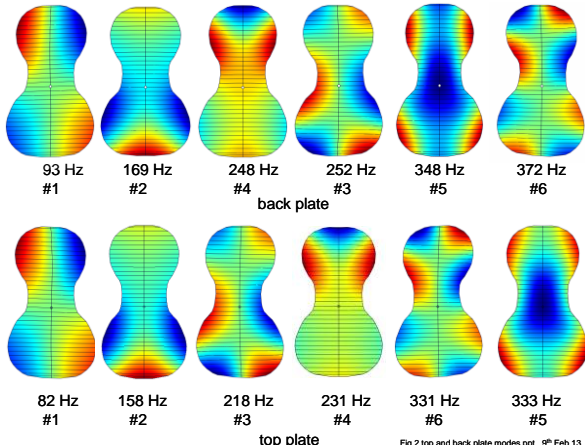


Figure 2. The first six modes of the isotropic, uniform thickness, arched, top and back plates

Arching results in a strong coupling between the flexural waves perpendicular to and longitudinal waves parallel to the arched surfaces. This can double the frequency of the low frequency plate modes, though it is less important as the mode frequencies increase. Such coupling also induces significant in-plane edge contractions

and extensions, responsible for the coupling between the bending and breathing modes of the assembled instrument. The plate frequencies are therefore strongly dependent on how they are supported by the ribs both in plane and perpendicular to their edges. Figure 2 demonstrates that even relatively small changes in arching profile along the lengths of the top and back plate can reverse the order of the mode frequencies, even though their arching heights are the same.

The individual modes are either symmetric or anti-symmetric about the longitudinal central axis. This remains true for their coupled motions in the assembled instrument in the absence of the symmetry-breaking bass bar and offset soundpost.

Mode #1 is a torsional mode of relatively little acoustic importance. Modes #2 and #4 involve flexural bending vibrations in the lower and upper bouts of the freely supported plates. Mode #4 is higher in frequency because it is confined to a smaller area. Mode #2 is often referred to as the *X*-mode, on account of the shape of its nodal lines, though they never cross. As we will show later, the coupled #2 and #4 modes result in the largely non-radiating bending mode component of the *B1*- and *B1*+ signature modes. Mode #3 is a higher frequency torsional mode, which includes some bending. Both plates vibrating in the same direction result in the *CBR* mode of the assembled instrument. Mode #5 is often referred to as the ring-mode on account of the nodal line around the central region of the plate moving in the same direction. When coupled by the ribs, the resulting mode with the two plates vibrating in opposite directions is responsible for the breathing mode components of the *B1*- and *B1*+ signature modes. These modes are directly or indirectly (by excitation of the *A0* f -hole resonance) responsible for virtually all the sound radiated by the violin and other stringed instruments at frequencies over their first two octaves.

4. RIB COUPLING

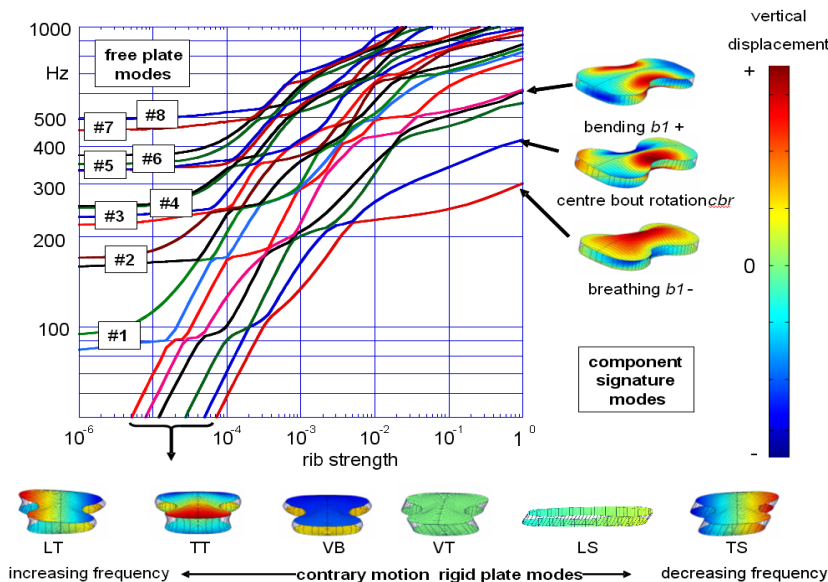


Figure 3. The computed transformation of the first eight modes of the free plates to those of the fully assembled but empty, doubly-arched, guitar-shaped, shell structure without f -holes, as a function of rib coupling strength increased from a very small value to that of a typical violin (see text). The mode frequencies are also strongly perturbed by the illustrated contrary rotational and linear vibrations of the rigid plates about their three orthogonal axis, with unperturbed vibrational frequencies proportional to (rib strength)^{1/2}.

Figure 3 illustrates the rather complicated way in which the modes of the individual free plates are transformed into those of the assembled instrument on increasing the rib coupling strength from close to zero to a representative normal value.

The coupling strength is proportional to $E_{rib}(t/h)^3$, where E_{rib} is an effective elastic constant across the ribs of height h and thickness t . In computing the dependence on rib strength the rib height (3 cm) and thickness (1mm) were held constant, while E_{rib} was scaled from a very small value to 10 GPa - a typical value for maple. For the cello, with its significantly larger rib height to thickness ratio, the rib coupling strength will be significantly weaker than that of the violin. This will result in relatively larger stretching and bending of the ribs.

The density was also simultaneously scaled by the same factor to maintain the frequency of the wave-guide like, flexural, rib modes between the two plates at a high frequency, typically >5-10 kHz for the violin and > 800 Hz for the cello (Stoppini, private communication) - well above the range of frequencies considered here. At lower frequencies, the ribs simply act as a series of parallel cantilevered springs inhibiting plate separation and bending around their edges. Nevertheless the isolated rib garland can easily be bent and twisted about its length with very little energy. Such vibrations are involved in the *CBR* and *B1-* and *B1+* signature modes of the violin, with large amplitude twisting and bending of the ribs along their edges, but relatively small amounts of stretching and bending between opposing plate edges.

Figure 3 shows that the ribs have a major influence on the frequencies and shapes of the low-lying modes of the assembled shell, which become the signature modes of the fully modeled instrument. The coupling is especially strong between front and back plates sharing the same symmetry and closely spaced frequencies.

As well as the coupled free plate modes, there are six additional modes derived from the twelve zero-frequency degrees of freedom of the two isolated plates describing their rigid body displacements and rotations along and about their three orthogonal symmetry axis.

Six of these modes become the six whole-body displacements and rotations of the assembled instrument. The remaining six modes are transformed into modes with the displacements and rotations of the rigid plates in opposite directions, as illustrated in Figure 3. Because such modes involve the stretching and compression of the ribs, their frequencies increase with rib strength as $\sqrt{E_{ribs}/M_{plates}}$, where M_{plates} will be a mode-specific effective plate mass. Their unperturbed frequencies therefore increase with slope $1/2$, when frequency and rib strength are plotted on logarithmic scales, as in figure 3.

At intermediate coupling strengths, these modes cross and will couple to any of the flexural wave modes of the top and back plates sharing a similar symmetry. This results in considerable veering and splitting of several mode frequencies in the cross-over region. The splitting of modes is proportional to the rib-induced interaction strength. Modes not sharing a common symmetry do not interact. Their frequencies simply cross.

At full coupling strength, the influence of the rigid plate modes are less important but still account for the small amounts of rib stretching, rotations and twisting observed in experimental modal analysis measurements.

Because of these interactions, the dependence of shell mode frequencies on rib strength is rather complicated and difficult to interpret. In figure 5a-d below, we have therefore extracted those parts of the dispersion curves that identify the transformations of specific plate modes to the low frequency modes of the assembled instrument.

Before leaving figure 3, it is important to note that, despite the complexity of the modal frequency plot, the number of modes of the assembled structure is always conserved and is equal number of the number of initially non-interacting modes considered. It is therefore possible to follow a single mode of the interacting system from that of the original uncoupled modes as they adiabatically (smoothly) transformed to those of the fully coupled structure. However, as the coupling strength increases each mode will increasingly include additional coupled component vibrations of the initial system.

As a result of the increase in modal frequencies on increasing rib strength, the number of low frequency modes of the shell at full coupling strength is relatively small. The five lowest modes are illustrated in Figure 4. These will be referred to as the *cbr* (centre bout rotation), *b1-* (breathing), *b1+* (bending), *ld* (longitudinal dipole) and *td* (transverse dipole) modes of the empty shell. Note the use of small letters to denote what become contributing component modes of the *CBR*, *B1-* and *B1+* normal modes of the assembled instrument.

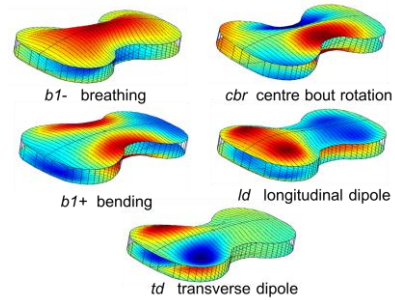


Figure 4. The first five component or basis modes of the assembled guitar-shaped box at full rib-coupling strength.

The transformation of each of the first five free plate modes into the above modes of the freely supported assembled shell will now be described.

Figure 5a illustrates the influence of rib coupling on the #1 torsional plate mode. Even at vanishingly small coupling, the top and back plates are coupled together to form two new *normal* modes with the plates twisting in either the same or opposite directions. The mode with both plates twisting in the same direction avoids stretching the ribs. Its stored energy, hence frequency, therefore increases far less rapidly with rib coupling strength than the contra-twisting mode, which rapidly rises in frequency into the multiplicity of flexural plate modes above a kHz. The mode with plates twisting in the same direction also transforms into a relatively unimportant acoustic mode just below a kHz.

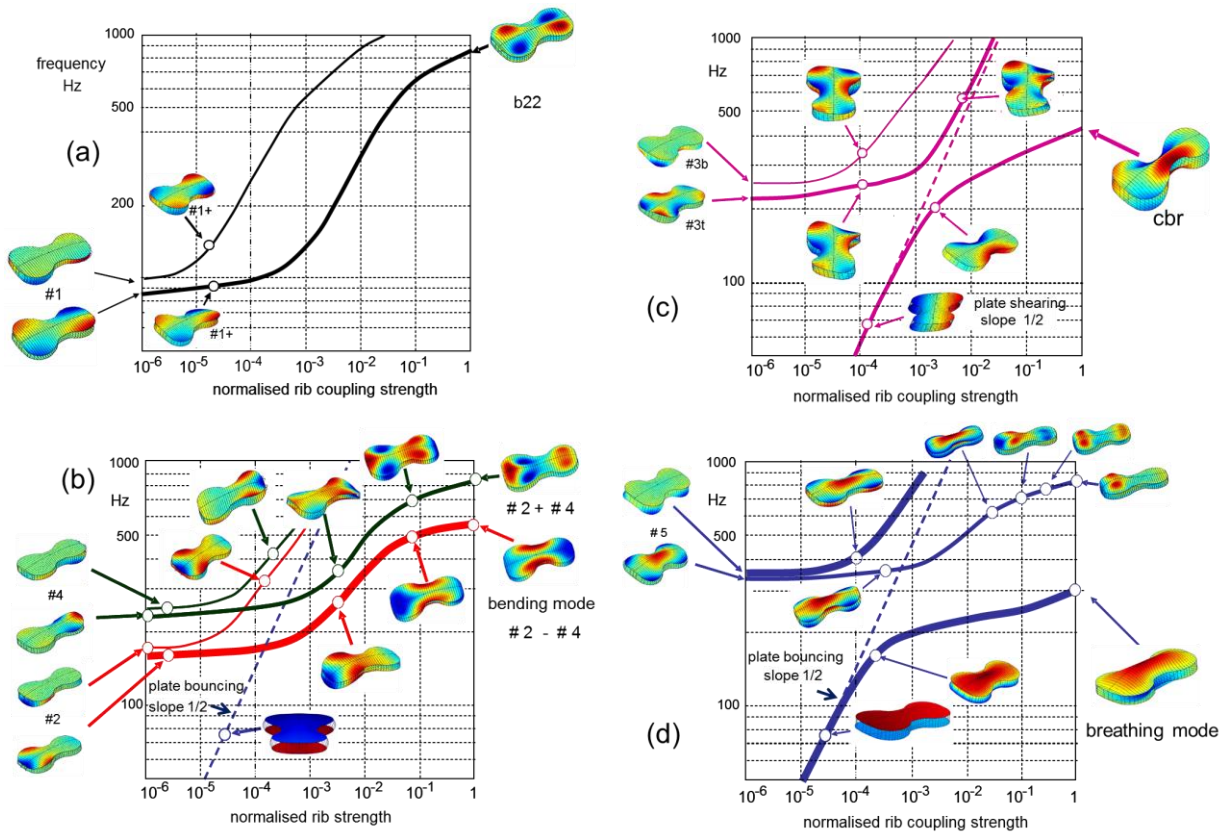
Figure 5b illustrates a similar initial behaviour for the coupled #2 and #4 free plate modes, with the frequency of the modes with plates flexing in opposite directions increasing much more quickly than when flexing in the same direction, avoiding stretching the ribs. In contrast, the pair of modes with plates vibrating in the same direction are coupled together by the ribs to form the important component, antilastic (bending in opposite directions), *b1+* bending mode of the *B1-* and *B1+* signature modes of the fully instrument.

The #3 modes illustrated in figure 5c behave in much the same way, with plates vibrating in opposite senses increasing rapidly in frequency. However, in this case, the lower frequency mode, with the two plates vibrating in the same direction, crosses and interacts strongly with the more rapidly increasing frequency shearing mode of the rigid plates. After crossing, the emerging *cbr* mode still describes the centre bout rotation, but retains a significant amount of coupled shearing motion. This results in a central bout rhombohedral distortional vibration, explaining the origin of the *CBR* name. However, it is its rotation rather than its shear motion that is important in coupling to the rocking bridge. It would therefore arguably be more appropriate to refer to this mode as a centre bout rotation mode. Because such motion involves little change in volume, the *cbr* mode, which transforms into the *CBR* mode of the fully assembled instrument,

usually only plays a minor role in the radiation of sound.

Figure 5d illustrates the very strong interaction of the #5 plate modes with the rising frequency bouncing mode. Inspection of the resulting mode shapes identifies the lowest and highest frequency branches (the thicker solid line) as the in- and out-of-phase of the coupled #5 “ring-modes” with the rigid plate bouncing mode (involving the stretching and compression of the ribs). The lower mode becomes the all-important component *b1*-breathing mode responsible for almost all the sound radiated directly or indirectly (via its excitation of the *f*-hole Helmholtz resonance) by instruments of the violin family over their lowest two octaves. On crossing the rising bouncing mode frequency, the breathing mode re-emerges at a significantly lower frequency than its initial value, as expected for any pair of strongly coupled oscillators. In contrast, the mode with the initial #5 ring plate vibrations in the same phase simply crosses the bouncing mode frequency and becomes another relatively unimportant acoustic mode in the transition region just below 1 kHz.

In every case, the mode frequencies are still rising for a coupling factor of unity corresponding to typical values. This reflects the increasing rigidity of the supporting ribs, which increasingly constrain the bending and stretching of the flexural waves around the plate edges.



Figures 5a-d. Extracted dispersion curves for the transformations of the first five free plate modes to the lowest frequency modes of the assembled shell.

6. *f*-HOLES AND ISLAND AREA

The open *f*-holes on the front plate and the island area between them play a major role in the sound of the violin and related instruments, as recognized by Cremer [1, chpt.10]. Firstly, the open holes result in the *A0* Helmholtz cavity resonance, which boosts the sound of all members of the violin family over their first octave or so. Secondly, the free *f*-hole edges define the shape of the island area, which strongly influences the penetration of flexural waves from the lower and outer bouts towards the two feet of the rocking bridge, which excite them. As shown later, the penetration of flexural waves into the island area and resulting excitation of radiating modes is also strongly influenced by the strength and position of the soundpost in the island area.

6.1 The island area

Figure 6 illustrates the influence of the *f*-holes on the low frequency flexural wave modes of the shell. The frequencies were computed as a function of *f*-hole strength varied by simultaneously decreasing the elastic constant and density of the *f*-hole areas by the same factor from unity (no *f*-holes) to 10^{-5} (effectively open *f*-holes).

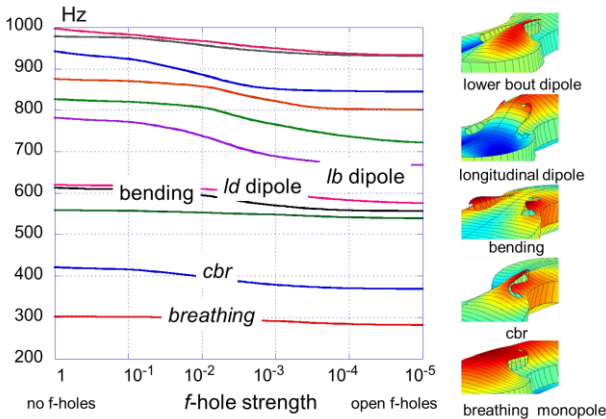


Figure 6. The influence of the *f*-holes on the empty shell mode frequencies and shapes in the island area.

Opening the *f*-holes has little effect on mode shapes, but lowers their frequencies by increasing their penetration into the island area towards the free edges of the island area. As a result the inside edges of the *f*-holes can vibrate with large amplitudes, almost independently from the outer edges constrained by the ribs, as illustrated.

6.2 The Helmholtz *A0* resonance.

The combined open area *A* of the two *f*-holes together with the volume *V* of the internal cavity, if rigid, would form a Helmholtz resonator, with resonant frequency

$$2\pi f_{\text{Helmholtz}} = c_o \sqrt{gA^{1/2}/V}, \quad (1)$$

where c_o is the velocity of sound within the cavity and *g* is an *f*-hole, shape-dependent, constant of order unity. The resonance is excited by the volume-changing, *b1*-component of the cavity wall flexural wave vibrations.

Although the Helmholtz frequency, proportional to the velocity of sound in air, is independent of pressure, the strength of the coupling between the induced cavity pressure fluctuations and flexural plate vibrations that excite them is proportional to the ambient pressure. Figure 7 illustrates the dependence of the shell modes and *A0* *f*-hole frequencies on ambient pressure.

The computations assume a uniform acoustic pressure within the cavity induced by the cavity volume changes. These are computed self-consistently and describe the interaction of the shell wall vibrations with the air inside the cavity via their coupling to the *f*-hole Helmholtz *A0* resonance. For uniform pressure changes, only volume only volume changing shell wall vibrations can provide such coupling.

Figure 7 illustrates the very strong perturbation of the *A0* and *b1*- breathing modes, as the ambient pressure is increased. Neglecting, interaction with the enclosed air is equivalent to zero ambient pressure. The previously computed empty shell breathing mode frequency of around 300 Hz is then very close to that of the modeled ideal Helmholtz resonator at 307 Hz. On increasing the internal pressure to a “normal” value, the coupling is so great that the predicted *A0* mode would drop to around 200 Hz, while the breathing mode would be raised to around 400 Hz. Analytic models, based on inertial rather than spring-like coupling, provided via the acceleration of the cavity walls and the pressure within the cavity, predict that the product of the resulting *a0* and *b1*- mode frequencies should remain constant, as confirmed by the computations.

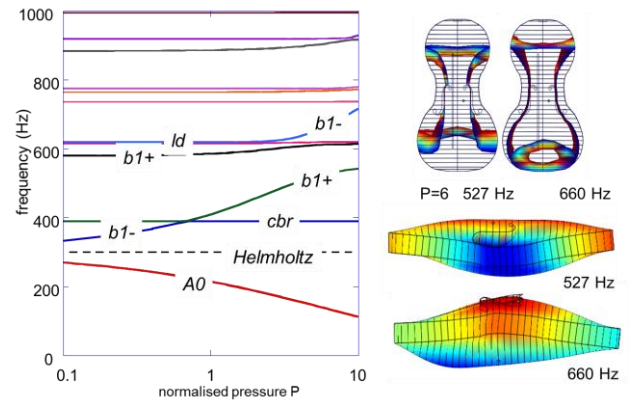


Figure 7. The variation of the low frequency modes of the empty guitar-shaped shell with *f*-holes cut into the top plate, as a function of ambient pressure normalized to normal air pressure. The dashed line indicates the unperturbed Helmholtz frequency of 307 Hz. Also illustrated are the mode shapes and reversed *baseball-like* nodal lines of the coupled *b1*-breathing and *b1*+ bending mode at high pressures.

As described in an accompanying paper (Gough [13]), the internal pressure driving the Helmholtz vibrations of the air in and out of the *f*-holes is reduced below the average pressure - because of the pressure drops associated with flow along the length of the cavity towards the *f*-holes. This reduces the effective pressure driving the *f*-hole resonance by a factor of ~ 0.7 , equivalent to reducing the effective ambient pressure by the same factor.

Measurements on violins with and without a soundpost show typical drops in $A0$ frequency from around 285 to 240 Hz, in qualitative agreement with the above model.

Initially, none of the other modes are significantly affected. However, as the pressure increases, the frequency of the $b1$ - breathing mode approaches and would otherwise cross that of the $b1$ + bending mode.

Because the bending and breathing modes similar but different amplitude in-plane contractions and extensions around the edges of the plates, they are relatively strongly coupled. They then form what become the dominant $B1$ - and $B1$ + signature modes of the assembled instrument with bending and breathing component vibrations in either the same or opposite phases, as illustrated in figure 7. This is one of the major results of the present investigation and has been confirmed by modal analysis measurements by Stoppani [invited talk at this conference] on many fine instruments.

The coupling of the bending and breathing modes automatically accounts for the characteristic reversal of the baseball-shaped nodal lines. Previous attempts to describe the $B1$ - and $B1$ + modes have suggested they involve the reversal of #2 and #5 free plate modes in the top and back plates, despite such modes having very different plate edge displacements. Furthermore, the coupling explains why the $B1$ - and $B1$ + modes are never coincident. It also explains the variations in relative monopole radiation strengths of such modes, which will depend on the contributing strengths of the component $b1$ - breathing (strongly radiating) and $b1$ + bending (weakly radiating) modes.

There is also a somewhat smaller interaction with the higher frequency longitudinal dipole mode illustrated in figure 4 and with a higher frequency mode just above 900 Hz.

6. THE SOUNDPOST

The soundpost wedged between the plates acts as a supporting beam exerting equal and opposite forces and couples to the plates across its ends. This imposes almost equal flexural wave plate displacements and slopes across its ends assuming intimate contact between the ends of the sound post and the plates.

As a result, the flexural wave displacements at both ends of the sound post tend to be extremely small. This involves a modification of the flexural mode of the empty shell by the addition of a localised flexural wave decaying exponentially as e^{-kr} at large distances, where $k = 2\pi/\lambda$ is given by the usual flexural wave dispersion relationship, $\omega \propto tk^2$ and λ is the characteristic wavelength of the standing waves excited in the upper and lower bouts. The existence of such waves is a characteristic feature of bending waves on thin-walled structures and have to be added to the more familiar wave-like functions at any boundary to satisfy the various boundary conditions involved –around the plate edges, along the f -hole slots and, in this case, at the ends of the soundpost.

When located within the island area, the soundpost then acts as a kind of gate inhibiting the penetration of flexural waves in the lower bout past the sound post into the upper bout and vice versa. Because the flexural wave

amplitudes are rather small and vary rapidly close to the soundpost ends, its position relative to the two feet of the bridge which excite the shell vibrations is crucial. The strength with which any vibrational mode of the shell can be excited by a horizontal bowing force at the bridge is critically dependent on the wave amplitudes at the two feet of the bridge, which excite such vibrations. The crucial role of the soundpost on the sound of an instrument is why the English call it the soundpost, while the French and Italians more imaginatively refer to it as the soul and spirit of the instrument.

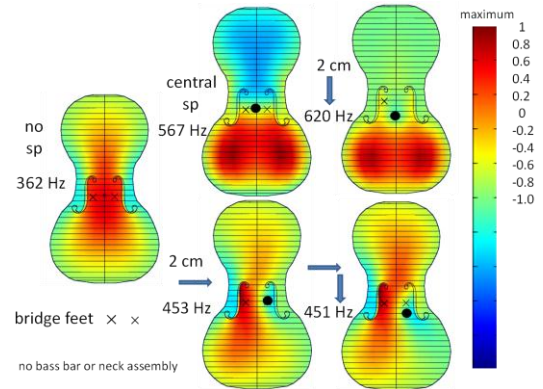


Figure 8. An overview of the influence of soundpost position on the frequencies and top plate mode shapes of the $b1$ - breathing mode, for the empty shell without bass bar or other attachments, but coupled to the internal air via the f -hole resonance. Modes are illustrated for the empty shell, a centrally placed soundpost, when it is shifted longitudinally and transversely separately by 2 cm, and when shifted in both directions together.

The gate-like function of the soundpost is evident from Figure 8. Without a soundpost, the breathing mode extends through the island area to both upper and lower bouts, with the strongest amplitudes at the upper end of the lower bout and within the island area. Because there is little restriction on the area available for the waves, the frequency is relatively low. In contrast, the nodal regions created by the centrally placed sound post strongly inhibits the penetration of waves from the lower to the upper bout, though there is still some penetration on both sides of the soundpost resulting in weak displacement of opposite sign in the upper bouts. Because of the restriction on available area for the lower bout vibrations, the frequency of the mode is increased very significantly from 362 to 567 Hz. Displacing the soundpost 2 cm towards the lower bouts further inhibits penetration past the sound post and the area available for vibrations is decreased, resulting in a further increase in mode frequency from 567 to 620 Hz. Because all the above modes are symmetric about the central axis, they cannot be excited by a horizontal force causing a symmetric rocking of the bridge. The violin would radiate very little sound.

Moving the soundpost 2 cm sideways from its central position has two affects. It allows the lower bout waves to penetrate more easily through the bass side of the island area, significantly lowering the mode frequency from 567 to 453 Hz. The sideways displacement also introduces a large asymmetry of the wave motion across the two feet

of the bridge. A horizontal bowing force can then couple strongly to the modes via the asymmetrically rocking bridge and hence radiate strongly. Because the waves already penetrate relatively easily past the offset sound post, moving it further away from the bridge results in a relatively small change in mode frequency from 453 to 451 Hz, though it results in quite a large change in the wave motion within the island area and upper bouts, particularly in the region of the two bridge feet.

Although such displacements are very much larger than the mm or so used by violin makers when setting up a violin, the computations illustrate how such changes are likely to change the sound of an instrument, both by changing the frequency of the important breathing mode component and the strength with which it can be excited by horizontal bowing forces at the bridge.

In view of the relatively large variations in $B1-$ and $B1+$ signature mode frequencies and their component breathing mode responsible for the radiated sound, even amongst many fine Cremonese instruments, it is somewhat surprising that many luthier's believe there is a "correct" position for the soundpost within a mm or so of its placing relative to the treble side foot of the bridge. This may well be true for fine instruments with properties already optimized for the "correct" soundpost position. But one has to ask why use is not made of the sensitivity of sound quality on soundpost position to "improve" the sound of poorer quality violins by much more adventurous shifts than the mm or so usually considered.

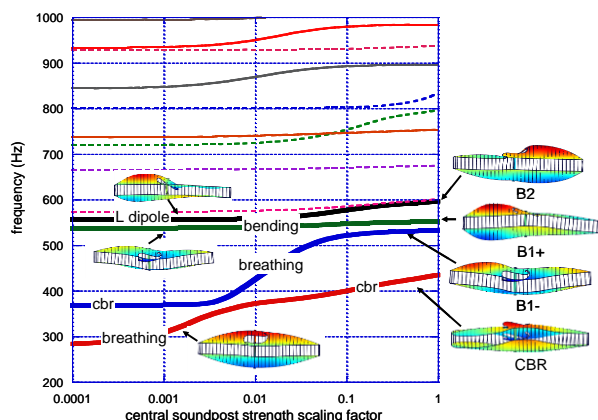


Figure 7. Influence on mode frequencies and shapes of the strength of a centrally placed soundpost.

Figure 7 illustrates the dependence of mode frequencies on the strength of a centrally placed soundpost in-line with the f -hole notches, for the violin body shell without bass bar, fingerboard/neck assembly and neglecting coupling to the Helmholtz cavity resonance. As with the ribs, the soundpost strength is varied by simultaneously increasing its elastic constant and density to maintain its resonant modes at a high value - above ~ 5 kHz for the violin, to avoid complications from the multiplicity of soundpost modes for small elastic constants. A centrally placed soundpost only significantly perturbs symmetric shell modes.

As already illustrated above, the most important feature is the strong increase in frequency of the $b1-$ breathing mode. This first crosses the frequency of the cbr

mode with a small amount of veering and splitting of mode frequencies indicating an inherent coupling between the bending and cbr mode. The frequency then continues to rise, approaching and otherwise crossing the initially higher frequency $b1+$ bending mode. In the cross-over region one again observes the formation of the $B1-$ and $B1+$ modes observed on increasing the internal cavity pressure.

Figure 8 illustrates the strong dependence of the low frequency mode frequencies and interactions on shifting the soundpost along the central axis. By shifting the sound post over relatively large distance both below and above the central bridge position, one can identify separate non-interacting lower and upper bout breathing modes, which cross without any significant veering or splitting, both of which are coupled to the bending mode, which is scarcely affected by soundpost position. As the centrally placed soundpost is moved along the length, the lower and upper bout breathing mode frequencies change in opposite directions, as the areas in which they are constrained to vibrate increase and decrease. At a given position - about a cm behind the bridge, their frequencies would coincide. However, close to the bridge the lower bout breathing mode is lower in frequency than the upper bout mode,

For the violin, only the $B1-$ and $B1+$ and sometimes the CBR modes radiate strongly in the monopole signature mode regime, though $B2$ and other higher frequency modes can introduce additional weak resonant features.

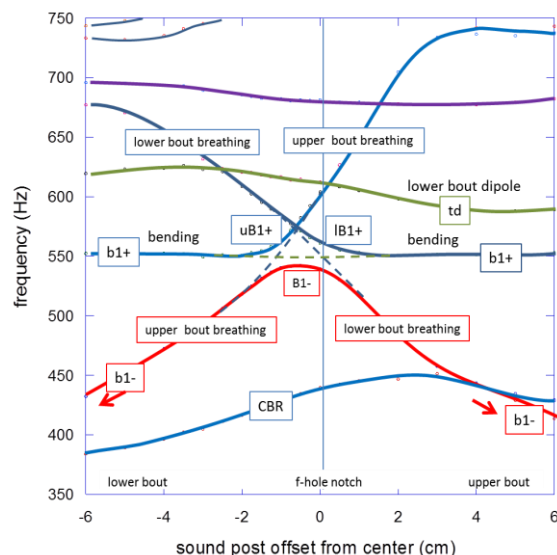


Figure 8. Influence on mode frequencies of shifting the soundpost along the central axis.

The influence on the frequencies of the low-lying modes, on displacing the initially centrally placed sound post away from the central axis is illustrated in figure 9. As described above, this results in a significant decrease in the breathing mode frequency, but only has a small influence on the bending mode. However, as the $b1-$ breathing mode frequency crosses the cbr mode there is a significant veering of the modes describing their coupled vibrations, implying quite a strong coupling between them.

At present, the physical origin of the coupling involved between the breathing and *cbr* component modes is not understood, but is likely to be a factor in determining why the *CBR* mode of the assembled instrument makes a significant contribution to the radiated sound in some instruments, but not in others.

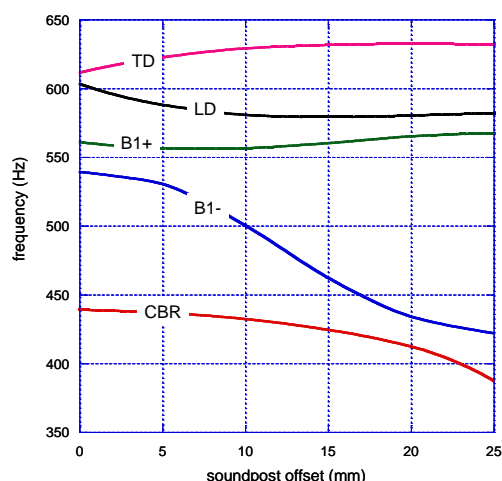


Figure 9. The influence of offsetting a centrally placed sound post from the central axis.

The added influence of anisotropy, bass bar, finger-board-neck assembly and coupled air modes will be described in a later paper, but do not significantly change the predictions symmetries and coupling interactions in any significant way.

8.SUMMARY

A wave mechanical model to describe and explain the vibrational modes of all members of the violin family has been introduced. The model is based on finite element computer simulations of the generic properties of the modes of vibration of thin-walled, doubly-arched, guitar shaped, shallow, box-like shell structures. The connection of such modes to the free plate modes had been demonstrated, in addition to their dependence on rib-coupling, the *f*-holes, Helmholtz cavity resonance, and sound post strength and position.

9. REFERENCES

- [1] K.D. Marshall, Modal analysis of a violin, *J. Acoust. Soc. Am.*, vol.77, no.2, 1985, pp. 695-709.
- [2] M. Schleske, "On making Tonal Copies of a Violin" CAS Journal Vol.3, No.2, (Series II), November 1996, pp. 18-28.
- [3] G. Bissinger, Structural acoustics of good and bad violins, *J. Acoust. Soc. Am.*, Vol. 124, No. 3, 2008, pp. 1764-1773.
- [4] G. Stoppani, invited paper at this conference.
- [5] G.A. Knott, *A modal analysis of the violin*, MSc Thesis, Naval Postgrad. School, Monterey, pp 507-550, in Research Papers in Violin Acoustic 1975-1993 (ed. C.M Hutchins and V. Benade), Acoust. Soc. Am., 1997.
- [6] G.W. Roberts, *Finite element analysis of the violin* (PhD Thesis, Cardiff,), pp. 575-590 in Research Papers in Violin Acoustic 1975-1993 (ed. C.M Hutchins and V. Benade), Acoust. Soc. Am., 1997.
- [7] O. Rogers and P. Anderson, Finite element analysis of a violin corpus, CASJ, Vol.4., no.4. (Series II), November 2001, pp. 43-49.
- [8] J. Bretos, C. Santamaria and J.A. Moral, Vibration patterns and frequency responses of the free plates and box of a violin obtained by finite element analysis, *Jnl. Acoust. Soc. Am.*, vol. 105, no.3, 1999, pp.1942-1950.
- [9] L. Cremer, *The Physics of the Violin*, MIT, 1984
- [10] S. Zygmuntowicz, The Titian Strad, *The Strad*, February, 2009, pp 30-34.
- [11] J. Curtin, Tap tones and weights of old Italian violin tops, *Jnl. Violin. Soc. Am.*, Vol 20, no.2, 2006, pp. 161-174.
- [12] C.M. Hutchins, The acoustics of violin plates, *Sci. Am.*, October 1981, pp170-186
- [13] C. Gough, Acoustic characterization of violin family signature modes by internal cavity measurements (contributed paper at this conference).

ACOUSTIC CHARACTERIZATION OF VIOLIN FAMILY SIGNATURE MODES BY INTERNAL CAVITY MEASUREMENTS

Colin Gough

School of Physics and Astronomy, University of Birmingham, UK
profgough@googlemail.com

ABSTRACT

The sound radiated by the signature modes of stringed instruments in their lowest two octaves is shown to be directly related to the sound pressure at the “acoustic centre” of their hollow bodies, where the nodal lines of the longitudinal $a1$ and transverse $a2$ internal dipole modes of the air cavity intersect. Pressure measurements inside the instrument itself can therefore be used to characterise the acoustic radiating properties of the signature modes of instruments of any size without contamination from resonances of the surrounding space or interference from external noise. This is illustrated by high resolution, acoustic, location-independent, measurements on violins, a viola and several double basses, which could all have been made in the luthier’s workshop without any sophisticated equipment.

1. INTRODUCTION

The sound of all instruments of the violin family in their lowest two octaves is strongly influenced by the low frequency $A0$, CBR , $B1$ - and $B1$ + signature modes, responsible for prominent resonances in both admittance and acoustic measurements. These modes are closely related to the f -hole Helmholtz resonance, centre bout rotation and a mixture of component bending and strongly volume-changing breathing mode vibrations (Gough [1]). The breathing mode component of the signature modes is responsible directly and indirectly (via excitation of the air bouncing in and out of the f -holes) for almost all the sound radiated by stringed instruments in their lowest two octaves, which is strongly correlated with the perceived quality of an instrument (Bissinger [2]).

Over the last few years, a large number of measurements of the radiated sound have been made on many violins including many Stradivari, Guarneri and outstanding modern instruments. A common method has been to measure the acoustic impulse response, with the radiated sound excited by a sharp tap at the bridge measured at typical distances between around 25 cm to 40 cm from the violin (Curtin [3]). This is already a compromise, as close to the violin the sound includes a strong, near-field component (instrument length \sim 34-35 cm).

Copyright: © 2013 First author et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

The near-field sound is important to the player, with their ears close to the instrument, but decays much faster than the radiated sound contributing very little sound to the distant listener. Considerably larger measurements distance would be needed for comparable quality measurements for larger members of the violin family like the cello and double bass. However, as the radiated sound intensity falls as $1/r^2$, contamination of the measurements from acoustical resonances of the surrounding space increases. As a result, very few measurements of the acoustic properties of comparable quality exist for larger instruments like the cello and double bass.

An alternative method of identifying the low frequency acoustic characteristics of even the largest of instruments is to measure the internal sound pressure, as this is directly related to the radiated sound in the monopole limit. This was first described and demonstrated by Jansson, Morsett and Guttler [4]. Somewhat surprisingly, this method has not been widely adopted as a characterization tool by makers or researchers, despite the simplicity of making such measurements.

The volume into which the internally radiated sound is confined is relatively small and extremely well defined. This results in easily reproducible high sound pressures, which are virtually free of any contamination from resonances of the surrounding acoustic space or extraneous noise. Such measurements can then be made on instruments of any size - even in a small luthier's workshop. Effectively, the hollow cavity of the instrument itself becomes the recording studio.

The relationship between the internal and radiated sound pressures is described by standard loudspeaker-enclosure theory originally used by Caldersmith [5] to describe the acoustical properties of the guitar and later by Schelleng [6] for the violin. A generalized model will be described, which accounts for both the sound radiated and the internal pressure excited by all four $A0$, CBR , $B1$ - and $B1$ + signature modes.

Following a brief description of the underlying theoretical model, examples of the radiated and internal sound pressures of the violin are described. Measurements on the violin, a viola and a number of double basses are then presented

The examples highlight the potential value of such simple measurements, especially for the larger stringed instruments, where the normal techniques used to measure their radiating properties at low frequencies are impracticable because of their large size.

2. MODEL

The standard loudspeaker enclosure model assumes a simple vibrating piston source of area S exciting pressure fluctuations $p = \gamma P_0 \Delta V / V$ in an enclosure of volume V at ambient pressure P_0 , where ΔV is the inward volume change and γ is the ratio of adiabatic and isothermal heat capacities for air. The pressure induces the Helmholtz resonance of a plug of air bouncing in and out of the cavity through one or more holes of total area A cut into the walls.

For the multi-resonant instruments of the violin family, the model can be generalised by the equivalent circuit shown in Figure 1. The currents represent the rate of volume flow u_{shell} induced by the flexural vibrations of the shell walls of thickness t , while $u_{\text{Helmholtz}}$ represents the induced flow of air in and out of the f -holes. The kinetic inductance of the m -th mode L_m is m_m / S_m^2 , where the effective mass at the driving point r_o is given by $m_m = M \psi_m(r_o)^2$, where the hollow shell mass $M = \int_S t \rho \psi_m(r)^2 dS$ is given by the mass-normalised flexural mode shape $\psi_m(r)$, with the integral is taken over the surface of the shell. The effective piston area $S_m = \int_S \hat{\psi}_m(r) \cdot \hat{n} dS$ describes the volume change associated with the inward flexural wave shell vibrations, where \hat{n} is a unit vector perpendicular to the surface towards the inside of the hollow cavity of the instrument. The neck, fingerboard and all other attached components introduce weak perturbations, mostly affecting signature mode frequencies rather than mode shapes.

The cavity capacitance $C_{\text{cavity}} = V_{\text{res}} / \gamma P_0$, where the voltage across it represents the induced pressure p . The mode plate capacities $C_m = S_m^2 / L_m \omega_m^2$ and f -hole inductance $L_{f\text{-hole}} = 1 / C_{\text{cavity}} \omega_{\text{Helmholtz}}^2$, where ω_m and $\omega_{\text{Helmholtz}}$ are the uncoupled resonant angular frequencies of the plate and Helmholtz resonator (i.e. measured at zero ambient pressure and rigid cavity walls respectively). Damping is included by introducing mode specific Q -factors, such that $\omega_m^2 \rightarrow \omega_m^2 (1 + i/Q_m)$ for the individual resonant mode frequencies involved.

The above standard model assumes a uniform acoustic pressure p within the cavity. In practice, the Helmholtz mode acoustic pressure varies along the central length of the cavity as $p \phi_H(r)$, with the highest pressure at the upper bout end, the lowest opposite the f -holes and an intermediate value at the lower bout end, as described later. The coupling between the cavity wall modes and Helmholtz resonator then depends on the mode-dependent overlap integral $\int_S \hat{\psi}_m(r) \cdot \hat{n} \phi_H(r) dS$.

As described in an accompanying paper [1], the CBR , $BI-$, $BI+$ normal modes involve combinations of component volume-changing breathing, centre bout rotation and bending vibrational modes. Computations suggest that only the volume-changing breathing mode component both radiates and couples strongly to the Helmholtz pressure fluctuations. This implies that each signature mode will contribute in equal measure to both the radiated and internal sound pressure, by an amount determined by the contribution of the component breathing mode vibrations in each mode. Nevertheless, the acoustic pressure driving

the plug of air through the f -holes will be reduced by the factor $S \phi_H(r_c)^2 / \int_S \phi_H(r)^2 dS$, where to a good approximation $\phi_H(r_c)$ is the normalized pressure at the acoustic centre of the shell driving the f -hole vibrations. This is equivalent to an increase in the effective cavity volume from V to V_r by the inverse factor.

The model describes the uncoupled flexural vibrations of the signature plate modes shell and Helmholtz air vibrations. It has four independent degrees of freedom, resulting in the four $A0$, CBR , $BI-$ and $BI+$ independent normal modes, describing the coupled motions of the previously uncoupled plate and Helmholtz component modes.

The monopole sound source generated by the net volume flows in and out of the cavity, $(u_{\text{plates}} - u_{\text{Helmholtz}})$, determines the radiated far field sound pressure proportional to $\omega (u_{\text{plates}} - u_{\text{Helmholtz}})$. In contrast, the internal cavity pressure p_{internal} is given by $(u_{\text{plates}} - u_{\text{Helmholtz}}) / i \omega C_V$. Hence, the ratio of radiated sound pressure to the internal pressure is proportional to ω^2 at all frequencies in the monopole signature mode regime (see also reference [4]).

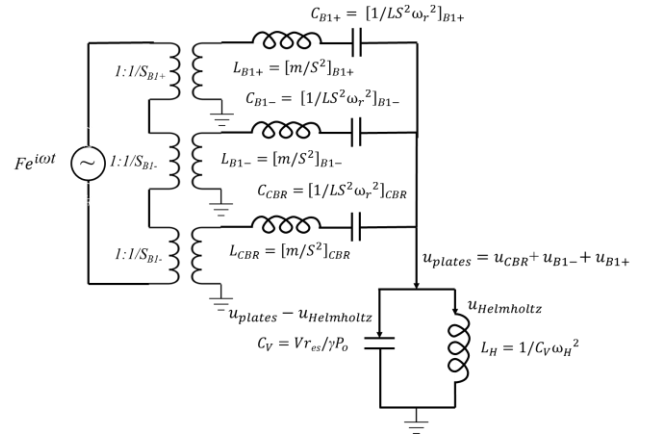


Figure 1. Equivalent circuit describing the coupling of the CBR , $BI-$ and $BI+$ signature modes measured in vacuum to the rigid-walled Helmholtz f -hole resonator.

At low frequencies, the outward flow of air out through the f -holes must match the inward flow from the flexural vibrations of the shell - the so-called *toothpaste* effect. Hence, the sound radiated by the CBR , $BI-$ and $BI+$ signature modes at low frequencies has to be matched in opposite polarity by the low frequency response of the $A0$ mode.

As a first approximation, we assume contributions from the higher order cavity modes can be neglected. It is then possible to describe the radiated sound over the whole of the monopole signature mode frequency range, simply from the CBR , $BI-$ and $BI+$ modes, with the addition of an $A0$ mode resonance of opposite polarity to cancel their low frequency response.

This is illustrated in figure 2 by Curtin-rig measurements [4] of the radiated sound from a violin, described over the whole signature mode region by first fitting resonances to the observed $BI-$ and $BI+$ peaks and then adding an $A0$ resonance of opposite polarity to give the required pressure cancellation at low frequencies. The only adjustable parameter required to fit the data below 400 Hz is then the $A0$ resonant frequency and its Q -value,

which is only important within a semi-tone or two of its resonance. Despite the neglect of the higher order cavity air modes excited, the fitted curves are in excellent agreement with the radiated pressure over the whole of the signature mode region below ~ 800 Hz. An additional small peak from the $A1$ cavity air resonance at ~ 500 Hz has also been added.

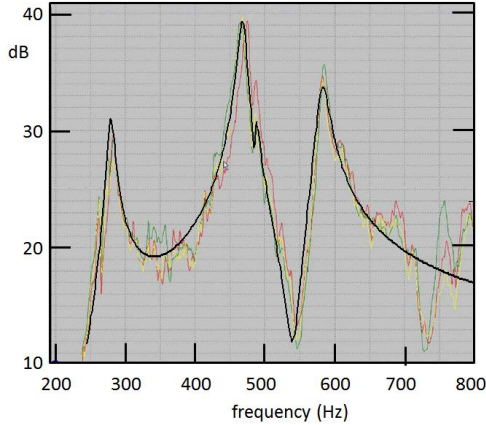


Figure 2. Radiated sound pressure at 12 azimuthal angles around the “Miles” violin fitted to the $A0$, $B1$ - and $B1+$ modes (and a small $A1$ internal cavity mode) using the generalized loudspeaker enclosure model (measurements courtesy of Borman).

Fits of similar quality have been made for several other instruments justifying the above approximation. For this particular violin, like many others, there was no significant contribution from the CBR mode. Because of the quality of the fit, one would expect the predicted frequency-squared dependence for the ratio of radiated to internal sound to be maintained throughout the isotropic, monopole radiation regime, for all instruments of the violin family.

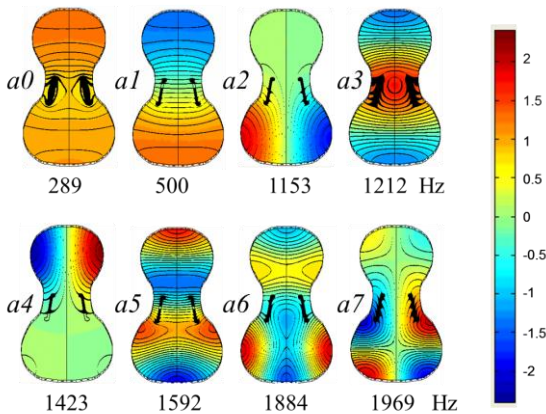


Figure 3. Computed central plane pressure of the first eight cavity air modes for a rigid, arched, violin cavity with f -holes. The pressure dependence on the central plane is only weakly perturbed by the f -holes.

Figures 3 and 4 illustrate the computed spatial variations of pressure of the first eight cavity air modes of a rigid walled, arched, violin shell with f -holes cut to give an $a0$ resonance at 289 Hz. Note the significant pressure changes along the length of the cavity for the $a0$ mode. This arises from the flow of air towards the f -holes from

both the upper and lower bouts. Because the $a0$ pressure is a minimum in line with the f -holes, there is little pressure difference along their length, in contrast to the assumptions made in the 2-degrees of freedom Shaw model [7] for the $a0$ and $a1$ air modes, which also neglects pressure variations in the lower bouts from flow towards the f -holes.

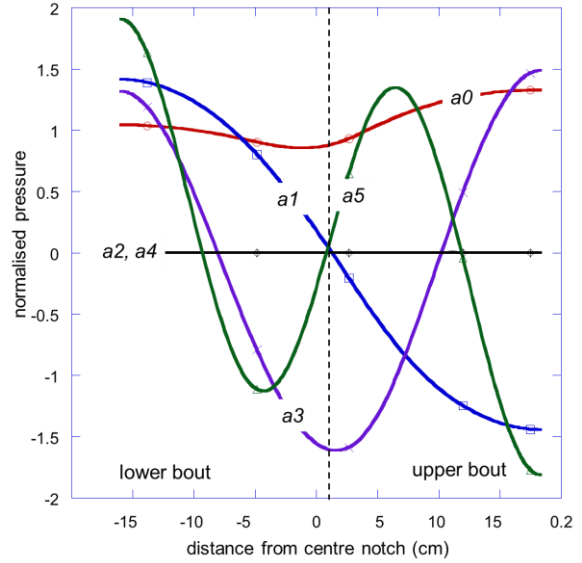


Figure 4. Spatial dependence of the internal pressure of the first six cavity air modes along the central axis of an arched violin shell with open f -holes measured from a position parallel to the central f -hole notches towards the upper bouts.

The $a0$, $a1$... cavity modes are the independent, non-interacting, normal modes of the rigid cavity body, which are perturbed to form the non-interacting $A0$, $A1$... normal modes, which includes their coupling to the flexural vibrations of the shell which excite them.

Our present interest is the acoustic pressure at the *acoustic centre* of the cavity, almost in-line with the f -hole notches. This is close to the nodes of the six lowest frequency cavity modes below ~ 1.6 kHz, other than $a0$ at ~ 290 Hz and $a3$ at ~ 1.2 kHz.

In the signature mode regime below 1 kHz, the sound pressure at the acoustic centre of the violin will therefore be dominated by the $a0$ mode. Because of the need for a uniform pressure within the cavity at low frequencies, additional contributions from the higher order $a1$ and $a3$ modes must also be excited. However, by definition, the $a1$ component will not contribute to the sound at the acoustic centre of the cavity, while the $a3$ contribution will be relatively weak because of its significantly higher frequency ~ 1.2 kHz. This justifies the assumptions made in fitting to the measurements in figure 2.

Around and above 1 kHz or so, the $a3$ mode, with one full wavelength along the length, will be coupled to any plate vibrations with nearby resonant frequencies having non-zero overlapping mode shapes. This will result in a set of normal modes describing their coupled vibrations, which will contribute to the acoustic pressure at the acoustic centre and hence also radiate through the f -holes.

Ideally, a calibrated microphone should be used to measure the internal sound pressure excited using a cali-

brated impact hammer – in just the same way that impulse measurements of the radiated sound are made.

As the measurements are made inside the cavity of the instrument itself, the sound pressure is high. This results in a large signal to noise ratio, free from external noise and complications from excited room acoustics. The measurements are highly reproducible when made at different locations, as the cavity volume is always the same, wherever the measurements are made.

Accurate comparisons of the acoustic properties of different instruments can therefore be made even in a noisy environment. In the measurements described here, a subminiature microphone was used, either freely supported or positioned at the acoustic centre on a very light curved wooden or plastic beam sprung between the outer edges of the f -holes. The acoustic centre can easily be determined by eliminating any weak $A1$ resonance, though this is easily recognizable and only weakly contaminates the measurements when present.

3. VIOLIN MEASUREMENTS

The measurements in Figure 5 illustrate strong contributions to the internal sound field from the $B1-$ and $B1+$ mode resonances, a small contribution from the CBR mode just below 400 Hz, and a splitting of the strong $A0$ mode resonance, almost certainly from coupling to the $B0$ longitudinal neck-body bending mode, with another much weaker resonance at a slightly lower frequency probably from the sideways-yaw or bending vibrations of the neck and body.

The assumption that the violin acts as a monopole source of sound in the signature mode region assumes that $(ka)^2 < 1$, where a is the effective size of the instrument. Around and above 800 Hz ($ka \sim 1$), the frequency

dependence will flatten off to a constant value. In this higher frequency range, the measurements reveal the anticipated strong group of resonances from normal modes involving the coupled $a3$ cavity air resonance at around 1.2 kHz and nearby in frequency shell vibrations. Such shell modes are likely to radiate significantly through the f -holes, which are close to the anti-node of the $a3$ mode.

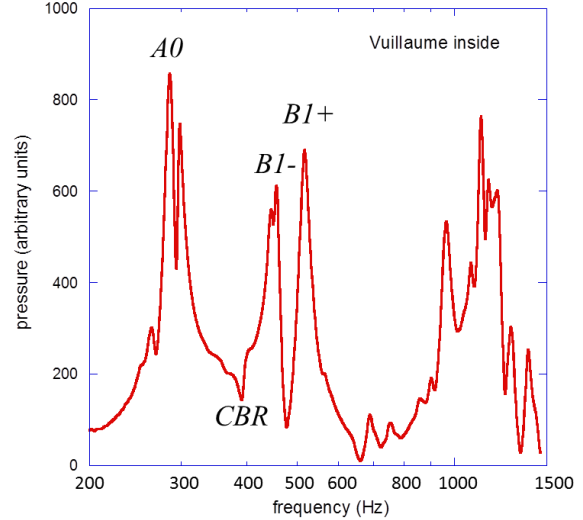


Figure 5. Typical measurements of the internal sound pressure spectrum at the acoustic centre of a Vuillaume violin with damped strings.

The measurements highlight the high quality of the internal sound measurements, which provide valuable information on the frequencies, damping and relative strengths of the four main signature modes – and a number of higher frequency modes as well.

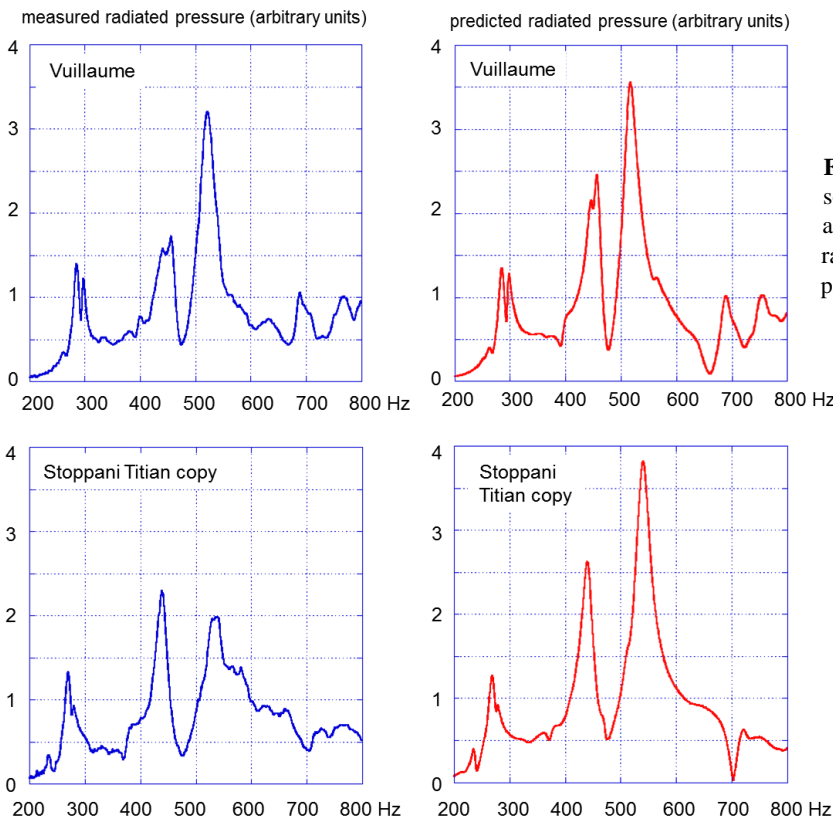


Figure 6. A comparison of the radiated sound pressure (left hand side) for two violins at a distance of ~ 30 cm and the predicted radiated pressure proportional to internal pressure $p \times \omega^2$.

Figure 6 compares the directionally averaged radiated sound pressure, measured at around 30 cm from two violins, with the radiated pressure derived from the predicted ω^2 - dependence of the measured sound at their acoustic centres. The radiation pressure was measured in the open air to eliminate contamination from room resonances – equivalent to making measurements in an anechoic chamber with a reflecting floor.

Despite the qualitative good agreement between the measured and derived frequency dependences, there are significant mode-dependent departures from the predicted ω^2 dependence, even allowing for the expected departures as ka approaches unity. Similar unexplained differences were observed by Jansson et al [4]. These could arise from systematic errors in the measurements themselves (near/far-field corrections, for example).

Such departures could also be explained, if the component bending modes of the $B1-$ and $B1+$ signature modes, vibrating in and out of phase with the breathing mode, contribute significantly to the radiated sound, despite the small volume changes and coupling to the internal air pressure predicted by shell model computations (Gough [1]). If so, the relative coupling strengths of the bending and breathing modes to the Helmholtz mode pressure fluctuation are likely to be very different from their coupling to the monopole radiation field modes. This would then account for the difference in relative heights of the $B1-$ and $B1+$ modes in the radiated and internal pressure measurements. Further investigations are required to clarify the origin of such differences.

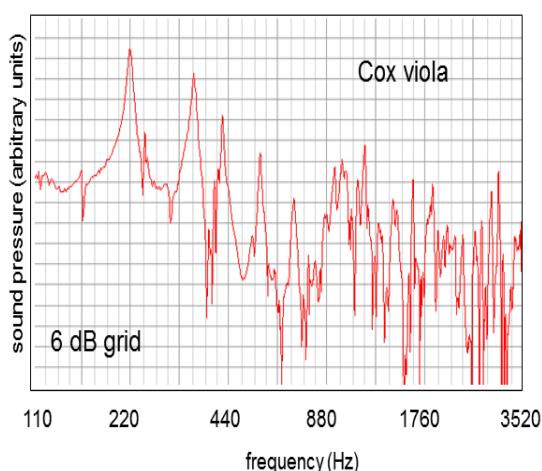


Figure 7. Acoustic centre pressure for a modern viola by Douglas Cox illustrating the high resolution, high signal to noise ratio and absence of contamination from room acoustics and external noise.

As an example of the use of internal sound measurements to characterize the acoustically important vibrational modes of instruments other than the violin, Figure 7 plots the acoustic centre pressure of a modern viola by Douglass Cox. This instrument has a number of extremely well defined modes in the signature mode region below

around 800 Hz and a clutch of strongly excited modes around 1 kHz and above.

Quite apart from their acoustic interest, the data illustrates the very high signal to noise ratio of such measurements, achieved using an inexpensive (~\$3) sub-miniature microphone. The data also illustrates the lack of any spurious resonances from room acoustics or extraneous noise from the relatively noisy environment of an Oberlin Acoustic Workshop. Reproducible, well-defined modal resonances are observed almost 50 dB below the dominant $A0$ resonance.

Interestingly, the measurements reveal four well-defined signature modes above the $A0$ resonance at 220 Hz. The shell model for string instrument presented separately (Gough [1]) suggests these are likely to be the CBR , $B1-$, $B1+$ and LD (longitudinal dipole/upper bout breathing) modes of the instrument.

4. DOUBLE BASS MEASUREMENTS

As examples of the use of internal air measurements to characterize the acoustic properties of very large instruments and as a personal tribute to the late Knut Guettler – an esteemed colleague, innovative acoustician and virtuoso bass player – some preliminary internal cavity measurements on a number of double basses will be described.

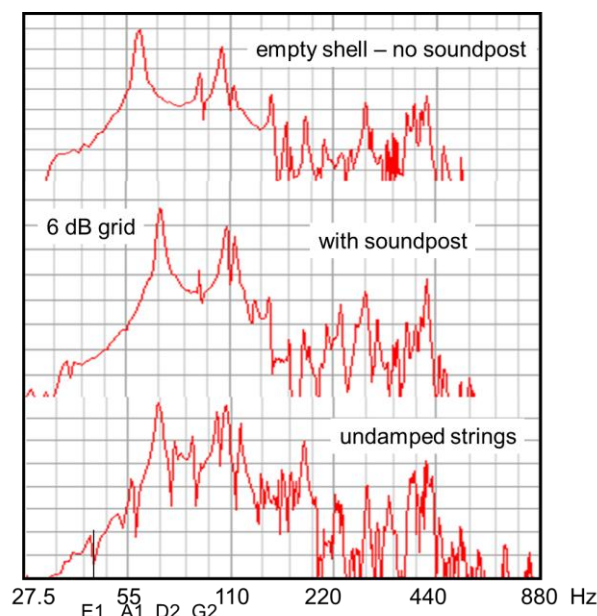


Figure 8. Internal sound measurements at the acoustic centre of modern Robbie McIntosh double bass without, with soundpost and undamped strings. The frequencies of the open strings are also indicated

The measurements were made at the 2012 Bass Oberlin Workshop in collaboration with the Violin Acoustics Workshop, where a dozen or so double basses were being temporarily stored in a single room, in various states of assembly. With invaluable assistance from several young makers, the following measurements were made in a single evening. The measurements should simply be seen

as an initial foray into the world of double bass acoustics. Future investigations including modal analysis measurements need to be made to confidently identify the principal modes observed.

Figure 8 shows internal pressure measurements for an outstanding modern bass by Robbie McIntosh, as judged by all the bass players and makers present. Measurements were made on this and other instruments without a sound post, with a soundpost, and set-up with a bridge and both damped and undamped strings.

The main features of the measurements on this instrument were reproduced on almost all dozen or so instruments, illustrated by some additional examples in Figure 9.

With the notable exception of the old European flat-back bass, the internal sound pressure, hence radiated sound at low frequencies, was dominated by only two strong resonances, rather than the three dominant $A0$, $B1$ - and $B1$ modes of the violin. The positions and relative strengths of the two prominent acoustic modes were similar for almost all the empty and set-up instruments. Additional analysis could extract comparative mode strengths, resonant frequencies and Q -values for all the observed resonant modes.

The lower mode of the empty McIntosh bass shell was raised on insertion of the soundpost from 60 Hz ($C2$) to 69 Hz ($C\#2$) and the upper mode from 105 Hz to 109 Hz, with two relatively small resonances of unidentified origin on either side.

Earlier admittance and laser modal analysis measurements by Askenfelt [7] and Brown [8] exhibited very similar peaks in a number of arched-back double basses and identified them as $A0$ resonances around 65 Hz and “top plate resonances” (equivalent to a $B1$ -breathing mode [1]) around 115 Hz, consistent with the above measurements. In making acoustic radiation measurements in an anechoic chamber, Brown noted that ‘disturbances from standing waves and other acoustic problems were problematic...especially below 100 Hz where location dependent deviations in response of up to 10 dB were observed’. Furthermore, he points out that the wavelength of the lowest note on a double bass is around 8m, while measurements had to be made at around 1m. The measured acoustic pressure at such frequencies were therefore strongly influenced by the non-radiating near field with pressure varying as $1/r^2$ rather than the $1/r$ dependence of the radiating sound. By deducing the radiating sound from internal cavity measurements one avoids all such problems.

In all the measurements, the bass was hand-held resting on a soft support in an upright position. Because of the existence of only a single mode above the Helmholtz $A0$ resonance and the relatively insensitive of any of the low frequency modes to the way the bass was held or supported, it seems likely that the component bending mode of the bass plays a less important acoustic role than for the violin and viola.

Also note the importance of the coupled string vibrations on the acoustical properties of the bass. The strings produce strong, narrow, resonances in the internal and

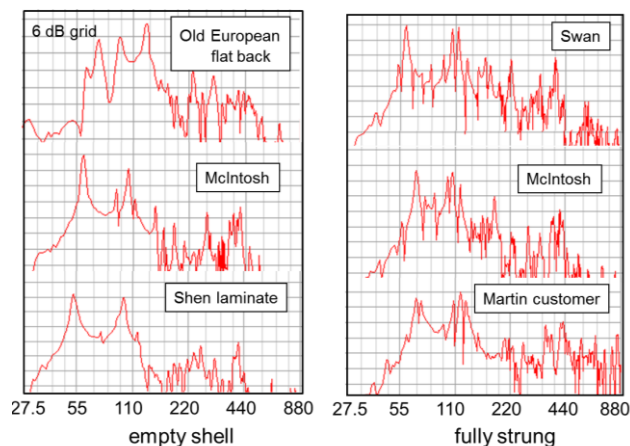


Figure 9. Internal sound measurements for a number of empty and fully-strung double basses

radiated sound. These will not only contribute to the liveliness of the bowed instrument when played with vibrato, but will also ring for much longer when plucked, as well as exciting slowly decaying transients at the start and end of any bowed note.

For all instruments of the violin family, measurements with damped strings help to identify the acoustical properties of the instrument itself, without the added complexity of the vibrating strings. However, if one wants to compare the sound quality of an instrument with physical measurements, the strings should arguably be left undamped - as they would be under normal playing conditions. For the player and listener, the transient response of an instrument associated with the undamped strings is important at the start and end of bowed notes of any length and also notes played with vibrato. This is particularly so for the bass, where the quality of plucked notes is often just as important as bowed notes.

5. CONCLUSIONS

Measurements of the sound pressure at the acoustic centre of the violin and related instruments are shown to characterise the acoustic modes of stringed instruments of any size at low frequencies, free of contamination from room acoustics and external noise.

A generalized model relating the internal to radiated sound pressures has been described. Demonstrations of the value of such measurements have been presented for the violin, viola and double bass. Using an inexpensive sub-miniature microphone with a laptop and soundcard, measurements are shown to be easily made in any environment including the luthier’s workshop.

Acknowledgments

I am particularly grateful to Anders Buen for drawing attention to the earlier double bass measurements of Andrew Brown, to all the students and bass makers at the 2012 Oberlin Bass Workshop who assisted with the bass measurements and set-ups, and to George Stoppani and Terry Borman for providing some of the experimental data.

6. REFERENCES

- [1] C. Gough, Vibrational modes of the violin family, in *Proc. SMAC2013*, Stockholm, 2013.
- [2] G. Bissinger, Structural Acoustics of good and bad violins, *J. Acoust. Soc. Am.*, vol. 124, no.3, 2008, pp. 1764-1773.
- [3] J. Curtin, Measuring Violin Sound radiation using an Impact Hammer, *J. Violin Soc. Am.: VSA Papers*, vol.22, no.1, 2009, pp. 186-209.
- [4] E.V. Jansson, L.A. Morset and Knut Guettler, Prediction of violin radiation properties in the 200-700 Hz range, *Proc. ISMA2001, Perugia*, pp 123-126, 2001.
- [5] G. Caldersmith, Guitar as a reflex enclosure, *J. Acoust. Soc. Am.*, vol. 63, 1978, pp. 1566-1575.
- [6] J. Schelleng, The violin as a circuit. *J. Acoust. Soc. Am.*, vol. 35, no.3, 1963, pp. 326-338.
- [7] E.A.G. Shaw, Cavity resonance in the violin: Network representation and the effect of damped and undamped rib holes, *J. Acoust. Soc. Am.*, vol. 7, no.1, 1990, pp. 398-410.
- [8] A. Askenfelt, Eigen modes and tone quality of the double bass, *STL-QPSR*, Vol 23, no.4, 1982, pp.19-174, with data reproduced in *The Science of String Instruments* (ed. T.D. Rossing), Springer, Chpt. Double Bass (Askenfelt), pp 259-278, 2010.
- [9] A.W. Brown, *Acoustical studies of flat-backed double bass*, PhD Thesis, Vienna, April 2004, and Documentation of double bass plate modes using the laser vibrometer, *Proc ISMA 2004, Nara*, 344-347, 2004.

Sensor Based Fatigue Recognition in Violin Playing

Tobias Grosshauser
Institute of Electronics
WearLab
ETH Zurich
grotobia@ethz.ch

Ulf Grossekaethöfer
Cognitive Interaction Technology
Center of Excellence
Bielefeld University
u.grossekaethoefer@tue.nl

Gerhard Tröster
Institute of Electronics
WearLab
ETH Zurich
troester@ife.ee.ethz.ch

ABSTRACT

Fatigue in daily instrumental exercising is a common problem in musical instrument learning. Musicians practice long and pause concepts are underrepresented resulting into performance shortcomings or even work-related injuries. The basic idea of the following investigation was to detect fatigue based on sensor data with sensors fixed on the musical instrument. Reliable fatigue detection would increase practicing efficiency, which results in faster progress in musical instrument learning while minimizing pain or preventing aftereffects due to too much physical stress. In this paper, we tested a sensor setup that is unobtrusively mounted to a violin and the frog of the violin bow. The sensors were 9 degree of freedom (9 DOF) acceleration sensors and gyroscopes. In 2 studies, amateur and professional violinists played defined music excerpts and scales at different speeds during sensor recordings. A one-hour uninterrupted playing regime with interspersed subjective evaluations of playing effort and tiredness was used. To test for changes in fatigue level, a scale and a segment of a music piece were intermittently played every 10 minutes. Sensor data were recorded and specific movement patterns recognized. Results show that the present setup can uncover features associated to individual fatigue during violin playing for amateurs, potentially helping to prevent injuries and to optimize practice regimes in these musicians.

1. INTRODUCTION

Daily exercising is a fundamental every-day task of musicians to learn their instrument. In this paper we examine violinists during a practice situation. In playing the violin, muscular work of the upper and lower arms is of pivotal importance. While movements of the lower arm are mainly involved in several displacements along and across strings, the upper arms, in addition to their involvement in arm displacements to allow fingering and bowing play a prominent role in holding the instrument (in the case of the left, fingering arm) and the bow (for the case of the right bowing-arm). In terms of effort,

holding tasks (also called static work tasks), are well known to be the most exhausting ones, resulting very soon into decreased muscle output due to fatigue (see for example [1]). Well known features of muscular fatigue are less functional capacity of the affected muscles together with the perception of increased difficulty to perform a determined task [2]. In low force tasks, fatigue has been shown to develop gradually with the associated perception of greater effort with a concomitant impossibility to accurately estimate the force of contraction [3]. The shoulder-neck region has been shown to be in particular ergonomic disadvantage regarding cardiovascular factors due to its elevated position relative to the heart level. This appears to be particularly the case for static work at lower levels of maximal force output [4]. In general, according to Brandfonbrener [5] and Lederman [6] musculoskeletal problems are frequent among professional instrumental musicians as well as music students and serious amateurs. The described scenario makes muscular fatigue of the shoulder -neck region a particular problem to be considered during daily exercise in violin practice, something being mandatory for the progress of instrumental learning in music. Muscular fatigue can result in several training limiting aspects. For example, in other domains, it has been shown that matching forces becomes less accurate and less consistent during fatiguing conditions in comparison to non-fatiguing ones, degrading force sense and, in so doing, postural control [7, 8]. Muscle fatigue can also be accompanied by soreness which has also been shown to interfere with the capacity to match a determined force and limb position [9]. Further exercising under fatigue conditions may lead to important delays in learning and in degraded quality of movements because it has been shown that muscle afferents change their firing pattern under fatigue [10]. For the reasons explained above, a sensor based set up able to index fatigue during violin practice appears of practical importance. We concentrate on the left and right arms as it has been recently shown that fatigue induced through repetitive motion of the arm affects the shoulder but not fingers' position sense [11]. Additional movements of the whole body or other parts of the body increasing with the time practiced are captured with the sensors attached to the violin.

We present here a sensor set up able to track changes in violin position and acceleration due to extensive exercise and fatigue of the arm during a 1-hour session uninter-

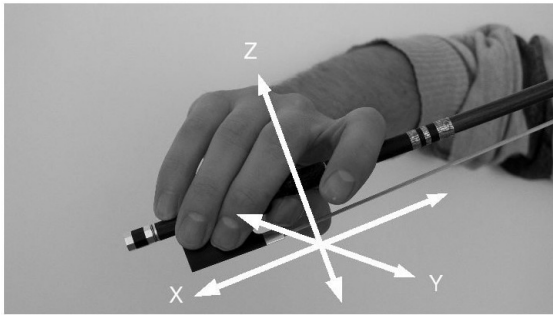


Figure 1. Depicted are the three motion axes while bowing the violin. In our setup, only the data of x- and y-axes are used.

rupted work on the violin. We hypothesized that fatigue subjective ratings of violinists would agree with changes in instrument's acceleration and position due to fatigue, as measured at the violin and characterized by additional acceleration patterns not present at baseline.

2. TECHNICAL DESCRIPTION, MEASUREMENTS AND STUDY DESIGN

Today, several technologies for motion and gestures' detection during instrumental musical playing exist. Diverse works presented several approaches to objectively capture gestures, particularly those associated to the bowing of string instruments. The most used measuring methods are based on the use of acceleration sensors and gyroscopes. Among others, the first sensors applied to violins, bow and violin gestures were the acceleration sensor on the bow in Bevilacqua et al. [12], left hand pressure and position sensor by Grosshauser et al. [13] and [14]. Similar approaches exist in the field of hyperinstruments with sensor based technology for new expression possibilities and the commercially available K bow.

2.1 Application and Description of the Used Sensors

All setups in our approach described in the following sections consist of a micro controller board and a 9 DOF MPU9150 from InvenSense (3 axes accelerometer, 3 axes gyroscopes and 3 axes magnetometer) sensor board placed on the violin (similar to Fig. 3 and Fig. 2) and on the frog of the bow. The assignment of the three axes to the bowing movement is shown in Fig. 1. The acceleration data of the right hand are used to identify the playing style. If the musician would play much louder or softer, the acceleration data would differ. Similar data of each individual player means the same playing style during the standardized measurement course every 10 minutes. The violin sensor is attached to the violin as depicted in Fig. 3. This allows fast removal and the same position on every individual violin.

Due to latency issues, all setups are realized with wire-based data transfer, but wireless data transfer extended to a wireless sensor network would be also possible. This allows for higher data transfer rates without dropouts.

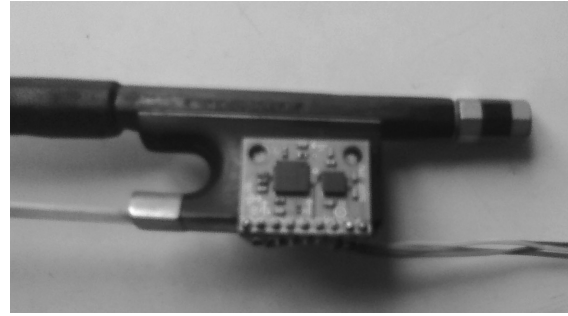


Figure 2. Depicted is the acceleration board fixed to the bow. Fixation is to the inner side (from the violinist's perspective) of the frog, to minimize disturbances while playing.



Figure 3. Shown is the sensor board attached to the violin. This typical "chin rest" mount allows stable fixation at the same position on every violin.

2.2 Measurements and Experiment Setups

Two independent sets of measurements with 6 participants each were performed on the individual's violin and bow. An acoustical metronome paced playing speed at 60 bpm. To induce task-related fatigue, all participants performed one hour of uninterrupted playing regime practicing in their own preferred manner. No special instructions were given on how to practice. All the segments played were placed on a music stand. All participants performed standing as this is the common practice position for violin performers with the exception of those instances in which orchestral and some chamber music ensembles play. All measurements were performed in an acoustically isolated room and musicians were left alone while exercising to most vividly resemble a typical practice situation. Before measurements, volunteers were allowed to warm-up in their preferred way for a brief period of time.

The experimenter entered briefly (ca. 20 sec) the room after each 10 min of uninterrupted practice to gather probes for changes in fatigue level. To this aim, a H- major scale and the first sequence of Kreutzer Etude Nr. 2. were played

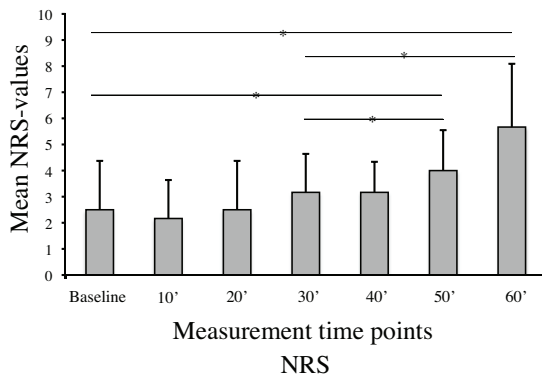


Table 1. Depicted are the mean NRS values across measurement time points. Line with asterisks denote participants' significant increments of fatigue perception as computed with rank statistics. Error bars correspond to standard deviations.

at a fixed metronome beat, and subjective evaluations (the rating scale is based on Likert's scale [15]) of playing effort and tiredness was performed. Total measurement set up lasted for approximately 70 min, whereby 60 min divided into 6 blocks of 10 min. were devoted to practice and the rest to the intermittent fatigue evaluations. After the measurement the subjects filled a form containing some questions regarding the experiment and asked for further comments.

2.3 Results of the First Measurement

To identify fatigue sensations of participants attributable to uninterrupted violin practice, the single values of the NRS for every measurement point were used to compute a group statistic by means of non-parametric statistics as explained below (see Table 1).

To identify objective differences in violin position and acceleration during uninterrupted practice, we calculated a total position and acceleration index score. All the single values for any participant were then considered for the group statistic containing 7 measurement time points for the test segments as described under 2.1. In the case of sensor data and for simplicity, we analyzed three points in time here: Baseline, after 30' and 60' of interrupted practice. For statistical analyses, non parametric tests for related samples were used, because number of participants was low and data distributions did not conform the normality criterion. The Friedman's test, a Two-Way Analysis of Variance by Ranks was used to uncover differences including more than two measurement time points. The related samples Wilcoxon Signed Rank Test was used for single comparisons. For all analyses a significance level of $p = 0.05$ was chosen. Calculations were run by means of SPSS19.

2.3.1 Subjective ratings

Table 2 shows the subjective NRS values used to quantify fatigue for the single participants.

The Friedman's test including 7 measurement times

Code	Sex	Age	Playing Experience	Skills	Sensors' influence	Baseline	10'	20'	30'	40'	50'	60'
P1	f	30	>10	professional	none	2	2	3	2	2	3	4
P2	f	22	>10	advanced	none	1	1	1	2	2	3	3
P3	f	24	>10	advanced	none	1	1	1	3	3	3	9
P4	f	19	>10	advanced	none	3	2	2	3	4	4	4
P5	f	24	>10	Performing Arts/Master	none	6	5	6	6	5	7	6
P6	m	36	>10	professional	none	2	2	2	3	3	4	8
Mean						2.6	2.2	2.6	3.2	3.2	4.0	5.2
SD						2.1	1.6	2.1	1.6	1.3	1.7	2.4

Table 2. Subjective NRS values for fatigue for the single participants, demographic data and evaluation of measurement setup. Please note that the sensor setup did not influence participants' practice. SD = Standard deviation.

Code	Baseline	30'	60'	Number of datapoints
P1	9865	10107	12271	986
P2	6783	7225	7005	995
P3	2434	4071	4289	850
P4	7806	10038	10787	986
P5	3696	4520	4628	986
P6	2781	4403	4203	986

Table 3. Single sensor measurement values of participants containing the added position change and acceleration of the violin during playing across three measurement time points at baseline, after 30' and 60'.

(baseline and after 10, 20, 30, 40, 50 and 60) was significant, indicating that over all, there was a clear increment in fatigue sensations during uninterrupted practice ($p = 0.000$).

Further statistical analyses with the Wilcoxon Signed Rank Test revealed that this significance was due to an increase of test values from baseline to 50 ($p = 0.024$) and 60min ($p = 0.042$) as well as from 30 to 50 ($p = 0.025$) and 60min ($p = 0.042$). All other comparisons including baseline and 10, 20, 30 and 40min as well as 50 to 60min comparisons were not significant (all p -values greater than $p = 0.05$).

2.3.2 Objective Data

Table 3 shows the single measurement values. The Friedman's test including 3 measurement times (baseline, after 30 min and after 60 min) was significant, indicating that position and acceleration values increased from baseline to last measurement ($p = 0.009$). The position change and acceleration values are calculated by summarizing the differences of each absolute sensor value of each axis absolute values. The acceleration and violin position sensor values are exemplary shown of one participant in Fig. 4 and the values of all subjects in table 3. Further the acceleration values of the bow hand are shown. The similarity of the bow acceleration values shows, that the phrases are played similar in the first and last recording.

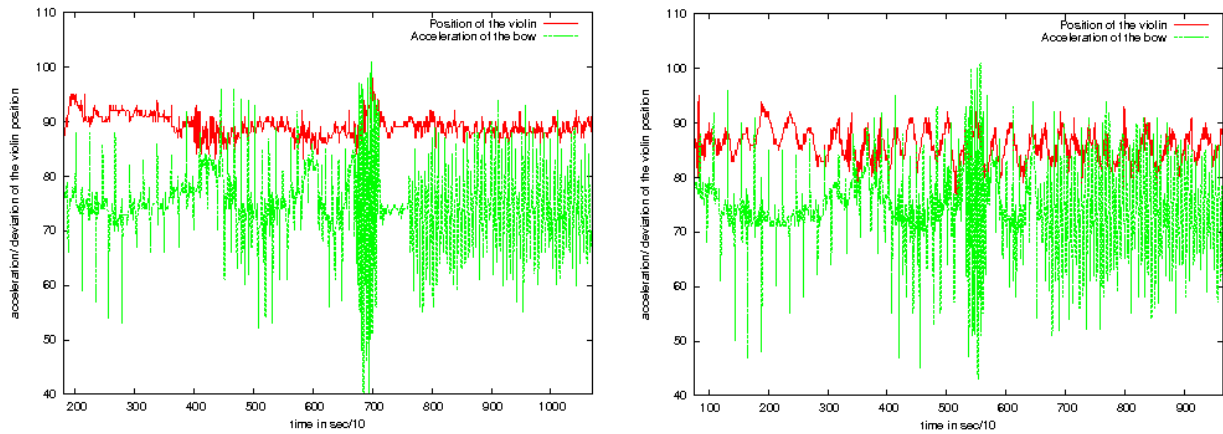


Figure 4. Depicted is the increase of the motion and position change of the violin for a representative participant. Please note, that the acceleration value of the bowing hand does not increase between first and last measurement, indicating the same playing style of the technical exercises.

Nr.	Sex	Age	Experience	Type	Influ
1	m	60	>10	prof	no
2	f	26	>10	adv	no
3	f	13	5 – 10	beg	no
4	f	54	>10	am	no
5	f	34	>10	prof	no
6	m	12	7	beg	no

Table 4. Demographic data of the single participants of the second test group.

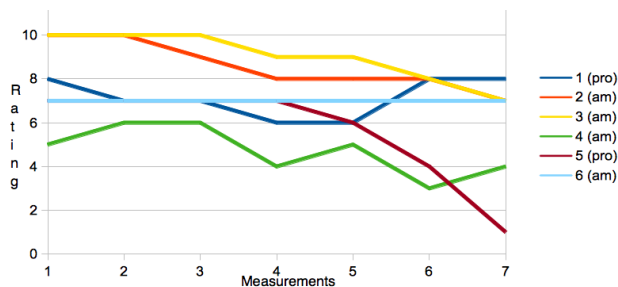


Figure 5. Subjective values for fatigue for the single participants.

2.4 Results of the Second Measurement

A second measurement was conducted with different test subjects to investigate the particular difference between amateur and professional musicians. The statistical data of the participants is shown in table 4.

2.4.1 Subjective Ratings

Fig. 5 shows the individual perceived fatigue of each participant during the second experiment. The stability of 6 (am) is exceptional, but might be caused by the age. The data of participant 1 (pro) shows a decreasing fatigue, caused by an unplanned extra pause.

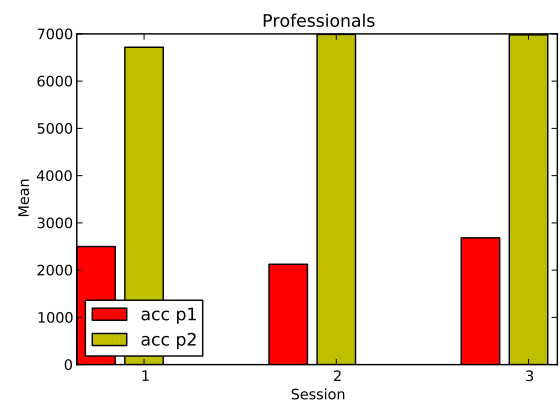


Figure 6. Mean of the x-axis data of the accelerometer sensor from the professional violinists (Session 1 baseline, 2 after 30 min, 3 after 60min).

2.4.2 Discussion of the Measurement Data

We used data from the x-axis of the accelerometer sensors to derive mean and standard deviation values for each participants sessions. The accelerometer measurements correspond to positions of the violin in relation to the earth center, i.e., the sensor mostly measures gravitational acceleration. In this context, the mean value indicates the participants postures and how they position their instruments during the recording. The standard deviation value corresponds to how intense the instrument was moving.

The results of the second measurement are illustrated in Fig. 4. Fig. 6 shows mean values for the first, middle and last session, for both professional violin players that participated in our study. Fig. 7 displays the matching standard deviation values. In contrast, Fig. 8 and 9 contain similar plots for amateur players.

These results indicate that both professional players are able to provide stable body postures during the complete course of the experiment. The average accelerometer values, as well as their standard deviation stay on a constant level. Please note the different valuation of the violinists

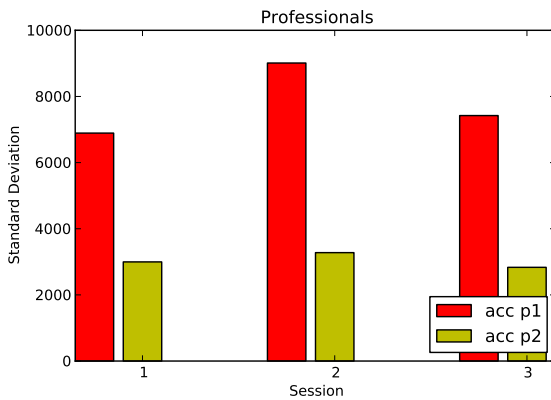


Figure 7. Standard deviation of the x-axis recordings of the accelerometer sensor from the professional violinist (Session 1 baseline, 2 after 30 min, 3 after 60min).

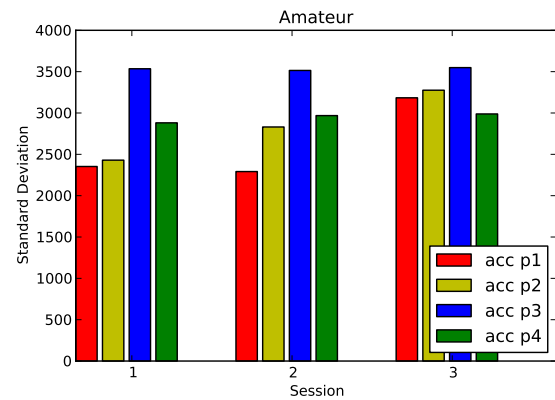


Figure 9. Standard deviation of the x-axis data of the accelerometer sensor from the amateur violinist (Session 1 baseline, 2 after 30 min, 3 after 60min).

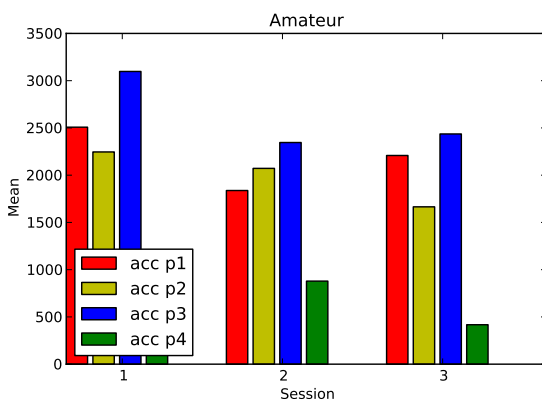


Figure 8. Mean of the x-axis data of the accelerometer sensor from the amateur violinist (Session 1 baseline, 2 after 30 min, 3 after 60min).

— an indicator how different the players hold their instruments. Only the measurements of professional participant *p1* reveal increased movements during the medium session.

In contrast, the mean values of the accelerometer recordings of the amateur violin players uncover a different picture. The results indicate that amateurs considerably vary their location during the recordings. All four participants from this group show substantial changes in their posing as represented by average accelerometer values. The standard deviations do vary greatly for 2 out of 4 players (*p1* and *p2*) while the remaining measurements stay comparatively stable.

3. CONCLUSION

Musicians' practice overly long placing themselves at the risk of performance and practice shortcomings and bodily pain. The described basic measurement's setup represents a further step to gathering objective data regarding fatigue during instrumental music making. Evaluation was performed here with a sensor setup fixed to the individual violin of each participant. This is of importance, be-

cause changes of instrument being used impose new challenges to the player, which might potentially result in fatigue attributable the instrument being used and not to the task itself. This is an important aspect to be considered in any measurement of fatigue, which has been shown to be highly task and context dependent. More over, even if subjective data on fatigue perception was congruent with the objective data, violinists, in comparison to the sensor data, underestimated their fatigue with their subjective ratings. Therefore, wearable technology might be of great value not only for musicians but also for the development of experimental settings aiming at assessing complex processes like for example muscular fatigue during instrumental playing. In addition, we believe that the presented measurement setup and sensors make possible to measure parameters of fatigue independent of playing and exercising situation. Next steps will also consider the inclusion of smaller inertial sensors and tests with other instrumentalists like for example flutists. Larger data series revealing fatigue features uncovered by means of unobtrusive sensor technologies might potentially help to prevent injuries and to optimize practice regimes of musicians.

4. REFERENCES

- [1] R. Birbaumer, N & Schmidt, *Biologische Psychologie*. Springer, 1996, no. 14, ch. Arbeit, Sport, Rehabilitation, Alter, Umwelt.
- [2] B. K. Barry and R. M. Enoka, "The neurobiology of muscle fatigue: 15 years later," *Integr Comp Biol*, vol. 47, no. 4, pp. 465–473, Oct 2007. [Online]. Available: <http://www.hubmed.org/display.cgi?uids=21672855>
- [3] H. W. Jones LA, "Effects of fatigue on force sensation," *Exp Neurol*, vol. 81, no. 3, pp. 640–650, 1983.
- [4] W. Hildebrandt, J. Herrmann, and J. Stegemann, "Fluid balance versus blood flow autoregulation in the elevated human limb: the role of venous collapse," *Eur J Appl Physiol Occup Physiol*, vol. 69,

- no. 2, pp. 127–131, 1994. [Online]. Available: <http://www.hubmed.org/display.cgi?uids=7805666>
- [5] A. G. Brandfonbrener, “Musculoskeletal problems of instrumental musicians,” in *Hand clinics*, vol. 19(2), 2003, pp. 231–239.
- [6] R. J. Lederman, “Neuromuscular and musculoskeletal problems in instrumental musicians,” in *Muscle and nerve*, vol. 27(5), 2003, pp. 549–561.
- [7] U. Proske, J. E. Gregory, D. L. Morgan, P. Percival, N. S. Weerakkody, and B. J. Canny, “Force matching errors following eccentric exercise,” *Hum Mov Sci*, vol. 23, no. 3-4, pp. 365–378, Oct 2004. [Online]. Available: <http://www.hubmed.org/display.cgi?uids=15541523>
- [8] N. Vuillerme and M. Boisgontier, “Muscle fatigue degrades force sense at the ankle joint,” *Gait Posture*, vol. 28, no. 3, pp. 521–524, Oct 2008. [Online]. Available: <http://www.hubmed.org/display.cgi?uids=18434157>
- [9] U. Proske, N. S. Weerakkody, P. Percival, D. L. Morgan, J. E. Gregory, and B. J. Canny, “Force-matching errors after eccentric exercise attributed to muscle soreness,” *Clin Exp Pharmacol Physiol*, vol. 30, no. 8, pp. 576–579, Aug 2003. [Online]. Available: <http://www.hubmed.org/display.cgi?uids=12890182>
- [10] J. L. Taylor, J. E. Butler, and S. C. Gandevia, “Changes in muscle afferents, motoneurons and motor drive during muscle fatigue,” *Eur J Appl Physiol*, vol. 83, no. 2-3, pp. 106–115, Oct 2000. [Online]. Available: <http://www.hubmed.org/display.cgi?uids=11104051>
- [11] K. Emery and J. N. Côté, “Repetitive arm motion-induced fatigue affects shoulder but not endpoint position sense,” *Exp Brain Res*, vol. 216, no. 4, pp. 553–564, Feb 2012. [Online]. Available: <http://www.hubmed.org/display.cgi?uids=22124803>
- [12] F. Bevilacqua, N. Rasamimanana, E. Fléty, S. Lemouton, and F. Baschet, “The augmented violin project: research, composition and performance report,” in *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*. Paris, France, France: IRCAM — Centre Pompidou, 2006, pp. 402–406.
- [13] T. Grosshauser, “Low force pressure measurement: Pressure sensor matrices for gesture analysis, stiffness recognition and augmented instruments,” in *8th International Conference on New Interfaces for Musical Expression NIME08*, S. G. Volpe, A. Camurri, Ed., 2008.
- [14] T. Grosshauser, U. Großekathöfer, and T. Hermann, “New sensors and pattern recognition techniques for string instruments,” in *International Conference on New Interfaces for Musical Expression, NIME2010*, Sydney, Australia, 2010.
- [15] R. Likert, “A technique for the measurement of attitudes,” in *Archives of Psychology* 140, 1932, pp. 1–55.

New Devices for Testing the Stiffness Characteristics of Free Violin Plates

Ruiqing Jia

Department of Mechanical Engineering
China University of Mining and Technology
ruiqingjia@yahoo.com

Lei Fu

bjfulei.good@163.com

Ailin Zhang

Department of Engineering
University of Cambridge
az304@cam.ac.uk

ABSTRACT

The mode tuning of violin plates has been demonstrated to play an important role in building an instrument. As an aid to tune free plates to the right resonance frequencies, modern measuring techniques such as the Chladni-pattern method and holographic interferometry show luthiers how the eigenmodes and frequencies change with adjusting the stiffness and mass of the plate intuitively. However, such methods cannot provide quantitative data or be recorded by luthiers effectively. Thus, the art of tuning is still dominated by empirical knowledge of violinmakers. In this article, three new devices have been developed to test the effective twisting and bending stiffness of free plates instead of traditional manual tests. Experiments are performed on four pairs of violin plates to show that these devices could reliably produce recordable data related to stiffness characteristics. In addition to a detailed calculation of the stiffness of the tested plates, potential directions for future investigation are also discussed.

1. INTRODUCTION

Traditional tuning of soundboards of bowed string instruments to an ideal pitch is usually performed by luthiers with extensive skills. Adjustments are made to the stiffness and mass of plates by scraping away wood in the tuning process due to a high dependence of eigenfrequencies on the stiffness characteristics. For assessing the stiffness characteristics that give rise to eigenmodes, luthiers twist, bend or push free plates slightly by hand and then judge by experience; in practice this requires appropriate skills and a long-term training.

Research has progressed to the point where suitable geometrical modifications of free plates can be suggested during the tuning. However, few of the methods provide quantitative data for ordinary makers. The variations in configurations of eigenmodes generated by small modifications of free plates might be too subtle to detect. Another problem is that the same geometrical adjustments might have entirely different effects on different plates since the vibrational properties of wood

vary from plate to plate.

This paper presents a new method to help violin makers get more informed and recordable data in the making of soundboards. Three new devices, which can closely mimic manual testing, are designed to assess the stiffness characteristics of free violin plates that dominate vibration modes 1, 2 and 5. Following some background presented in Section 2, detailed descriptions of the design and functionality of the three devices are outlined in Section 3. Experimental procedures and results are then provided in Section 4 and 5. Conclusions are given in the final section.

2. BACKGROUND

Savart [1] was the first researcher to look closely at the “tap tone” of violin plates by investigating valuable Italian violins. During the 1930s, Backhaus [2, 3] mapped out different vibration patterns at their main resonance frequencies by an electrical method. Hutchins [4, 5] then verified the findings of Savart and further summarized the relation of tap tones to acoustic properties of free plates in the early 1950s. She [6] also developed a modified Chladni-pattern technique, for visibly displaying mode shapes of free violin plates. In her study the manual tests used by makers were explained as an effective method to check the principal stiffness characteristics of modes 1, 2 and 5. In the late 1960’s, a new testing technique was devised by Stetson [6] with his research on holographic interferometry. By applying this highly-sensitive technology to a pair of free top and back plates, he showed contour maps of vibration amplitude of vibration modes at discrete frequencies.

The question then arose: how can a maker effectively use these scientific investigations to make correct decisions during the actual process of construction? Many efforts have been made to find answers to this question. Following the measurements of Backhaus, Meinel [7] tested instruments made by himself with controlled alteration of various factors, especially the thickness of the plates, during the 1930s. He explored the correlation between thickness and the vibration nodal lines, to provide the possibility of improving unassembled plates deliberately by shaving the surface. Roger [8] whom Hutchins had worked with since 1979 contributed to transfer Hutchins’s knowledge of plate tuning to a complex computer model by Finite Element analysis. Taking advantage of this model, he successfully

tuned thousands of plates. But his tuning process was too impractical for ordinary luthiers. Another notable work was carried out by Schleske [9] in 1990s, which documented the influence of thickness graduation on the eigenfrequencies based on results of Jansson [10]. He intentionally combined the modification process of free plates with measurements of the input admittance and mode shapes although no clues were discovered about the relation of the tuning of free plates to vibration modes of the corpus.

3. DESCRIPTION OF NEW DEVICES

3.1 Twisting Stiffness Testing Device

In the design of this device, the following criteria were taken into account: the prime requisite is a device that can produce a precisely controlled moment, which is easily reproducible and reliable, to the plate; then yield a recordable response. The testing device described here can fulfill above requirements. The essential elements of the twisting stiffness testing device are shown schematically in Figure 1. It consists of two clamping elements, a loading wheel, and a rigid shaft attached by bearing connections to a hexagonal metal base. Each vulcanized plastic clamp can firmly hold a small area of the plate in a manner resembling that of a human maker, without damage. To save weight, the metal base was made of hollow aluminum alloy blocks with minimized size.

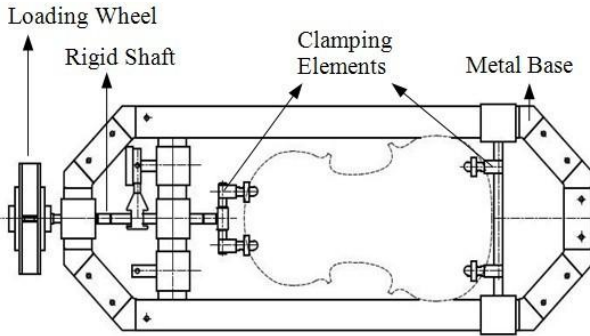


Figure 1. Schematic and mechanical structure of the twisting stiffness testing device for measuring the main stiffness of mode 1.

To provide controllable moments, a loading wheel with a display meter scale, as shown in Figure 2, was used to give deformation reminiscent of vibration mode 1. A rigid shaft which passes through the centres of the wheel and the left clamping element is attached by threaded connection to the wheel. A known weight is fastened at the end of a rope, which is wrapped around the wheel. When the rope is pulled by the weight, a torque is transferred to the left clamps by the shaft. The right clamps hold the plate motionless. The advantage of using a loading wheel is that the exact amplitude of torque M_e can be calculated precisely as

$$M_e = mgr \quad (1)$$

where m is the known weight and r the radius of the wheel. To additionally save weight, the wheel employed

in the testing device has a hollow rim made of ABS plastic.

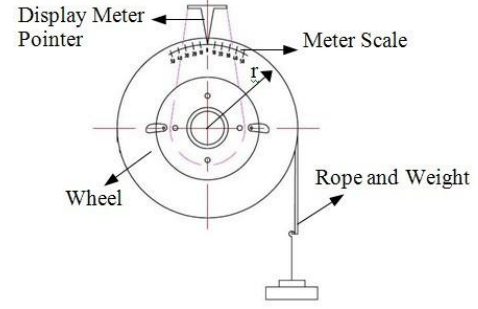


Figure 2. Sketch of the loading wheel used in the twisting stiffness testing device.

The resulting angle of twist θ can be read from a display meter pointer, as illustrated in Figure 2. However, for improved accuracy of measuring, a dial indicator as shown in Figure 3 was also set up on the shaft through a measuring rod. It can measure the small distance variation δ at the end of the rod caused by rotation of the shaft and plate when the torque is applied. If θ is assumed to be small then $\sin\theta$ is approximately equal to θ . Thus if the length of the measuring rod is R ,

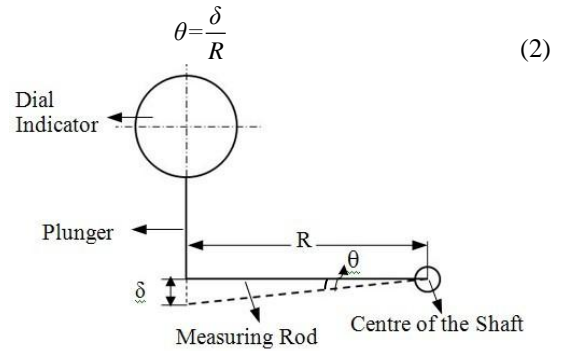


Figure 3. Cross-sectional view of the dial indicator set up on the shaft.

3.2 Bending Stiffness Testing Devices

The essential elements of the two bending moment testing devices are shown in Figure 4 and 5. The first feature to note is the use of two loading wheels and rigid shafts set up with dial indicators. As discussed in the previous section, this arrangement was used in order to closely mimic the actions of makers when they are bending free plates. For this purpose, the same weights were suspended from both loading wheels so that two equal moments can be applied together to bend the tested plate. The loading wheels, clamps, and rigid shafts are the same as those described in Section 3.1. The metal bases were also made out of the same material mentioned above with minimized size. For both devices, the bending moments are given by equation (1). The computing of bending curvatures will be presented in Section 5.2 with more details.

The bending stiffness testing device in Figure 4 was designed to measure the relative stiffness characteristics for mode 2. In order to accomplish this, both clamps hold the same end of the plate and known weights are added to each wheel for squeezing and bending the plate as a

maker does. This design allows the cross-grain stiffness of both upper and lower areas of a free plate to be tested.

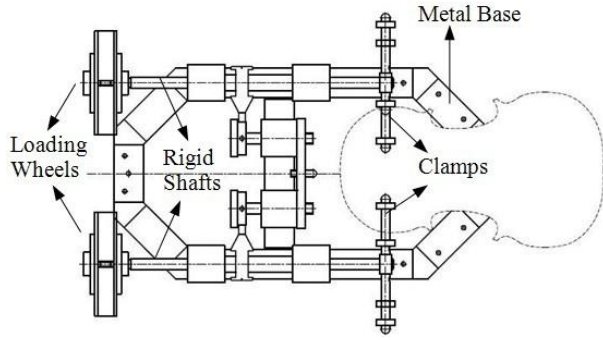


Figure 4. Schematic and mechanical structure of the bending moment testing device for measuring the stiffness of mode 2.

To provide a bending moment necessary to initiate mode 5, another bending stiffness testing device shown in Figure 5 was designed. For measuring the long-grain stiffness of unassembled plates, luthiers always hold the two ends of the plate and push down the arch in the middle. This action is readily achievable through the arrangement of two wheels and clamps as shown below.

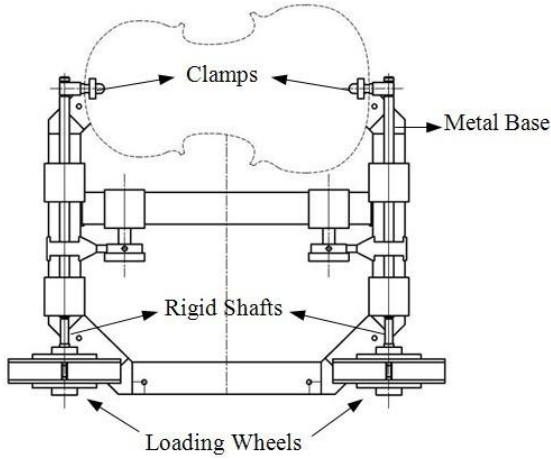


Figure 5. Schematic and mechanical structure of the bending moment testing device for measuring the stiffness of mode 5.

4. EXPERIMENTS

4.1 Experiment Procedures

Four pairs of free top and back plates left in the white were used for measurements. All plates were not quite finished so that they were thicker than usual plates. The tested top plates had a thickness between 2.4 and 4 mm. The thickness of back plates was between 3 and 4.3 mm. The spruce for the top plate had 7-9 rings/cm and a density of 450 kg/m³. The maple for the back plate had 5-6 rings/cm and a density of 510 kg/m³.

Before the measurement, the tested free plate must be held static in a balanced position by the testing device, as shown in Figure 6. The clamping points need to be symmetric about the central line of the tested plate. The plunger of dial indicator must be zeroed and adjusted

appropriately so that its tip is perpendicular to the measuring rod.

During the experiments, certain known weights were added. The loading weight increased gradually from 0.5 kg up to 2 or 3 kg in steps of 0.5 kg. The numerical reading showed by the dial indicator was recorded for each loading weight. As for measurements of mode 2 depicted in Figure 6 (b), the relative stiffnesses of both ends of the plate were measured. Therefore, each plate was tested in four ways: one twisting experiment and three bending experiments.

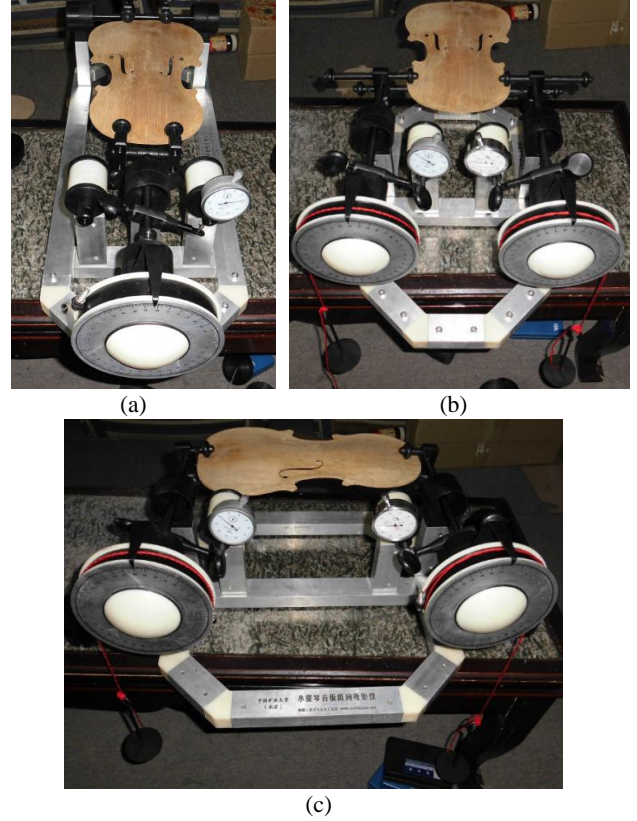


Figure 6. Measurements in progress: (a) Measurements of the main stiffness of mode 1; (b) measurements of the relative stiffness of mode 2; (c) measurements of the main stiffness of mode 5.

4.2 Twisting Stiffness

Figure 7 illustrates a plate fixed at one end and twisted at the other end due to the action of torque M_e in the mode 1 experiments. When twisted, the plate imaged as a horizontal line AB moves through an angle θ to AB'. Thus the resulting angle of twist θ , which is determined by the angle δ measured by the dial indicator, can be calculated as

$$\theta = \frac{M_e L}{S_e} \quad (3)$$

where L denotes the distance between two clamping points A and B and S_e the effective twisting rigidity .

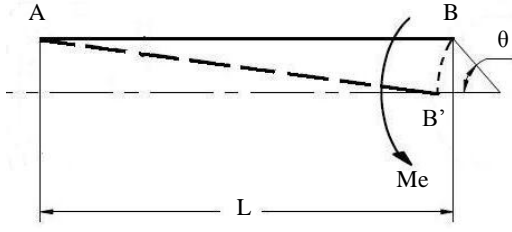


Figure 7. Torque applying on a free plate.

Thus we might write the effective twisting stiffness S_e as

$$S_e = kL \quad (4)$$

where k is the ratio of torque M_e to corresponding angle θ .

4.3 Bending Stiffness

In the bending experiments, the bending moments M are imposing on both clamping points A and B as shown in Figure 10. The resulting bending curvature $1/\rho$ can be shown to be

$$\frac{1}{\rho} = \frac{M}{S_b} \quad (5)$$

where S_b is the effective bending stiffness characteristic, also called flexural rigidity.

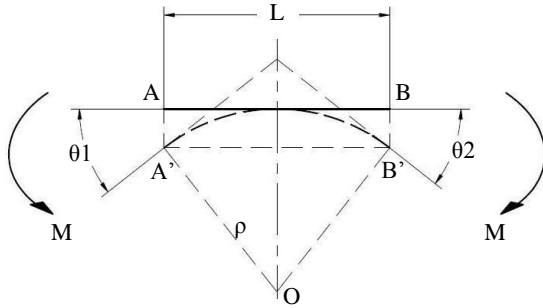


Figure 10. Bending moment applying on a free plate.

If A'B' is the deflection curve of the plate, and O is the center of curvature, the radius of curvature ρ is given by

$$\frac{1}{\rho} = \frac{\theta_1 + \theta_2}{L} \quad (6)$$

where θ_1 and θ_2 are bending angles determined by two dial indicators separately and the distance L between two clamping points. Thus the effective bending stiffness S_b can be written as

$$S_b = M\rho = \frac{ML}{\theta_1 + \theta_2} \quad (7)$$

5. RESULTS

5.1 Twisting Stiffness

Figure 8 is a scatter plot of data from the four tested top plates in twisting experiments for mode 1. The torque M_e is graphed on one axis while the angle of twist θ on the other one. The blue, red, green and yellow points denote four groups of data obtained from belly plate 1, 2, 3 and 4 separately. The tendencies for all data are very

pronounced, which suggest variables M_e and θ are strongly correlated in the tested region. Similarly, the results of four back plates scattered in Figure 9 also indicate the relation between the twisting moment and angle should be linear. These results demonstrate good accuracy of measurements made with this device.

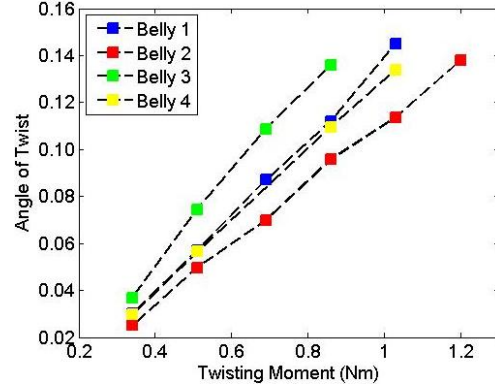


Figure 8. Results of top plates measured by the twisting stiffness testing device.

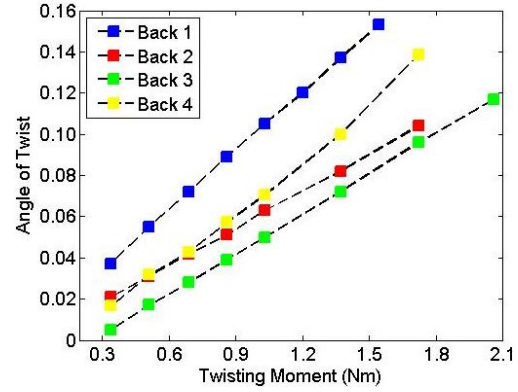


Figure 9. Results of back plates measured by the twisting stiffness testing device.

By means of linear least-squares fitting, best-fit straight lines through the data can be generated. The values of k , which are the line slopes, will be subsequently determined and used for calculation of stiffness. Thus the effective twisting stiffness characteristics of four pairs of free plates are calculated by equation (4) as shown in the table below.

Mode 1	1	2	3	4
Belly	7.4964	7.9854	7.4289	8.5775
Back	6.7631	7.1498	8.3228	8.5883

Table 1. Effective twisting stiffness of mode 1.

5.2 Bending Stiffness

Scatter points shown in Figure 11 and 12 demonstrate sets of data obtained from bending experiments of mode 2. For all figures, the bending moment M is graphed on one axis and the bending curvature $1/\rho$ is on the other one. The blue, red, green and yellow points indicate experimental results of plate 1, 2, 3 and 4 respectively.

Results for the upper end of the plates are illustrated in Figure 11 (a) and Figure 12 (a) while those for the lower end are in Figure 11 (b) and Figure 12 (b). As expected, a general tendency for the points to rise to the right of the graph can be seen from all figures. This clear linearity justifies the capabilities of the testing device for mode 2. On basis of this, the cross-grain stiffness characteristics of four pairs of free plates shown in Table 2 and 3 are calculated by equation (7).

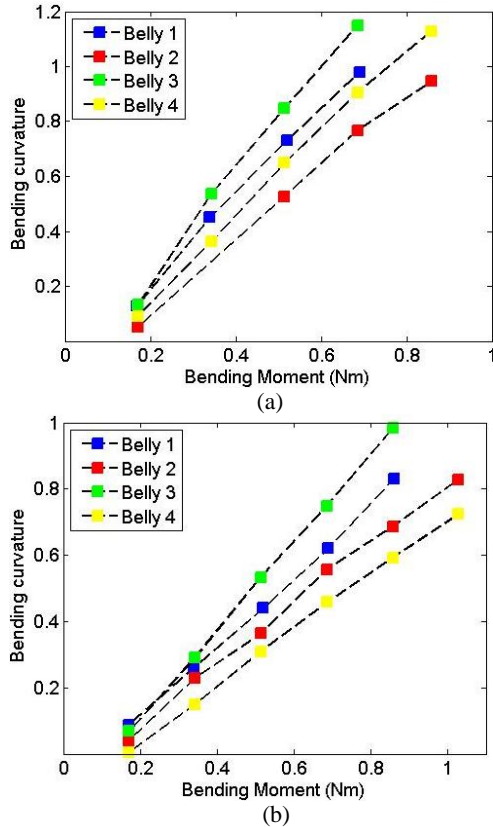


Figure 11. Results of top plates measured by the bending stiffness testing device for mode 2: (a) the upper end; (b) the lower end.

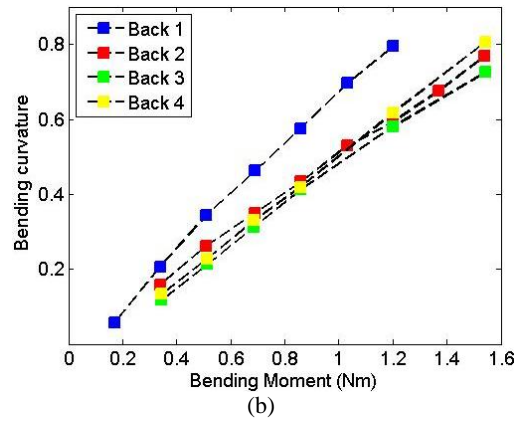
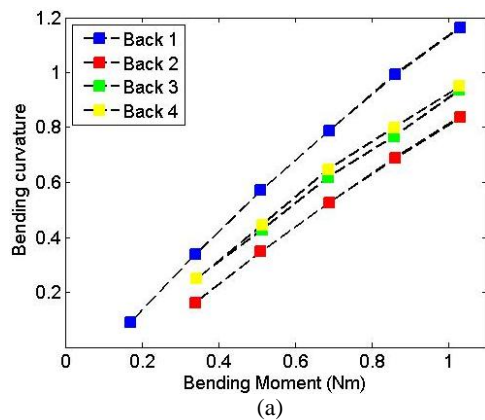


Figure 12. Results of back plates measured by the bending stiffness testing device for mode 2: (a) the upper end; (b) the lower end.

Mode 2 upper	1	2	3	4
Belly	0.9326	1.0566	0.7500	1.1550
Back	1.3987	2.0069	1.9537	1.7822

Table 2. Upper end effective bending stiffness of mode 2.

Mode 2 lower	1	2	3	4
Belly	0.6042	0.7549	0.5073	0.6287
Back	0.7917	1.0125	0.9995	0.9715

Table 3. Lower end effective bending stiffness of mode 2.

Figure 13 and 14 show results obtained from bending experiments of mode 5. Again, the measured results of bending curvature are approximately proportional to the measured bending moment. Thus the testing device for mode 5 is capable of its intended performance. After linear fitting, the stiffness characteristics along the grain of the plates given in Table 4 can also be obtained by equation (7).

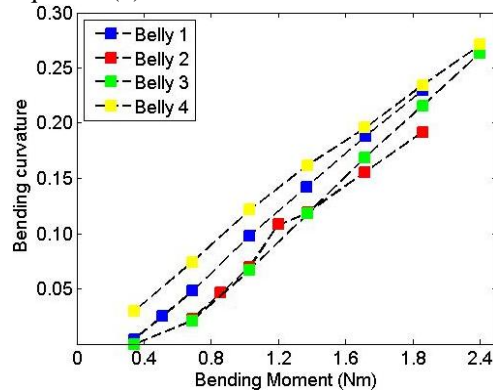


Figure 13. Results of top plates measured by the bending stiffness testing device for mode 5.

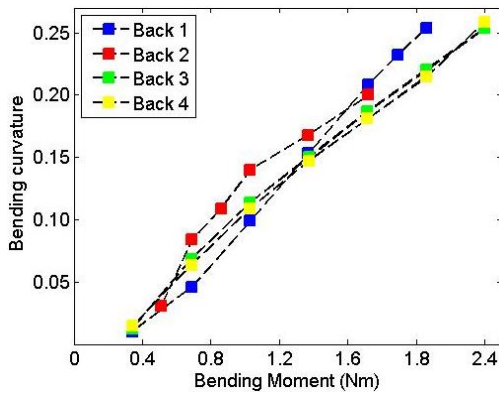


Figure 14. Results of back plates measured by the bending stiffness testing device for mode 5.

Mode 5	1	2	3	4
Belly	1.8986	2.4105	1.6195	2.0648
Back	3.2823	5.2338	4.8098	3.6027

Table 4. Effective bending stiffness of mode 5.

6. CONCLUSIONS AND DISCUSSION

In this paper three new testing devices have been constructed to test effective stiffness characteristics of free plates in a human-like way. Their performance capabilities are examined by a series of experiments. With the mechanical features described in Section 3, and the additional features mentioned above, these devices allow the same gestures used by violinmakers to be executed on free violin plates and provide recordable and quantitative data. The experimental results gave a clear indication of the fact that applied moments and resulting angles or curvatures have a strong positive correlation. Thus the reliability of measurements carried out with these new devices is also ensured. Overall the performances of three devices are shown to be easily adequate for their intended purposes.

It can be seen that the devices and experimental methods presented here are a good first step. The accuracy and ease of use suggest the possibility of further study: based on the effective stiffness of nearly finished free plates, research might shed light on the influence from the distribution of thickness and arching of plates on the stiffness and eigenfrequencies of plates; in particular it will aim to help discover especially sensitive places where wood should be scraped.

Future development of these devices will examine the following aspects: the accomplishment of testing and calibration standards, and use of a digital readout system.

To explore what might be learned by mining data obtained by these devices two further steps will be taken. A wide range of measurements will be carried out in collaboration with makers in the Division of Violin Making at the Central Conservatory of Music in Beijing and Shanghai Conservatory of Music. As a consequence, luthiers might significantly benefit from quantifying the specific stiffness characteristics of several modes of lower frequency in their empirical work. The feedback

data from luthiers will be subsequently combined with input admittance measurements and finite element analysis to analyze how eigenfrequencies are related to the dimensions, mass and stiffness of plates. Thus applications of the devices presented in this paper can range from the practical tuning process to theory validation studies.

Acknowledgments

Our sincere thanks are due to Beijing Zhigao Violin Making Studio and Central Conservatory of Music in China for loans of instruments. The author would like to thank Professor Jim Woodhouse for helpful discussions.

7. REFERENCES

- [1] F. Savart, "Rapport sur un mémoire relatif à la construction des instruments à cordes et à archet," *Ann. Chim. Phys.* vol. 2, pp. 225, 1819.
- [2] H. Backhaus, "Über die Schwingungsformen von Geigenkörpern," *Z. Phys.* vol. 62, pp. 143-166, 1930.
- [3] H. Backhaus, "Über die Schwingungsformen von Geigenkörpern, II," *Z. Phys.* vol. 72, pp. 218-255, 1931.
- [4] C.M. Hutchins, "A history of violin research," *J. Acoust. Soc. Amer.* vol. 73, pp. 1423-1425, 1983.
- [5] C.M. Hutchins, A.S. Hopping, and F.A. Saunders, "Subharmonics and Plate Tap Tones in Violin Acoustics," *J. Acoust. Soc. Amer.* vol 32, pp. 1443-1449, 1960.
- [6] C.M. Hutchins, "The acoustics of violin plates," *Scientific American* vol. 245, no.4, pp. 170-181, 1981.
- [7] K.A. Stetson, "New Design for laser-speckle interferometer," *Optical and Laser Technol.* vol. 2, pp. 179-181, 1970.
- [8] H.F. Meinel, "Über die Beziehungen zwischen Holzdicke Schwingungsform, Körperamplitude und klang eines Geigenkörpers," *Elekt. Nachrichten-Technik.* vol 14, pp.119-134, 1937.
- [9] O.E. Rodgers, "Initial results on finite element analysis of violin backs," *Catgut Acoust. Soc. Journal* vol.1, no. 46, pp.19-23, 1986.
- [10] M. Schleske, "The influence of changes in the distribution of mass and stiffness of violin-plates on eigenfrequencies and mode-shapes of the violin," *Proc. Stockholm Music Acoust Conference 1993*, pp. 420-27, 1993.
- [11] E.V. Jansson, "On the tuning of free plates," *Catgut Acoust. Soc. Journal* vol. 2, pp. 27-32, 1989.

Enhanced simulation of the bowed cello string

Hossein Mansour

CIRMMT, CAML

Schulich School of Music

McGill University

Montreal, Québec H3A 1E3, Canada

hosse-

in.mansour@mail.mcgill.ca

Jim Woodhouse

Cambridge University Engineering

Department

jw12@cam.ac.uk

Gary P. Scavone

CIRMMT, CAML

Schulich School of Music

McGill University

Montreal, Québec H3A 1E3, Canada

gary@music.mcgill.ca

ABSTRACT

A detailed time-domain simulation is implemented to model the bowed cello string. Building on earlier simulation models [1-3], several new features have been added to make the model more realistic. In particular, a large number of body modes, both transverse polarizations of the string motion, the longitudinal vibrations of the bow hair and the effect of the sympathetic strings are included. These additional features can be turned on and off in the model to evaluate their relative importance. To the best of our knowledge this is the first time that the second polarization of the string and the effect of the sympathetic strings have been included in a bowed-string simulation. The compliance of the bow-hair was accounted for in previous studies but without considering its own vibration properties [4]. Different features of the model are turned on and the classic Schelleng minimum bow-force [5] is calculated for combinations of bow-bridge distance and different notes being played on the string. The main finding is that all features reduce the minimum bow-force to some extent. This reduction is almost frequency independent for the case of the second polarization and the longitudinal bow-hair vibration, but clearly frequency-dependent for the sympathetic strings case.

1. INTRODUCTION

The time-domain simulation of the bowed string has been the subject of many studies in recent years and as a result, many features of bowed strings have been explained qualitatively and to a certain extent quantitatively [6, 7]. Different mechanical details of the bowed string have previously been investigated in the hope of making the numerical models match the experimental measurements over a wide range of bowing gestures of musical interest. Among those details added to the model are torsional vibration of the string [8], bending stiffness of the string [9], effect of a close body resonance resulting in a “wolf note” [1], longitudinal bow-hair compliance [4], and the effect of the bow’s finite width [2]. This study is aimed at adding more such details to the model. Specifically, modal properties of the body, dual-polarization of the string, longitudinal vibration of the bow-hair, and the effect of the sympathetic strings are taken into account.

2. MATERIAL AND METHOD

2.1 Basic model of the bowed string

The simulation results in this paper are based on modeling the best-understood musical string, a cello D string (Thomastik ‘Dominant’). For this particular string a reasonably complete set of calibration data is available, covering transverse vibration frequencies and damping factors, torsional frequencies and damping factors, and bending stiffness. In short, the characteristic impedance of the string in the transverse direction is $Z_D = 0.55$ Ns/m and a constant Q -factor of 500 is assumed and implemented in the reflection functions using the method proposed in [8], which only reflects the intrinsic damping of the string. The reflection functions from the bridge side and the finger side are modified as suggested in [9], to take into account wave dispersion due to the bending stiffness of the strings. The string tension, bending stiffness, and cutoff frequency [9] were respectively 111 N, 3×10^{-4} Nm², and 3×10^4 Hz. Torsional vibrations are also taken into account and parameters were extracted from [8]. The torsional wave has the characteristic impedance of 1.8 Ns/m, constant Q -factor of 45, and propagation speed of 1060 m/s. The position of the bowed point on the string is denoted by the dimensionless quantity β , which is the fractional distance of the bow from the bridge (i.e. $\beta = \frac{\text{bow-bridge distance}}{\text{string length}}$). The time-step used for all simulations is 5×10^{-6} second.

All simulations of this study are done using the old velocity dependent friction model (also known as Friedlander’s friction curve [10]) whose parameters are extracted from the constant slipping experiments reported in [11]. Although now known to be wrong in some respects, the friction-curve model has been so far remarkably successful in describing, at least qualitatively, many observed aspects of bowed-string behavior [3]. The physically more accurate visco-plastic friction model is more computationally demanding, and still involves unresolved research questions for its details [12]. Incorporating such a model is a topic for future research.

2.2 Model of the body

Calibrated measurements have been made on the C -string corner of two mid-quality cellos. A miniature hammer (PCB Model 086E80) and LDV (Polytec LDV-100) were used to make the full set of measurements in the bowing plane (i.e. X-X, Y-Y, X-Y, and Y-X where X represents

the bowing direction and Y is the direction normal to that in the bowing plane).

Compensation for the strings is made as described in [13] such that the admittances are adjusted to account for all strings. For a cello, each string has a vibrating length on both sides of the bridge, so that if all of these are damped then the sum should be taken over eight semi-infinite strings. Equation (1) is used to deduce $Y_{\text{body}}(\omega)$ from the measured $Y(\omega)$.

$$\frac{1}{Y(\omega)} = \frac{1}{Y_{\text{body}}(\omega)} + \sum_j Z_j \quad (1)$$

Since both polarizations of the strings are taken into account, each admittance matrix (i.e. $Y_{\text{body}}(\omega)$ and $Y(\omega)$) is a 2x2 ideally symmetric matrix and $\sum_j Z_j$ in Equation (1) can be calculated as:

$$\sum_j Z_j = (\sqrt{\rho_C T_C} + \sqrt{\rho_G T_G} + \sqrt{\rho_D T_D} + \sqrt{\rho_A T_A}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (2)$$

where indices C , G , D , and A represent the different strings of the cello, ρ is the density of the strings with units kg/m and T is the tension of the strings.

As discussed in [13], using Equation (1) will result in the admittances being compensated too generously since the strings cannot be perfectly damped with conventional damping methods. As expected, adding the compensation to the admittances has resulted in sharper peaks with higher Q -values.

Consequently, for Cello #1, which is used the most in the simulations, 89 modes were used to synthesize the admittance up to the frequency of 5 kHz; among these, 8 were below 100 Hz. Those low frequency modes were included in the curve fitting process to keep their residual effect, but were later removed from the simulations as any relative bridge-corpus motion is unlikely over that low frequency range. The curve fitting process was performed using ME'Scope, a commercial modal analysis package. Top priority was given to fit the X-X admittance as precisely as possible, the second priority was given to fitting the X-Y (or alternatively Y-X) data, and finally the Y-Y admittance was reconstructed according to the parameters enforced by X-X and X-Y and Equations (3) repeated from [14]. The measured admittances and their reconstructions are compared in Figure 1.

$$Y_{xx}(\omega) = \sum_k \frac{i\omega \cos^2 \theta_k}{m_k(\omega_k^2 - i\omega\omega_k/Q_k - \omega^2)} \quad (3.1)$$

$$Y_{yy}(\omega) = \sum_k \frac{i\omega \sin^2 \theta_k}{m_k(\omega_k^2 - i\omega\omega_k/Q_k - \omega^2)} \quad (3.2)$$

$$Y_{xy}(\omega) = Y_{yx}(\omega) = \sum_k \frac{i\omega \cos \theta_k \sin \theta_k}{m_k(\omega_k^2 - i\omega\omega_k/Q_k - \omega^2)} \quad (3.3)$$

The final result of the reconstruction is four 81-element vectors containing the natural frequencies, Q -factors, modal masses (m), and mode angles (θ) that were fed into the bowed-string model. The same procedure was applied to the measurements made on Cello #2 with a slightly different number of modes. It is noteworthy that θ indicates the angle between the principal direction of vibration for each mode with respect to the bowing direction. This angle is deduced from the relative expression of

each mode in the measured admittances of different directions (calculated from Equations (3.1) to (3.3)).

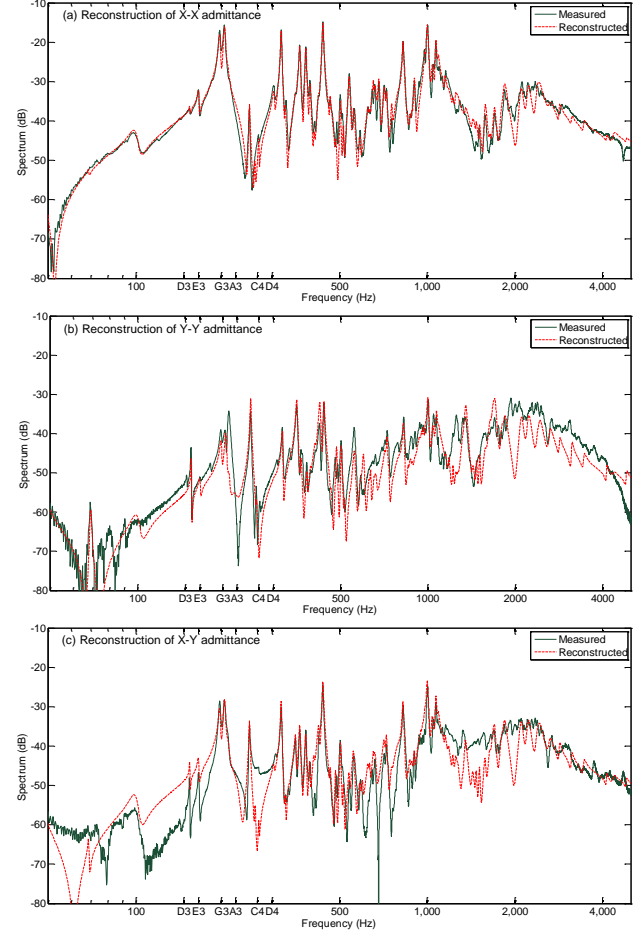


Figure 1: Measured admittances on the C-string side of the cello (solid green line) together with their reconstruction with 81 modes (dashed red line). X represents the bowing direction and Y direction is perpendicular to that. (a) X-X measurement, (b) Y-Y and (c) X-Y (or alternatively Y-X); for easier comparison some musical scales are also shown on X-axis.

The body modes are included in the bowed-string model using an IIR technique described in [14]. The time step used in our simulations allows for 40 time-samples in each period of the highest simulated body mode. Each (k^{th}) mode is modeled as an independent resonator and its damped free oscillation is modified by the value $F_{\text{Excit}}(k) * h/m_k$ in each time step where h is the time step and $F_{\text{Excit}}(k)$ is the instantaneous force applied by the string(s), projected in the principal direction of mode k .

F_{Excit} itself is a function of the incoming wave velocities from the bowed string in the X and Y directions (V_{Xin} and V_{Yin}) as well as the velocity of the bridge notch in the X and Y directions ($V_{XBridge}$ and $V_{YBridge}$) which can be approximated by their value at the previous time step as:

$$V_{XBridge} = \Re \{ \sum_k A_k(\text{Bridge}) \cos(\theta_k) \} \quad (4.1)$$

$$V_{YBridge} = \Re \{ \sum_k A_k(\text{Bridge}) \sin(\theta_k) \}, \quad (4.2)$$

where $A_{(\text{Bridge})}$ is the complex vector of mode velocities. F_{Excit} , the vector of modal forces, can thus be calculated from

$$F_{\text{Excit}} = 2 * Z_D * (V_{\text{Xin}} * \cos \theta + V_{\text{Yin}} * \sin \theta) - 2 * (V_{\text{XBridge}} * \cos \theta + V_{\text{YBridge}} * \sin \theta) * (Z_D + Z_C + Z_G + Z_A), \quad (5)$$

which will reduce to the form of Equation (6) if we only consider a single polarization of the string (i.e. the bowing direction):

$$F_{\text{Excit}} = 2 * Z_D * V_{\text{Xin}} * \cos \theta - 2 * V_{\text{XBridge}} * \cos \theta * (Z_D + Z_C + Z_G + Z_A). \quad (6)$$

In both Equations (5) and (6) the first term corresponds to the reaction force of the bridge resulting in the phase reversal of the reflected velocity wave of the bowed string and the second term allows for the fact that movement of the bridge sends out velocity waves in all strings whether the strings are damped or not and regardless of the amount of incoming velocity wave from the strings. It is noteworthy that this effect was accounted for when the admittance was being measured with the strings damped but was taken out later by admittance compensation. The factor of two for the first term represents the difference of the incoming and outgoing waves and for the second term relates to the fact that a cello has strings extending on both sides of the bridge (as opposed to a guitar for example).

2.3 Second polarization of the string

The movement of the bridge notch is not necessarily in the bowing direction for all body modes. Also, the bow-hair is not rigid enough to suppress all the string vibrations normal to it. As shown in Equation (5), the X and Y polarizations of the bowed-string are coupled together via the bridge. Incoming waves in the X and Y directions add up to excite modes that are not necessarily lined up in either direction. Projected velocities of each mode in the X and Y directions ($V_{\text{XBridge}}(k)$ and $V_{\text{YBridge}}(k)$) add, consequently, to the reflected waves at the bridge.

In this regard, the bow-hair needs to be flexible in its transverse direction to allow for this second polarization of the string motion. A pair of traveling waves is considered to model this bow hair transverse vibration. The bow hair ribbon is assumed to have mass per unit length 0.0077 kg/m, length 0.65 m and total tension 60 N [15]. A Q-factor of 20 is used for the transverse waves as estimated in [16]. This will result in a fundamental of 68 Hz for bow-hair vibrations in the transverse direction. The “bow β ” (distance from the contact point to the frog divided by the full length of the hair ribbon) is arbitrarily chosen to be 0.31 which does not change during the simulation as its dynamics are assumed to be much slower than the dynamics of the string itself.

The transverse vibration of the bow-hair is excited by the normal-to-bow vibrations of the string; hence the bow and the string are coupled at the contact point. The constraint at the contact point is that they share a common velocity and apply the same amount of force to each other in opposite directions. To find the unknown common velocity and the mutual force, the velocity of the string and the bow are first calculated in the absence of the other one. These values, called V_{Yh} and V_{bTh} , represent history of the string velocity in the Y direction and history of the bow-speed in the transverse direction. With simple

math it can be shown that the matched velocity (V_{match}) will be equal to

$$V_{\text{match}} = \frac{V_{\text{Yh}} * Z_D + V_{\text{bTh}} * Z_{\text{bT}}}{Z_D + Z_{\text{bT}}} \quad (7)$$

and the resulting fluctuating force in the contact region (F_{fluc}) will be

$$F_{\text{fluc}} = 2 * Z_D * (V_{\text{match}} - V_{\text{Yh}}). \quad (8)$$

This force is added to the nominal value of the bow-force, supplied by the player, to give the effective bow-force as in Equation (9). Since the bow-force is being dynamically updated for each time step, the friction curve should consequently be re-scaled.

$$F_{\text{eff}} = F_{\text{nom}} + F_{\text{fluc}} \quad (9)$$

2.4 Longitudinal bow-hair vibration

Bow-hair also has some degree of compliance in the longitudinal direction, which is excited by the fluctuating friction force between the bow and the string. The characteristic impedance of the hair ribbon in the longitudinal direction (called Z_{bL}) is about 10 Ns/m [4]; the wave speed in the longitudinal direction 2300 m/s [15]; and the Q-value approximated at 10 [16]. In the presence of bow-hair longitudinal vibrations, the nominal bow velocity will be modulated by the velocity of the contact point on the bow-hair relative to the bow stick. This relative velocity can be found from:

$$V_{\text{bFluc}} = V_{\text{L_intip}} + V_{\text{L_infrog}} + \frac{F_f}{2 * Z_{\text{bL}}} \quad (10)$$

and the effective bow speed can be calculated from

$$V_{\text{bEff}} = V_{\text{b}} - V_{\text{bFluc}}, \quad (11)$$

where F_f is the instantaneous friction force between the bow and the string, V_{b} is the nominal bow speed provided by the player, and $V_{\text{L_intip}}$ and $V_{\text{L_infrog}}$ are the incoming longitudinal velocity waves, from the tip and the frog respectively, arriving at the contact point. It is noteworthy that since the friction curve is a function of bow speed, it should be reconstructed with V_{bEff} instead of V_{b} at each time step.

The effect of the bow-stick modes was also taken into account elsewhere [17]. Briefly, the interaction of the longitudinal and transverse bow-hair vibration with the stick modes can be implemented in a similar fashion to the string’s dual polarization motion with the body modes. The effect of the bow-stick modes was, however, negligible when compared to the effect of the transverse and longitudinal vibrations of the bow-hair itself.

2.5 Sympathetic strings

Under normal playing conditions the bow continuously excites one string and the other three are free to vibrate, their excitation being provided by the moving bridge. This certainly has an effect on the playability of a note, particularly when one or more of the coupled strings have harmonics matching those of the bowed string.

Physical properties of those three strings were extracted from [18] and their vibrations are simulated using the same method as for the bowed string. Only a single polarization is considered for the sympathetic strings; however, their resistance against bridge vibration in the Y direction is also taken into account. The single polarization assumption is made based on the fact that most of the

body modes in the lower range of frequencies result in a rocking motion of the bridge which is approximately in the X direction for all strings. Based on the same assumption, a single body admittance is used for all four strings. Equation (12) shows the modified version of Equation (5) for the case with sympathetic strings:

$$F_{\text{Excit}} = 2 * Z_D * ((V_{\text{Xin}} - V_{\text{XBridge}}) * \cos \theta + (V_{\text{Yin}} - V_{\text{YBridge}}) * \sin \theta) + \sum_{j=C,G,A} 2 * Z_j * (V_{\text{jIn}} * \cos \theta - V_{\text{XBridge}} * \cos \theta - V_{\text{YBridge}} * \sin \theta) \quad (12)$$

The single polarization assumption results in an underestimate of the sympathetic strings' effect.

2.6 Summary of the model

The model of the bowed string allows for up to 6 different types of motion, each of which is modeled with a pair of superposed traveling waves: 1) vibration of the bowed string in the bowing direction; 2) vibration of the bowed string in the direction perpendicular to the bow; 3) torsional vibration of the string; 4) transverse vibration of the bow-hair; 5) longitudinal vibration of the bow-hair, and 6) single polarization vibrations of the three sympathetic strings. It also uses 81 independent resonators to simulate body modes. The code adjusts the effective length of the bowed string based on the chosen fundamental frequency.

3. RESULTS AND DISCUSSION

Different combinations of the above mentioned degrees of freedom are allowed and their relative importance can be evaluated. Schelleng maps are calculated for each case to see the effect on the playability of the cello. Schelleng calculated formulae for the maximum and minimum bow-forces between which the Helmholtz motion of the string is possible, and plotted the results in the plane of force against bowing position on a log-log scale [19]. The following simulations are made on the first 13 semi-tones played on the D string, giving the possibility to study the note-by-note variations of the instrument. Each map studied in this paper is composed of a grid of 7800 time-domain simulations, each 1 s long. The grid is formed by different combinations of the played note, β , and bow-force (20 values for β and 30 values of bow force, all exponentially spaced). β ranged from 0.02 to 0.22 and force ranged from 0.0005 N to 2.5 N. The bow velocity was chosen to be 0.05 m/s in all cases, although varying that value might form an interesting study in its own right [19]. Motion of the string in the bowing direction was always initialized with a proper sawtooth wave. In all cases torsional vibration of the string and its bending stiffness were taken into account. Each time domain waveform is automatically classified into Helmholtz, double slip, constant slipping, ALF [20] and Raman higher types [21] using the method proposed in [3] and modified in [12]. A separate function automatically calculates the minimum bow-force for each combination of played note and β . The result is a surface similar to the one shown in Figure 2. Slices of this surface in the fre-

quency and β directions are comparable to the lower limit of the classic Schelleng diagram and to a sampled version of the minimum bow-force proposed in [22]. Another interesting possibility with this 3-D Schelleng diagram is to find the geometric mean of the minimum bow-force when averaged over frequency or over β to see more global effects (see for example Figure 3).

3.1 The body effect

The minimum bow-force surface for Cello #1 is plotted in Figure 2. As was expected from the admittance, a mountain range is observed from F# to G#. To study the influence of body modes on minimum bow force the same plot is calculated for our second studied cello (called Cello #2). The results averaged over frequency and over β are compared for the two instruments in Figures 3a and 3b respectively. The average over β (Figure 3a) shows that these two instruments behave slightly differently on different notes, specially in the wolf range, but the average over frequencies (Figure 3b) shows that these differences are mostly local and neither of the two instruments is globally more “playable” in this sense than the other. The β averaging is taken over the first 13 exponentially spaced β values ranging from $\beta = 0.02$ to $\beta = 0.09$; beyond $\beta = 0.1$ or so, the minimum bow-force is mostly driven by string properties and Helmholtz usually breaks down to constant slipping with no sign of double slip. For the same reason, Figure 3b is only plotted up to $\beta=0.117$.

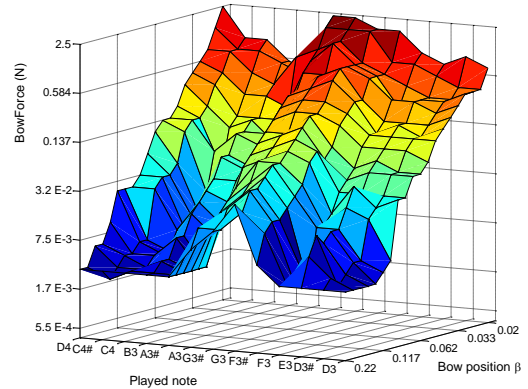


Figure 2: Minimum bow-force as a function of played note and β for Cello #1. Only the body is included in the model and all other options are turned off.

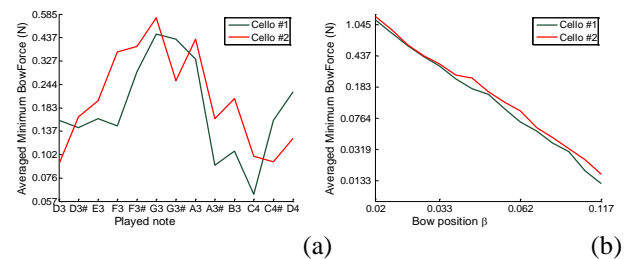


Figure 3: Minimum bow-force for Cellos #1 and #2 averaged over β (a), and over thirteen semitones of D3 to D4 (b). The instrument is being played with a rigid bow, and the other strings are damped

It is expected to see a stronger effect of body for the *C* string due to its closer impedance match to the body. However, no effort was made to simulate such effect in this study as the full mechanical properties of the *C* string were not available as much as they were for the *D* string.

3.2 Dual string polarization

The effect of the second string polarization together with compliant bow-hair in the transverse direction is studied in this section. Adding the second polarization can potentially affect the playability of the instrument in two ways: a) by modulating the effective bow-force according to Equation (9) and b) by absorbing energy from the string and damping the transients/disturbances that can potentially trigger a second, premature slip. The latter effect can only reduce the effective bow-force while the former can either increase or decrease it depending on the relative phase of bow-hair and stick-slip motions.

It can be seen from Figure 4 that the second string polarization has generally reduced the effective bow-force. The effect is even more visible for larger β s in Figure 4b. A possible explanation for this β dependency is that the amplitude of string vibration in both polarizations gets higher when the bow moves farther from the bridge, thus a stronger string-bowhair coupling can occur.

An interesting observation in Figure 4a is that the bow-force reduction is almost independent of the note being played. The fundamental frequency of the bow-hair in the transverse direction is around 68 Hz, so one may expect a further minimum bow-force reduction in integer multiples of this frequency (i.e. *G3#*, *C4#*). This did not happen, for two possible reasons: a) the bow- β used in this simulation is close to 1/3 and thus the third mode of the bow-hair can not contribute much; b) the bow-hair is heavily damped ($Q=20$), so fairly broad peaks are expected in the impedance of the bow-hair, which cannot create a strong frequency-dependent behavior (this agrees with the experience of the players that the playability of a particular instrument is not strongly dependent on bow tension or bow- β). Of course, a stronger frequency dependency may be expected if a strong body resonance with a mode angle close to 45° exists close to one of the played notes.

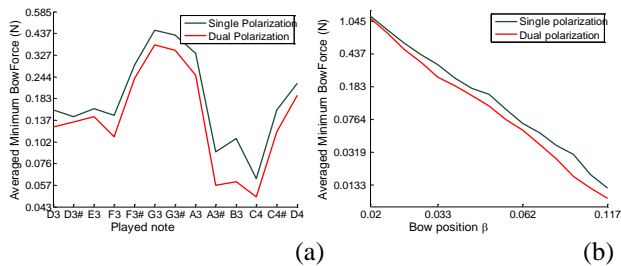


Figure 4: Effect of the second string polarization on the minimum bow-force of Cello #1, averaged over β (a), and over thirteen semitones of *D3* to *D4* (b). In both plots the green line is for the single polarization (same as Green in Figure 3.a and 3.b) and the red is for dual string polarizations

3.3 Longitudinal bow-hair vibration

The effect of the longitudinal bow-hair vibration on the minimum bow-force is plotted in Figures 5a and 5b. As can be seen, the only difference is a slight reduction, which is probably due to a small energy absorption from the string. The insignificance of this effect can probably be traced to the large impedance mismatch between the bow-hair in the longitudinal direction (around 10 Ns/m) and the string in the transverse direction (0.55 Ns/m). A much stronger effect was observed when the bow-hair characteristic impedance was reduced to 4 Ns/m. Also a stronger effect might be expected when studying faster transients of the string such as the ones studied in [23].

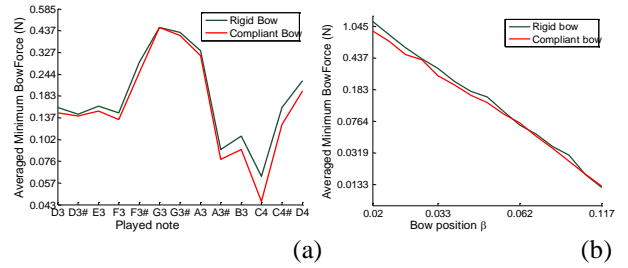


Figure 5: Effect of the longitudinal bow-hair vibration on the minimum bow-force of Cello #1, averaged over β (a), and over thirteen semitones of *D3* to *D4* (b). In both plots the green line is for a rigid bow (same as Green in Figure 3.a and 3.b) and the red is for a bow compliant in longitudinal direction

Another interesting observation from the grid of classified motions is more frequent occurrence of ALF notes when longitudinal bow-hair vibration was allowed. For the simulations over the *D3* to *D4* octave, 102 ALF occurrences were observed for a compliant bow as compared to 62 for a rigid bow. This point was emphasized by Mari Kimura in [24] “*The first secret is maintaining loose bow-hair.... You don’t want a lot of tension... You need enough elasticity on the bow-hair that you can really grab the string.*”

3.4 The sympathetic strings

The effect of three sympathetic strings is systematically studied in this section. An octave range is played on the *D* string of Cello #1 and the resulted minimum bow-force is compared for the cases of damped and undamped sympathetic strings (with a regular *C-G-D-A* tuning). The exact effect is somehow complicated (see [25] section 5) and needs further analysis but the general expectation would be a reduction in the minimum bow-force when one or more of the coupled strings are sympathetically tuned with the played note. In this regard, four notes in the octave *D3* to *D4* are expected to be affected:

- *G3* which is an octave above *G2* (twice the frequency)
- *A3* that is unison with the *A3* coupled string
- *C4* which is two octaves above *C2* (four times the frequency)
- *D4* which is an octave and a fifth above *G2* (three times the frequency)

This effect is clearly visible in Figure 6a. A consistent “deep canal” can be seen on $G3$ which has support from two coupled strings and also happens to be around the strongest body resonance represented by a mountain range in Figure 2.

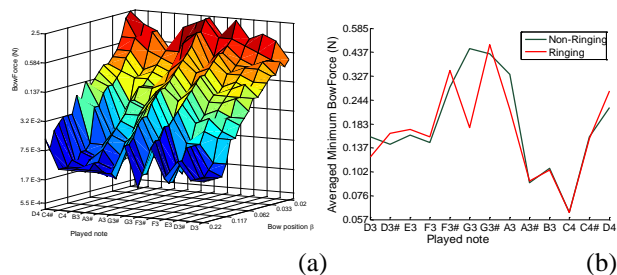


Figure 6: Minimum bow-force as a function of played note and β for Cello #1 similar to Figure 2 but being played when the other three strings were free to vibrate (a); β -averaged minimum bow-force with and without the sympathetic strings (b)

It is more revealing to look at the β -averaged minimum bow-force with and without the sympathetic strings (Figure 6b). A similar effect is observed for $A3$ to a lesser extent; however, $C3$ and $D4$ were not much affected by the sympathetic strings. A possible explanation for the $C3$ case is that its minimum bow-force was already quite low without the sympathetic strings; thus its dynamics are most likely being driven by the intrinsic damping of the string itself rather than the vibrations of the bridge. $D4$'s independency on the sympathetic strings might be associated with the fact that it only gets support from every third partial of $G2$, which are apparently not so effective.

4. CONCLUSIONS

New features have been added to the model of the bowed string. A detailed body model is implemented whose parameters were extracted from calibrated admittance measurements. The model takes into account the angle of body modes with respect to the bowing direction, which transforms the initial excitation of the string in the bowing direction to an excitation of the second polarization. A Schelleng minimum bow-force plot was produced that looks similar to ones found in the literature, with the difference that this diagram was calculated for different notes being played on an instrument. This offered the opportunity to explore the note-by-note variation of minimum bow-force on the same instrument or over different instruments. The second polarization of the string was taken into account by allowing the string to vibrate in both polarizations and the bow-hair to vibrate in its transverse direction. The result was a general reduction in minimum bow-force with no obvious frequency dependence, and with a stronger effect when bowing the string farther from the bridge. Moreover, the longitudinal vibration of the bow-hair was added to the model and its effect in modulating the effective bow-speed was studied. A general reduction in minimum bow-force was observed, similar to the second string polarization case, and the occurrence of ALF notes was found to be much more frequent compared to the rigid bow case. Finally, the ef-

fect of the sympathetically tuned coupled strings was studied. A noticeable reduction in minimum bow-force was observed at $G3$ which was the dominant wolf of the instrument before inclusion of the sympathetic strings. This effect and a minor reduction in $A3$ were justified by harmonic relations between the played notes and the sympathetic strings. The results are not further pursued here as the main focus of this article was to describe the theoretical background and modeling procedure. The results will be further discussed and experimentally validated in future studies.

Acknowledgments

The authors would like to thank the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) for partially support this project and making this collaboration possible. We thank two anonymous reviewers for detailed constructive comments. The first author is financially supported through a Richard H. Tomlinson doctoral fellowship to McGill University.

5. REFERENCES

- [1] M. E. McIntyre and J. Woodhouse, "On the fundamentals of bowed string dynamics," *Acustica*, vol. 43, pp. 93-108, 1979.
- [2] R. Pitteroff and J. Woodhouse, "Mechanics of the contact area between a violin bow and a string. Part II: Simulating the bowed string," *Acta Acustica united with Acustica*, vol. 84, pp. 744-757, 1998.
- [3] J. Woodhouse, "Bowed string simulation using a thermal friction model," *Acta Acustica united with Acustica*, vol. 89, pp. 355-368, 2003.
- [4] R. Pitteroff and J. Woodhouse, "Mechanics of the Contact Area Between a Violin Bow and a String. Part I: Reflection and Transmission Behaviour," *Acta Acustica united with Acustica*, vol. 84, pp. 543-562, 1998.
- [5] J. C. Schelleng, "The bowed string and the player," *The Journal of the Acoustical Society of America*, vol. 53, pp. 26-41, 1973.
- [6] L. Cremer, *The Physics of the Violin*. Cambridge MA: The MIT Press, 1984.
- [7] J. Woodhouse and P. M. Galluzzo, "The bowed string as we know it today," *Acta Acustica united with Acustica*, vol. 90, pp. 579-589, 2004.
- [8] J. Woodhouse and A. R. Loach, "Torsional behaviour of cello strings," *Acta Acustica united with Acustica*, vol. 85, pp. 734-740, 1999.
- [9] J. Woodhouse, "On the playability of violins. Part I: Reflection functions," *Acustica*, vol. 78, pp. 125-136, 1993.
- [10] F. Friedlander, "On the oscillations of the bowed string," in *Proc. Cambridge Philos. Soc.*, 1953, p. 516.
- [11] J. H. Smith and J. Woodhouse, "The tribology of rosin," *Journal of the Mechanics and Physics of Solids*, vol. 48, pp. 1633-1681, 2000.

- [12] P. M. Galluzzo, "On the playability of stringed instruments," Ph. D. thesis, Trinity College, University of Cambridge, Cambridge, UK, 2003.
- [13] J. Woodhouse and R. S. Langley, "Interpreting the Input Admittance of Violins and Guitars," *Acta Acustica united with Acustica*, vol. 98, pp. 611-628, 2012.
- [14] J. Woodhouse, "On the synthesis of guitar plucks," *Acta Acustica united with Acustica*, vol. 90, pp. 928-944, 2004.
- [15] A. Askenfelt, "Observations on the violin bow and the interaction with the string," in *Proceedings of the International Symposium on Musical Acoustics*, 1995.
- [16] C. E. Gough, "Violin bow vibrations," *The Journal of the Acoustical Society of America*, vol. 131, p. 4152, 2012.
- [17] H. Mansour, J. Woodhouse, and G. P. Scavone, "Time-domain simulation of the bowed cello string dual-polarization effect," in *Proceedings of the International Congress on Acoustics*, Montreal, Canada, 2013.
- [18] K. Guettler, "Some typical properties of bowed strings," <http://knutsacoustics.com/files/Typical-string-properties.pdf>, last accessed, 2013.
- [19] E. Schoonderwaldt, K. Guettler, and A. Askenfelt, "An empirical investigation of bow-force limits in the Schelleng diagram," *Acta Acustica united with Acustica*, vol. 94, pp. 604-622, 2008.
- [20] R. J. Hanson, F. W. Halgedahl, and K. Guettler, "Anomalous low-pitched tones from a bowed violin string," *The Journal of the Acoustical Society of America*, vol. 97, pp. 3270-3270, 1995.
- [21] C. V. Raman, "On the mechanical theory of vibrations of bowed strings," *Indian Assoc. Cult. Sci. Bull.*, vol. 15, pp. 243-276, 1918.
- [22] J. Woodhouse, "On the playability of violins. Part II: Minimum bow force and transients," *Acustica*, vol. 78, pp. 137-153, 1993.
- [23] K. Guettler, "On the creation of the Helmholtz motion in bowed strings," *Acta Acustica united with Acustica*, vol. 88, pp. 970-985, 2002.
- [24] J. Reel, "Mari Kimura on Subharmonics," in *Strings*, June 2009.
- [25] C. E. Gough, "The theory of string resonances on musical instruments," *Acustica*, vol. 49, pp. 124-141, 1981.

DIGITAL MODELING OF BRIDGE DRIVING-POINT ADMITTANCES FROM MEASUREMENTS ON VIOLIN-FAMILY INSTRUMENTS

Esteban Maestre

CCRMA – Stanford University
MTG – Universitat Pompeu Fabra
esteban@ccrma.stanford.edu

Gary P. Scavone

CAML/CIRMMT – McGill University
gary@music.mcgill.ca

Julius O. Smith III

CCRMA – Stanford University
jos@ccrma.stanford.edu

ABSTRACT

We present a methodology for digital modeling of D -dimensional driving-point bridge admittances from vibration measurements on instruments of the violin family. Our study, centered around the two-dimensional case for violin, viola, and cello, is based on using the modal framework to construct an admittance formulation providing physically meaningful and effective control over model parameters. In a first stage, mode frequencies and bandwidths are estimated in the frequency domain via solving a non-convex, constrained optimization problem. Then, mode amplitudes are estimated via semidefinite programming while enforcing passivity. We obtain accurate, low-order digital admittance models suited for real-time sound synthesis via physical models.

1. INTRODUCTION

String instruments, such as in the violin family, radiate sound indirectly: energy from a narrow vibrating string is transferred to a radiation-efficient body of larger surface area. To a large extent, sound radiation is produced due to the transverse velocity of the instrument body surfaces (e.g., the front or back plates), and such surface motion is transferred to the body through the force that the string exerts on the instrument's bridge. The way in which the input force at the bridge is related to the transverse velocities of the body surfaces depends on very complex mechanical interactions among the bridge itself, the sound post, the front and back plates, the air inside the body cavity, the neck, etc. Because of the importance of the bridge in mechanically coupling the strings and the body, the relation between applied force and induced velocity at the bridge has been an object of study for over forty years [1].

In the context of sound synthesis, we are interested in constructing efficient physical models of violin-family instruments. We aim to design digital filters that accurately represent the string termination as observed from the string-bridge interaction of real instruments. We model transverse string motion by means of two-dimensional digital waveguides [2] with orthogonal internal coupling.

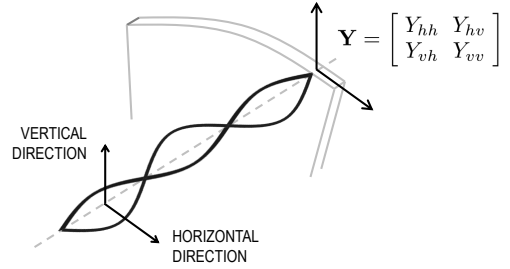


Figure 1. Two-dimensional bridge driving-point admittance.

The admittance is a physical measure used to map applied force to induced motion in a mechanical structure. In the frequency domain, the velocity vector $\mathbf{V}(\omega)$ and force vector $\mathbf{F}(\omega)$ at a position of the structure are related via the *driving-point* admittance matrix function $\mathbf{Y}(\omega)$ by $\mathbf{V}(\omega) = \mathbf{Y}(\omega)\mathbf{F}(\omega)$. In order to emulate horizontal and vertical wave reflection and transmission at the bridge, we need to construct a digital representation of matrix $\mathbf{Y}(\omega)$ in two dimensions, as represented in Figure 1, where sub-indexes indicate string (horizontal and vertical) polarizations, and $Y_{hv} = Y_{vh}$ (symmetric admittance). By taking measurements \mathbf{Y} from real instruments, one can pose this as a system identification problem where a parametric model $\hat{\mathbf{Y}}$ is tuned so that an error measure $\varepsilon(\mathbf{Y}, \hat{\mathbf{Y}})$ is minimized.

Leaving aside approaches based on convolution with measured impulse responses, a first comprehensive work on efficient digital modeling of violin bridge admittances for sound synthesis was by Smith [3], where he proposed and evaluated several techniques for automatic design of common-denominator IIR filter parameters from admittance measurements, making real-time violin synthesis an affordable task. However, while efficiency and accuracy can be well accomplished (also when applied to other string instruments [4]), positive-realness (passivity) [2] cannot be easily guaranteed with common-denominator IIR schemes, leading to instability problems when used to build string terminations. In that regard, the modal framework [5] offers a twofold advantage: (i) admittance can be represented through a physically meaningful formulation, and (ii) positive-realness can be guaranteed.

The modal framework has been used extensively to study the mechanical properties of violins and other string instruments [1], but only recently was applied to synthesize positive-real admittances by fitting model parameters to

Copyright: ©2013 Esteban Maestre et al.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

measurements. In a recent paper [6], Bank and Karjalainen construct a positive-real driving-point admittance model of a guitar bridge by combining all-pole modeling and the modal formulation: they first tune parameters of a common denominator IIR filter from measurement data, and use the roots of the resulting denominator as a basis for modal synthesis, so that positive-realness can be enforced. We tackle a similar problem for the case of violin-family instruments, but using the modal formulation throughout the complete fitting process.

The basic principle of the modal framework is the assumption that a vibrating structure can be modeled by a set of resonant elements satisfying the equation of motion of a damped mass-spring oscillator, each representing a natural mode of vibration of the system. Assuming linearity, the individual responses from the resonant elements (modes) to a given excitation can be summed to obtain the response of the system [7]. In theory, a mechanical structure presents infinite modes of vibration, and experimental modal analysis techniques allow to find a finite subset of (prominent) modes that best describe the vibrational properties as observed from real measurements. In general, admittance analysis via the modal framework begins from velocity measurements taken after excitation of the structure with a given force impulse function.

As introduced in [6], a useful set of structurally passive D -dimensional driving-point admittance matrices can be expressed in the digital domain as

$$\hat{\mathbf{Y}}(z) = \sum_{m=1}^M H_m(z) \mathbf{R}_m, \quad (1)$$

where \mathbf{R}_m is a $D \times D$ positive semidefinite (nonnegative definite) matrix, and each scalar modal response

$$H_m(z) = \frac{1 - z^{-2}}{(1 - p_m z^{-1})(1 - p_m^* z^{-1})}$$

is a second-order resonator determined by a pair of complex conjugate poles p_m and p_m^* [5, 6]. The numerator $1 - z^{-2}$ is the bilinear-transform image of s -plane zeros at dc and infinity, respectively, arising under the “proportional damping” assumption [5, 6]. It can be checked that $H_m(z)$ is positive real for all $|p_m| < 1$ (stable poles). In Section 3 below, we will estimate p_m in terms of the natural frequency ω_m (rad/s) and the half-power bandwidth B_m (Hz) of the m -th resonator, which are related to the z -domain pole p_m respectively by $\omega_m = \angle p_m / T_s$ and $B_m = \log|p_m| / \pi T_s$, where $\angle p_m$ and $|p_m|$ are the angle and radius of the pole p_m in the z -plane, and T_s is the sampling period [2].

Since the admittance model $\hat{\mathbf{Y}}(z)$ is positive real (passive) whenever the gain matrices \mathbf{R}_m are positive semidefinite, the passive bridge-modeling problem can be posed as finding poles p_m and positive-semidefinite gain matrices \mathbf{R}_m such that some error measure $\varepsilon(\mathbf{Y}, \hat{\mathbf{Y}})$ is minimized.

In [6], poles from an all-pole IIR fit are used as the modal basis to estimate \mathbf{R}_m . Once the common-denominator IIR filter has been estimated from measurement data, they find matrices \mathbf{R}_m as follows: First, they independently solve three one-dimensional linear projection problems, each corresponding to an entry in the upper triangle of matrix \mathbf{Y} .

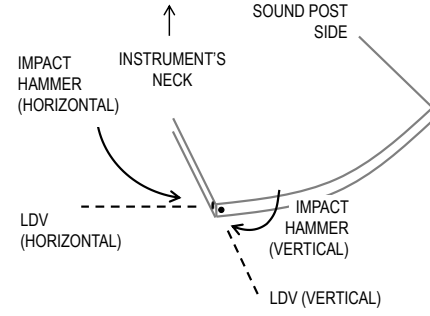


Figure 2. Illustration of the two-dimensional measurement.

This leads to three length- M modal gain vectors. Then, since simply rearranging such gain vectors as a set of M independent 2×2 symmetric gain matrices (matrices \mathbf{R}_m of Equation (1)) does not enforce passivity (all of the \mathbf{R}_m need to be positive semidefinite), the authors ensure passivity by computing the spectral decomposition of each \mathbf{R}_m , and recompose each matrix after discarding any negative eigenvalues and corresponding eigenvectors.

In this work, we deal with violin, viola, and cello bridge driving-point admittance measurements obtained from deconvolution of bridge force and velocity signals, acquired by impact excitation and motion measurement via a calibrated hammer and a commercial vibrometer (Section 2). Based on the modal formulation, we estimate modal parameters in the frequency domain via spectral peak processing and optimization of mode natural frequencies and bandwidths (Sections 3, 4, and 5). Then, we use semidefinite programming to obtain modal gain matrices by solving a matrix-form, convex problem while enforcing passivity via a non-linear, semidefinite constraint (Section 5).

2. MEASUREMENTS

We carried out zero-load bridge input admittance measurements on three decent quality instruments (violin, viola, and cello) from the Schulich School of Music at McGill University. The instruments were held in a vertical position by means of a metallic structure constructed from chemistry stands. Clamps covered by packaging foam were used to rigidly hold all three instruments from the fingerboard near the neck. While the bottom part of the body of both the violin and the viola rested firmly on a piece of packaging foam impeding their free motion during the measurements, the cello rested on its extended endpin. In order to lower the characteristic frequencies of the modes of the holding structure, sandbags were conveniently placed at different locations on the chemistry stands. Rubber bands were used to damp the strings on both sides of the bridge. Figure 3 shows a detail of the measurement setup.

Measurements of force and velocity were performed using a calibrated PCB Piezotronics 086E80 miniature impact hammer and a Polytec LDV-100 Laser Doppler Vibrometer (LDV), both connected to a National Instruments USB-4431 signal acquisition board. The location and orientation of the impact and the LDV beam are illustrated in Figure 2. Both the hammer and the LDV were carefully oriented so



Figure 3. Detail of the violin measurement setup.

that impacts and measurements were as perpendicular as possible to the surface of the edge of the bridge while not interfering with each other. Time-domain signals of force and perpendicular velocity were collected, delay-compensated, and stored before computing the admittance by deconvolution. For each of the three admittance matrix entries Y_{hh} , Y_{vv} , and Y_{hv} , several measurements were collected and averaged in order to use coherence as a means for selecting the most consistent set.

Plots in Figure 4 show the frequency responses of admittance measurements Y_{hh} , Y_{vv} , and Y_{hv} performed on the violin, viola, and cello. From the responses, it is possible to make a few observations. In the region going from 100 Hz to 1 kHz approx., the characteristic modes of violin-family instruments (as extensively studied in the literature [1]) clearly appear at expected frequencies, showing moderately low overlap. Well below 100 Hz, prominent peaks appear in all measurements. Attending to the literature and previous works on normal mode analysis and identification [1], no normal modes are expected to appear at such low frequencies, leaving us with the convincing possibility that these peaks must correspond to modes of the holding structure. This was confirmed after numerous measurement trials in which the configuration of the holding structure and its position in the room were altered. Above 1 kHz, higher

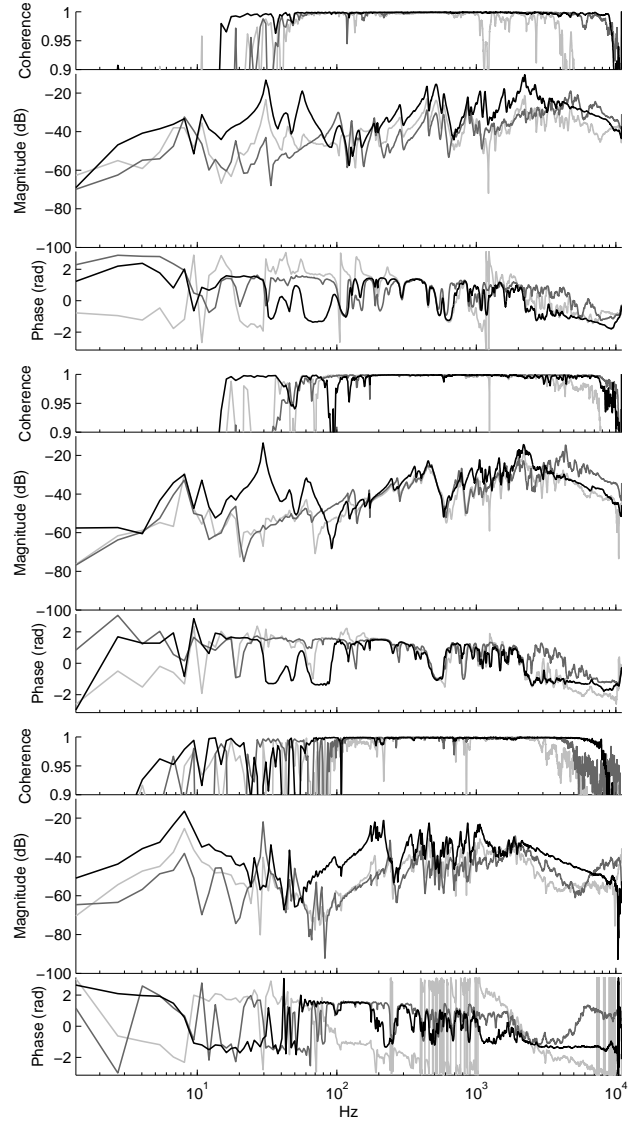


Figure 4. Admittance measurements and observed coherences. From top to bottom: coherence, magnitude, and phase of violin, viola and cello. In each plot: $Y_{hh}(f)$ (black), $Y_{vv}(f)$ (dark gray), and $Y_{hv}(f)$ (light gray) as measured at a sampling frequency of 22.05 kHz.

mode overlap leads to a broad peak (the so-called *bridge hill*), particularly prominent in the Y_{hh} measurements. It can also be observed that in the Y_{vv} responses, a second broad peak appears at a higher frequency region. Regarding phase, measurements corresponding to diagonal terms Y_{hh} and Y_{vv} present a response lying between $-\pi/2$ and $\pi/2$ (corresponding to positive-real functions), as opposed to off-diagonal terms Y_{hv} .

3. MODELING STRATEGY

Departing from admittance measurements in digital form and the M -th order modal decomposition described in Equation (1), our problem can be posed as the minimization

$$\begin{aligned} & \underset{\omega, \mathbf{B}, \mathbf{R}}{\text{minimize}} && \varepsilon(\mathbf{Y}, \hat{\mathbf{Y}}) \\ & \text{subject to} && \mathbf{C}, \end{aligned} \quad (2)$$

where $\omega = \{\omega_1, \dots, \omega_M\}$ are the modal natural frequencies, $\mathbf{B} = \{B_1, \dots, B_M\}$ are the modal bandwidths, $\mathbf{R} = \{\mathbf{R}_1, \dots, \mathbf{R}_M\}$ are the gain matrices, $\varepsilon(\mathbf{Y}, \hat{\mathbf{Y}})$ is the error between the measured admittance matrix \mathbf{Y} and the admittance model $\hat{\mathbf{Y}}$, and \mathbf{C} is a set of constraints. This problem, for which analytical solution is not available, can be solved by means of gradient descent methods that make use of local (quadratic) approximations of the error function $\varepsilon(\mathbf{Y}, \hat{\mathbf{Y}})$. Given the way our fitting problem is posed, five main issues need to be taken into consideration:

- **Design parameters.** From observation of each diagonal entry Y_{hh} and Y_{vv} of the measured admittance matrices (see Section 2), and by contrasting with relevant literature on modal analysis of violin family instruments [1], we make two assumptions. First, at low frequencies (approximately between 100 Hz and 1 kHz) modes of interest present relatively low overlap and can be identified and modeled individually. Second, at higher frequencies (above 1 kHz) high mode overlap leads to a broad peak (bridge hill) that can be modeled by a single, highly-damped resonance. This leaves us with three design parameters: frequencies ω_{\min} and ω_{\max} , and number of low-frequency modes M . Between frequencies ω_{\min} and ω_{\max} , M modes are identified and modeled individually, while above ω_{\max} the bridge hill is modeled by a single mode. When studying the two-dimensional case, we found that using two high-frequency modes provides a better basis for modeling the bridge hill (see Section 2 and Section 5.1).

- **Initial estimation.** Because the minimization problem, Equation (2) is not convex, it is very important to choose a starting point that is close enough to the global minimum. Therefore, it is crucial to carry out an initial estimation of mode parameters prior to optimization. The method, to be described in Section 4.1, is based on peak (resonance) picking from the magnitude spectrum, and a graphical estimation of each mode frequency and bandwidth.

- **Constraint definition.** Defining constraints on the parameters of the problem is motivated by two reasons: feasibility and convergence. First, a number of feasibility constraints are needed to obtain a realizable solution. Second, in order to ensure that the optimization algorithm will not jump into regions of the parameter space where it can get stuck in local minima, additional constraints need to be defined so that candidate solutions stay within the region of convergence.

- **Error computation.** Since there is no analytical expression for the gradient of this error minimization, it needs to be estimated from computing the error in different directions around a point in the parameter space. Therefore, we need to choose a convenient method for computing $\varepsilon(\mathbf{Y}, \hat{\mathbf{Y}})$ at any point in the parameter space.

- **Estimation of gains.** Once the M modal frequencies and bandwidths are optimized, it is necessary to perform an estimation of matrices \mathbf{R} of Equation (2) to complete the model of Equation (1).

4. ONE-DIMENSIONAL MODELING

The one-dimensional procedure presented here can be used both for Y_{hh} and Y_{vv} . First, once design parameters M , f_{\min} , and f_{\max} have been set, individual mode resonances

are identified from the magnitude spectrum through a peak picking iterative procedure. Then, an initial estimation of mode parameters (frequencies and bandwidths) is obtained via the half-power method. In a final step, mode parameters are tuned via numerical optimization.

4.1 Initial estimation

4.1.1 Peak selection

Peak selection in the low-frequency region is carried out through an automatic procedure that iteratively rates and sorts spectral peaks by attending to a salience descriptor. The high-frequency bridge hill resonance center frequency is selected via smoothing the magnitude spectrum.

4.1.2 Estimation of frequencies and bandwidths

For estimating modal frequencies, three magnitude samples (respectively corresponding to the corresponding maximum and its adjacent samples) are used to perform parabolic interpolation. For estimating bandwidths, the *half-power* rule [2] is applied using a linear approximation.

4.2 Error computation

For optimization routines to successfully approximate error derivatives, it is necessary to supply a procedure to evaluate the error function as a function of the model parameters, namely a vector \mathbf{x} . In our case, parameters are (see Equation 2) mode frequencies ω and bandwidths \mathbf{B} . Thus \mathbf{x} is constructed by concatenating elements in sets ω and \mathbf{B} , leading to $\mathbf{x} = [\omega \ \mathbf{B}]^T$. We work with frequency-domain representations of the admittance measurement Y and admittance model \hat{Y} . At the k -th iteration, parameter vector is $\mathbf{x}|_k$, and evaluating $\varepsilon(Y, \hat{Y}|_k)$ implies: (i) retrieving mode frequencies and bandwidths from $\mathbf{x}|_k$, (ii) estimating gains as outlined in Section 4.4, (iii) constructing a synthetic admittance $\hat{Y}|_k$ with computed gains, and (iv) computing error $\varepsilon(Y, \hat{Y}|_k)$.

Let vector $\mathbf{y} = [y_1, \dots, y_n, \dots, y_N]^T$ contain N samples of $Y(\omega)$, taken in $0 \leq \omega < \pi$. Analogously, let $\hat{\mathbf{y}}|_k = [\hat{y}_1|_k, \dots, \hat{y}_n|_k, \dots, \hat{y}_N|_k]^T$ contain N samples of $\hat{Y}(\omega)|_k$, constructed from parameter vector $\mathbf{x}|_k$. We compute the error $\varepsilon(Y, \hat{Y}|_k)$ as

$$\varepsilon(Y, \hat{Y}|_k) = \sum_{n=1}^N \left| \log \frac{|y_n|}{|\hat{y}_n|_k} \right|, \quad (3)$$

which can be interpreted as subtracting magnitudes when expressed in the logarithmic scale.

4.3 Constraint definition

Apart from providing a reliable initial point $\mathbf{x}|_0$ (see Section 4.1), we need to define a set of constraints to be respected during the search:

- **Mode sequence order.** A first important constraint to be respected during the search is the sequence order of modes (in ascending characteristic frequency) as they were initially estimated (see Section 4.1).

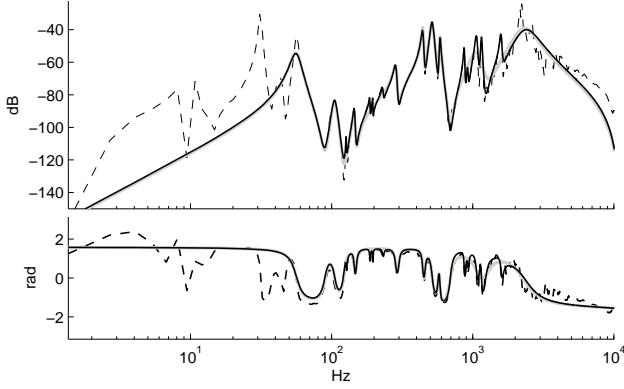


Figure 5. Synthetic admittance \hat{Y}_{hh} modeling example. Magnitude (top) and phase (bottom). Dashed curves: admittance measurement; solid gray curves: $M = 7$, $f_{\min} = 50$ Hz, $f_{\max} = 1300$ Hz; solid black curves: $M = 15$, $f_{\min} = 50$ Hz, $f_{\max} = 1700$ Hz.

- **Low-frequency region.** All of the M low-frequency modes must lie in the frequency region prescribed by design parameters ω_{\min} and ω_{\max} .
- **Bridge hill frequency region.** Analogous to the previous constraint, the characteristic frequency of the bridge hill resonance must be within the limits set by additional design parameters ω_{\min}^b and ω_{\max}^b .
- **Positive bandwidth.** Bandwidths must be positive.
- **Distance from initial estimation.** Assuming the initial estimation $\mathbf{x}|_0$ is close enough to the final solution, one can bound the search space around $\mathbf{x}|_0$.

All constraints are linear in parameters \mathbf{x} , and Equation (2) can be efficiently solved via sequential quadratic programming by means of Matlab's Optimization Toolbox [8].

4.4 Estimation of gains

At any k -th iteration of the optimization procedure, modal gains can be obtained via solving a constrained linear projection problem. By evaluating the individual frequency responses of the $M + 1$ resonators as constructed from parameters in vector $\mathbf{x}|_k$, modal gains are found via solving the one-dimensional version of Equation (1) in the frequency domain while imposing a nonnegative constraint on the gain vector. Results for a violin admittance Y_{hh} are shown in Figure 5.

5. TWO-DIMENSIONAL MODELING

The procedure for the two-dimensional case is based on the fitting procedure for the one-dimensional case. A main assumption is made: the modes of the full system \mathbf{Y} can be estimated by only attending to diagonal measurements Y_{hh} and Y_{vv} . First, diagonal entries Y_{hh} and Y_{vv} of measurement matrix \mathbf{Y} are used separately to obtain two sets of modal frequencies and bandwidths (Section 4). In a second step, the two sets of modes are merged into a common mode structure, which will be used as the basis for the model $\hat{\mathbf{Y}}$. After merging, mode parameters are re-optimized by simultaneously minimizing an error between (diagonal) measure-

ments Y_{hh} and Y_{vv} , and their respective model counterparts \hat{Y}_{hh} and \hat{Y}_{vv} , the latter two constructed from the common mode structure under optimization. Finally, we use semidefinite programming to estimate gain matrices \mathbf{R} for the full matrix model $\hat{\mathbf{Y}}$ by using all three measurements Y_{hh} , Y_{vv} , and Y_{hv} .

5.1 Initial estimation

Providing an initial estimation of mode frequencies and bandwidths consists of two main steps: independent tuning of a one-dimensional model for each entry in the diagonal of matrix \mathbf{Y} , and merging of mode estimations.

5.1.1 Diagonal entries

First, by means of the procedure outlined in Section 4, we carry out two one-dimensional fittings respectively corresponding to self-admittance (diagonal) measurements Y_{hh} and Y_{vv} . Design parameters are shared by both cases, except for the fact that two broad resonances are used for the high frequency region (see Section 2), meaning that two regions $(\omega_{hh,\min}^b, \omega_{hh,\max}^b)$ and $(\omega_{vv,\min}^b, \omega_{vv,\max}^b)$ are defined after observation of measurements.

5.1.2 Mode merging

Since many of the modes of the system get excited both in the horizontal and vertical directions (see Figure 4), the same mode may be estimated from both measurements. Joining the two sets of $M + 1$ modes independently obtained from the two one-dimensional fits leads to a set of $2(M + 1)$ mode candidates, from which pairs of numerically close mode estimations (i.e., corresponding to the same mode of the system) may appear. We merge the $2M$ mode estimations in $(\omega_{\min}, \omega_{\min})$ into a set of M' modes (with $M' \leq 2M$). First, we perform clustering on mode frequencies. Then, from each of the M' clusters, we keep the mode estimation presenting a natural frequency that is closest to the cluster centroid. From now on, the total number of modes, including the two bridge hill modes, will be referred to as M .

5.2 Error computation

Because optimization is carried out from measurements Y_{hh} and Y_{vv} simultaneously, we compute the diagonal modeling error ε' at iteration k as the sum of one-dimensional modeling errors when obtained through the common set of parameters (frequencies and bandwidths) $\mathbf{x}|_k$.

5.3 Constraint definition

Once initial mode estimates have been merged, constraints are defined analogously to the one-dimensional case. The only remarkable difference is the use of two broad resonances to represent the bridge hill.

5.4 Estimation of gains

We perform gain estimation by working with a frequency-domain expression of Equation (1) as described next. From a two-dimensional admittance measurement (symmetric) matrix \mathbf{Y} , let \mathbf{y}_{hh} , \mathbf{y}_{hv} , and \mathbf{y}_{vv} be complex-valued column

vectors each containing N frequency-domain samples of its corresponding entry in \mathbf{Y} , leading to a $2N \times 2$ matrix of the form

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_{hh} & \mathbf{y}_{hv} \\ \mathbf{y}_{hv} & \mathbf{y}_{vv} \end{bmatrix}. \quad (4)$$

Now, we proceed with rewriting the right-side of Equation (1) in matrix form as constructed from linear combinations of frequency-domain samples of the individual modal responses $H_m(\omega)$. First, we define a $N \times M$ matrix \mathbf{H} as

$$\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_m \dots \mathbf{h}_M], \quad (5)$$

where each \mathbf{h}_m is a complex column vector containing N samples of $H_m(\omega)$. With matrix \mathbf{H} , we construct a $2N \times 2M$ block-diagonal matrix \mathbf{B} defined as

$$\mathbf{B} = \begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}, \quad (6)$$

which can be interpreted as a two-dimensional modal basis. The next step is to set up a $2M \times 2M$ block-symmetric matrix \mathbf{R} as

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{hh} & \mathbf{R}_{hv} \\ \mathbf{R}_{hv} & \mathbf{R}_{vv} \end{bmatrix}, \quad (7)$$

where \mathbf{R}_{hh} , \mathbf{R}_{hv} , and \mathbf{R}_{vv} are $M \times M$ diagonal, real matrices. In the m -th entry of the diagonal of matrix \mathbf{R}_{hh} appears the gain from entry (1, 1) of the individual gain matrix \mathbf{R}_m in Equation (1). Analogously, matrix \mathbf{R}_{hv} will contain gains from the M entries (1, 2), and \mathbf{R}_{vv} from entries (2, 2). Now, with modal basis \mathbf{B} and gain matrix \mathbf{R} , it is possible to write an expression for model $\hat{\mathbf{Y}}$ as

$$\hat{\mathbf{Y}} = \mathbf{B}\mathbf{R}\mathbf{S}, \quad (8)$$

where \mathbf{S} is a $2M \times 2$ matrix of ones which acts as the summation of Equation (1). It is important to note that $\mathbf{R} \succeq 0 \Leftrightarrow \mathbf{R}_m \succeq 0 \forall m \in \{1, \dots, M\}$, implying that the model $\hat{\mathbf{Y}}$ will be passive if matrix \mathbf{R} is positive semidefinite. Now we are ready to express the modal gain estimation problem as an error minimization problem that includes a positive semidefinite constraint on matrix \mathbf{R} . If expressing the model approximation error $\varepsilon(\mathbf{Y}, \hat{\mathbf{Y}})$ as

$$\varepsilon(\mathbf{Y}, \hat{\mathbf{Y}}) = \|(\hat{\mathbf{Y}} - \mathbf{Y})\| = \|(\mathbf{B}\mathbf{R}\mathbf{S} - \mathbf{Y})\|, \quad (9)$$

where $\|\cdot\|$ represents a suitable matrix norm, the problem can be written as

$$\begin{aligned} & \underset{\mathbf{R}}{\text{minimize}} && \|(\mathbf{B}\mathbf{R}\mathbf{S} - \mathbf{Y})\| \\ & \text{subject to} && \mathbf{R} \succeq 0, \end{aligned} \quad (10)$$

which is a matrix norm minimization problem with a positive semidefinite constraint. This convex problem can be solved via semidefinite programming by means of CVX, a package for specifying and solving convex programs [9].

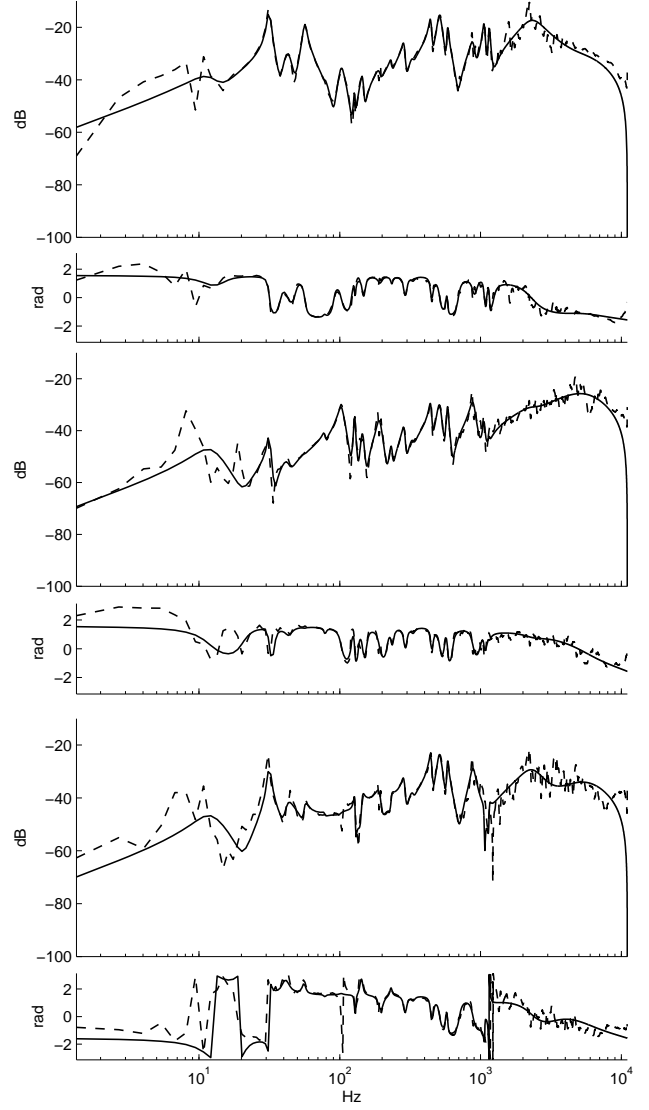


Figure 6. Violin synthetic admittance matrix $\hat{\mathbf{Y}}$ modeling example. Top to bottom, magnitude and phase plots of $\hat{Y}_{hh}(f)$, $\hat{Y}_{vv}(f)$, and $\hat{Y}_{hv}(f)$. $M = 18$, $f_{\min} = 8$ Hz, $f_{\max} = 1300$ Hz.

6. RESULTS AND DISCUSSION

Modeling examples for violin, viola, and cello are respectively shown in Figure 6, 7, and 8. In all three cases, with respective model orders between $M = 18$ and $M = 27$, high accuracy (including phase matching) can be observed between 100 Hz and 6 KHz, where measurement coherence was acceptable. In particular for the cello, the interaction between modes of the measurement apparatus and lower-frequency modes of the instrument made measuring and modeling a more difficult task. In general, both accuracy and convergence times are improved if carrying out the estimation on a warped frequency axis [3, 10]. Moreover, truncation of spectral domain samples above 6 kHz was needed in order to avoid artifacts caused by measurement limitations.

It is very important to include the lower frequency region (i.e., between 5 Hz and 100 Hz) in the fitting process

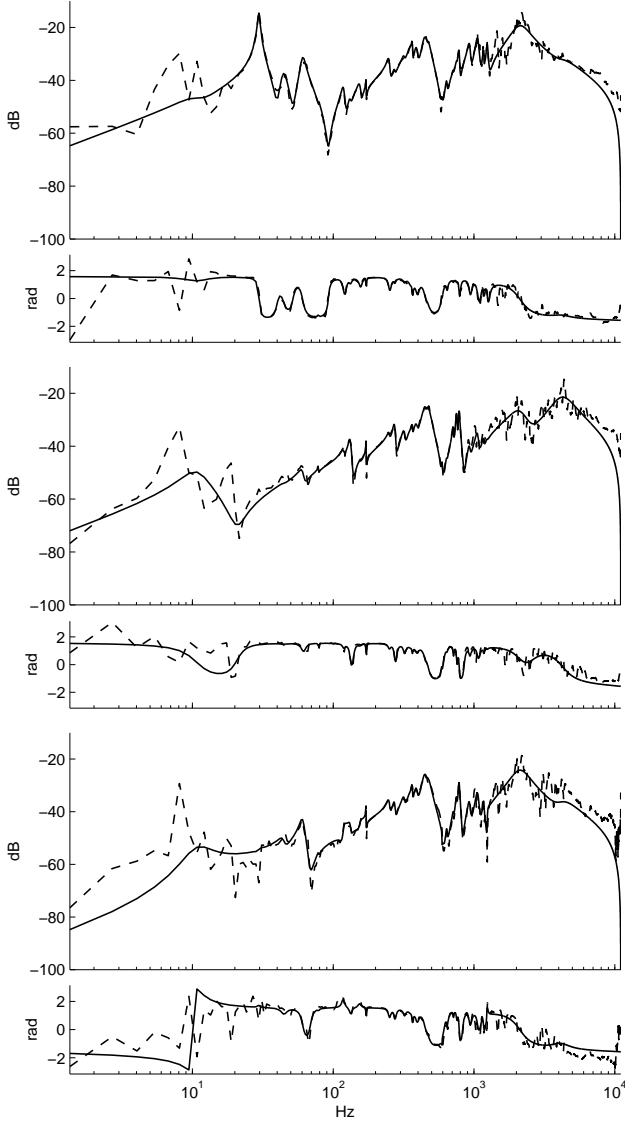


Figure 7. Viola synthetic admittance matrix \hat{Y} modeling example. Top to bottom, magnitude and phase plots of $\hat{Y}_{hh}(f)$, $\hat{Y}_{vv}(f)$, and $\hat{Y}_{hv}(f)$. $M = 22$, $f_{\min} = 8$ Hz, $f_{\max} = 1300$ Hz.

by setting design parameter ω_{\min} close to dc. This allows the modes of the measurement apparatus (prominent peaks below 100 Hz) to also be modeled, leading to a more consistent overall estimation that accounts for the interaction of such modes with the *real* modes of the instrument. Once the estimation is finished, those modes and their respective gain matrices can be discarded from Equation (1).

Regarding implementation, an elegant re-formulation of second-order sections proposed in [11] and later applied in [6] allows to maintain the parallel structure, leading to a straightforward realization as a reflectance. Our results from applying such re-formulation have been used to construct lumped terminations where four two-dimensional digital waveguides are coupled without the need for parallel adaptors (as in wave digital filters—see [6]). Example sounds, including one-pole filters to simulate string losses,

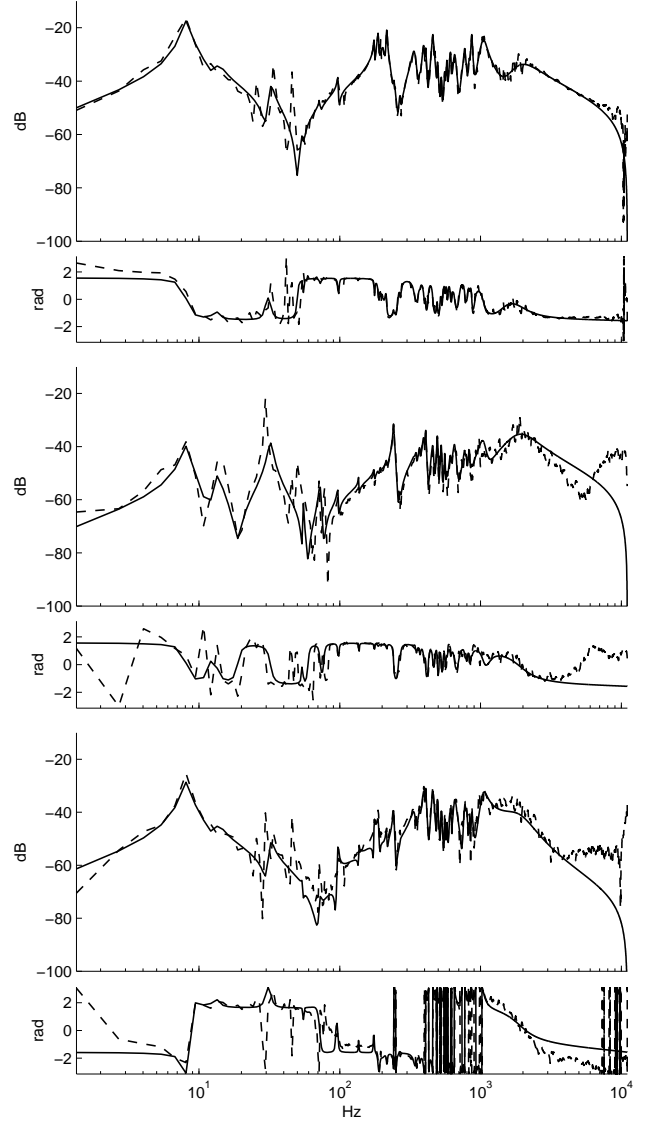


Figure 8. Cello synthetic admittance matrix \hat{Y} modeling example. Top to bottom, magnitude and phase plots of $\hat{Y}_{hh}(f)$, $\hat{Y}_{vv}(f)$, and $\hat{Y}_{hv}(f)$. $M = 27$, $f_{\min} = 8$ Hz, $f_{\max} = 900$ Hz.

are available online¹. The application of these models to bowed-string simulation with two-dimensional transverse string motion is imminent.

A potential improvement to the fitting method goes around embedding the semidefinite programming step as part of an outer loop in which mode parameters are estimated, although it would imply a higher computational cost. By increasing the model order and redefining design parameters it would be possible to represent the bridge hill region more accurately; yet, a perceptual evaluation might be needed to confirm improvements. Further tests might encourage the construction of statistical admittance models, where modal frequencies, bandwidths, and amplitudes follow empirically inferred distributions. An extension of the framework to include radiation measurements is currently under study.

¹ <http://ccrma.stanford.edu/~esteban/adm/smac13>

Acknowledgments

This work was partially funded by the Catalan Gov. through a Beatriu de Pinós Fellowship. Thanks go to Argyris Zymnis and Jonathan S. Abel for inspiring discussions.

7. REFERENCES

- [1] T. Rossing, *The Science of String Instruments*. Springer New York, 2010.
- [2] J. O. Smith, *Physical Audio Signal Processing, December 2008 Edition*. <http://ccrma.stanford.edu/~jos/pasp/>, accessed 2012, online book.
- [3] —, “Techniques for digital filter design and system identification with application to the violin,” Ph.D. dissertation, Stanford University, 1983.
- [4] M. Karjalainen and J. O. Smith, “Body modeling techniques for string instrument synthesis,” in *Proc. of the International Computer Music Conference*, 1996.
- [5] K. D. Marshall, “Modal analysis of a violin,” *Journal of the Acoustical Society of America*, vol. 77:2, pp. 695–709, 1985.
- [6] B. Bank and M. Karjalainen, “Passive admittance matrix modeling for guitar synthesis,” in *Proc. of the 13th International Conference on Digital Audio Effects*, 2010.
- [7] J. M. Adrien, “The missing link: Modal synthesis,” in *Representations of Musical Signals*, G. D. Poli, A. Piccialli, and C. Roads, Eds. MIT Press, 1991, pp. 269–267.
- [8] MATLAB, *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.
- [9] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 1.21,” <http://cvxr.com/cvx/>, Apr. 2011.
- [10] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, “Frequency-warped signal processing for audio applications,” *Journal of the Audio Engineering Society*, vol. 48(11), pp. 1011–1031, 2000.
- [11] M. Karjalainen, “Efficient realization of wave digital components for physical modeling and sound synthesis,” *IEEE Trans. Audio, Speech, and Lang. Process.*, vol. 16:5, pp. 947–956, 2008.

EVALUATING VIOLIN QUALITY: A COMPARISON OF PLAYER RELIABILITY IN CONSTRAINED VS UNCONSTRAINED TASKS

Charalampos Saitis^{1†} Gary P Scavone¹ Claudia Fritz² Bruno L Giordano³

¹Centre for Interdisciplinary Research in Music Media and Technology, McGill University

²Lutherie-Acoustique-Musique, Université Pierre et Marie Curie, UMR CNRS 7190

³Institute of Neuroscience and Psychology, University of Glasgow

[†]charalampos.saitis@mail.mcgill.ca

ABSTRACT

The overall goal of the research presented here is to better understand how players evaluate different qualities of the violin. To this end, we investigated intra- and inter-individual consistency in preference judgements by experienced violinists. Results from two previous studies that involved free-playing evaluative tasks showed that players are self-consistent in their preference for violins and tend to agree of what particular qualities they look for in an instrument (in this case, “richness” and “dynamic range”). However, the perception of the same attributes widely varies across individuals, thus likely resulting in large inter-individual differences in the preference for violins. A third study was conducted to further investigate the perceptual evaluation of richness and dynamic range in constrained- vs. unconstrained-playing tasks. Results indicated that specifying the musical material and technique removes a significant amount of inter-individual variability: the more focused the task, the more self-consistent violinists are and the more they agree with each other.

1. INTRODUCTION

For a period spanning more than ten years, Bissinger conducted a wide range of acoustical and structural dynamics measurements on 17 violins [1]. Those instruments were quality-rated from “bad” to “excellent” by a professional player and Bissinger himself. Attempts to quantify the characteristics of “excellent” violins were largely inconclusive, which led Bissinger to remark: “What truly defines violin excellence? If the answer is truly excellent violinists, then the reliability-reproducibility of their psychoacoustic judgements must draw more attention.” The research presented here takes this “contrarian viewpoint” (in Bissinger’s own wording) as a starting point and aims to quantify the extent to which skilled players are consistent at assessing violins and whether there is agreement between violinists.

We previously carried out two perceptual experiments based on a carefully controlled playing-based procedure for the perceptual evaluation of violins [2]. The first experiment was designed to examine both within-individual consistency

and between-individuals agreement across a certain range of violins. In the first session of the experiment, 20 skilled violinists were asked to freely play a set of 8 different violins, evaluate their quality, and order them by preference. Upon completing the task, participants had to comment on the ranking process and provide rationale for their choices. They had to repeat the ranking task 5 times and return for a second, identical session 3–7 days after having completed the first session. In total, players ranked each violin $5 \times 2 = 10$ times. Results showed that violin players are self-consistent when evaluating different violins in terms of overall preference. However, a significant lack of agreement between violinists was observed.

A second experiment was then conducted to investigate the origin of inter-individual differences in the preference for violins and to measure the extent to which different attributes of the instrument influence preference. Thirteen experienced violin players were asked to freely play a set of 10 different instruments and rate them according to *playability* (how easy they are to play), *response*, *richness*, *balance* (across all strings), *dynamic range* and preference. The rating attributes-criteria were determined based on the analysis of verbal data collected in the first experiment as well as the potential for the descriptors to be correlated with measured vibrational properties of the violin. Participants had to rate one violin on all scales at a time. The rating task was repeated 3 times. Results showed that the perception of the same violin attributes widely varied between individual players, while confirming the large inter-individual differences in the preference for the violins observed in the first experiment. Importantly, despite the variability in the evaluation of both preference and violin attributes, violinists appeared to strongly agree on their preference for violins with a rich sound and, to a lesser extent, a broad dynamic range. As such, what makes a violin good might, to a certain extent, lie in the ears and hands of the performer not because different performers prefer violins with largely different qualities, but because the perceptual evaluation of violin attributes widely considered to be important for a “good” violin vary across individuals. This important conclusion may explain the limited success of previous studies at quantifying the differences between “good” and “bad” violins from vibrational measurements.

From verbal responses collected in the first experiment, a classification scheme emerged that illustrates the complex links between the different player-typical concepts (e.g.,

response, timbre), properties (e.g., ease, richness), and underlying themes (handling, sound and their relevance to the individual) [3]. In particular, richness emerged as a key perceptual factor in violin quality, supporting the observations in the second experiment.

One of many hypotheses about the origin of the large inter-individual differences in violin preference is that players may take varying playing approaches to assess different attributes of the instrument. In the previous two experiments, no playing constraints were imposed on the evaluation process (e.g., specific repertoire). Participants were instead instructed to follow their own strategy with respect to what and how to play. To tease apart the effects of the playing skills of different individuals, a new experiment was designed to investigate the perceptual evaluation of richness and dynamic range in playing tasks based on prescribed musical material and techniques. The objective was to compare intra-individual consistency and inter-individual agreement in constrained (i.e., playing only certain notes on certain registers) versus unconstrained (i.e., playing a certain excerpt from the violin repertoire) tasks for the cases of richness and dynamic range. The prescribed evaluation materials and techniques were determined based on verbal data collected in an online survey that was conducted prior to the main experiment. We chose to focus on the perceptual characteristics of richness and dynamic range as they had been previously found to be highly correlated with violin preference.

2. METHOD

2.1 Participants

Sixteen skilled string players took part in this experiment (8 females, 8 males; average age = 32 yrs, SD = 8 yrs, range = 21–55 yrs). They had at least 17 years of violin experience (average years of violin training = 25 yrs, SD = 8 yrs, range = 17–48 yrs; average hours of violin practice per week = 15 hrs, SD = 11 hrs, range = 3–35 hrs), owned violins with estimated prices ranging from \$3K to \$70K, and were paid for their participation. Eleven participants described themselves as professional musicians and 10 had higher-level

Violin	Origin	Luthier	Year	Price
A	Italy	Contino	1916	\$71K
B	Switzerland	-	2003	\$30K
C	Denmark	Hjorth	1914	\$20K
D	Germany	Unknown	Unknown	\$10K
E	China	-	2011	\$2.7K

Table 1. Violins used in Study 3. Violin D was included in Study 1 (highest preference score) and Study 2. Its origin is based on a luthier’s informal appraisal, as there is no information regarding the make and age of this violin. The names of living luthiers are not provided for confidentiality purposes.



Figure 1. Richness-constrained task

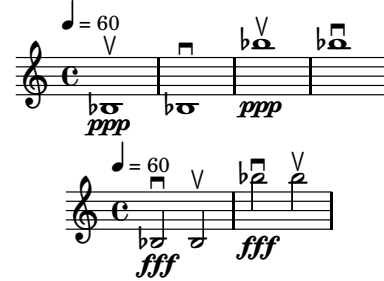


Figure 2. Dynamic range-constrained task

degrees in music performance (MMus, MA, DMus, DMA). They reported playing a wide range of musical styles [classical (81%), folk (13%), jazz/pop (6%), and contemporary (6%)] and in various types of ensembles [symphonic orchestra (38%), chamber music (31%), folk/jazz band (25%), and solo (19%)].

2.2 Violins

Five violins of different make (Europe, North America, China), age (1914–2011) and price (\$2.7K–\$71K) were used (see Table 2). They were chosen from two local luthier workshops in order to form, as much as possible, a set of violins with a wide range of characteristics. The violins had not been played on a regular basis as most were from the available sales stock of the workshops. The respective luthiers provided the price estimates and tuned the instruments for optimal playing condition based on their own criteria. Participants were given the option to either use a provided shoulder rest (Kun Original model), or use their own, or use no shoulder rest. The experiment took place in a diffuse room with a surface of 46.8 m² and reverberation time of about 0.3 s to help minimize the effects of room reflections on the direct sound from the violins [4]. All other experimental conditions (i.e., visual occlusion and choice of a bow) were as in the previous studies [2].

2.3 Tasks

For each one of the perceptual characteristics of richness and dynamic range, a constrained and an unconstrained task were designed. For richness-constrained, participants were asked to play certain notes on the G-string (see Fig. 1). They were instructed to play *détaché*, first *without vibrato* followed by a repetition *with vibrato* using the whole bow. For dynamic range-constrained, participants were asked to play the same note first in the lower and then in the upper register (see Fig. 2). They were instructed to play *détaché*, *without vibrato*, as soft and as loud as possible to obtain a clear sound (i.e., the sound doesn’t break).

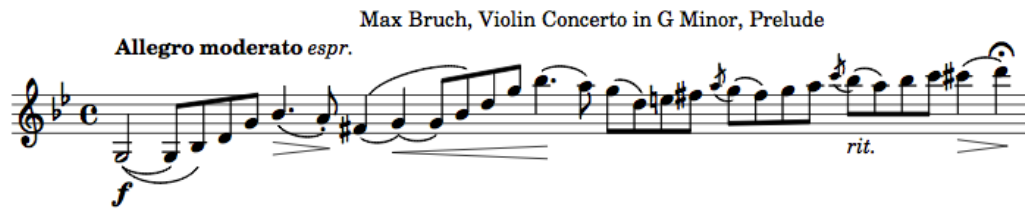


Figure 3. Unconstrained task

The unconstrained task was recurrent across the two criteria and involved playing the opening solo passage from Max Bruch's Violin Concerto No. 1 in *G* minor, Op. 26 (Movement I: Prelude; see Fig. 3). The particular excerpt was chosen because it incorporates the whole range of the instrument (as opposed to certain registers) as well as a variety of techniques and dynamics. The unconstrained task was also used for the evaluation of preference.

Unlike the free-playing approach adopted in our previous studies, the idea of constrained versus unconstrained playing in this experiment concerned the playing range of the instrument on which violinists were permitted to focus (1 or 2 strings versus all strings) as well as the playing technique they could apply (strict versus loose instructions) during the evaluation procedure. In this respect, the idea of "unconstrained" is not similar to that of "free." In the latter, which was not used in the current study but only in the previous two, the participants were encouraged to choose both their own materials and techniques—and those would often change from one trial to the next, whereas the musical material was common for all players in the former.

2.4 Procedure

The experimental session lasted two hours and was organized in three parts. The first part involved two training rankings with three violins, which were distinct from the five violins used in the actual study, to help participants familiarize themselves with each of the constrained-playing tasks respectively. In the second part, participants were asked to rank-rate (see next paragraph) the violins in terms of richness first and then dynamic range according to the respective constrained task. Each task involved three repetitions (trials) and all players carried out the two tasks in the same order. In the third part, participants were asked

to rank-rate the violins in terms of richness, dynamic range and preference according to the unconstrained task. Each of the three criteria was presented once in each of three subsequent blocks of trials. The order of presentation of the criteria within each block of trials was randomized (determined by computer calculations). In total, participants ranked-rated all violins $2 \times 3 + 3 \times 3 = 15$ times.

In each trial, participants were first presented with all violins placed on a table in random order (determined by computer calculations) by the experimenter. Participants were then asked to simultaneously rate each violin on the same unipolar discrete scale using separate, identical on-screen sliders, thus providing a ranking of the five violins at the same time (see Fig. 4). They had to move each slider (i.e., assess each instrument) before being allowed to move to the next trial. Participants were instructed to always rate their top choice as 1 and their lowest as 0. They were not allowed to assign the same rank-rating to two or more instruments. Participants were instructed to maximize evaluation speed and accuracy. They were encouraged to play their own violin whenever they needed a reference point during the experiment. To minimize fatigue, participants were encouraged to take breaks between trials whenever needed.

3. RESULTS

Three different analyses were carried out. Firstly, the measures of intra- and inter-individual consistency for each of the evaluation tasks were assessed and compared. Furthermore, a two-way repeated-measures analysis of variance was employed to investigate the effects of condition (i.e., constrained versus unconstrained) and attribute (richness versus dynamic range) on the measures of intra-individual consistency. The measures of intra- and inter-individual consistency recorded during this study were also compared with those recorded during Experiments 1 and 2. Secondly, the effects of participant characteristics (self-reported) on the measures of intra-individual consistency computed for each of the tasks were assessed. Thirdly, an overall score for each of the violins was derived.

3.1 Intra- and inter-player consistency

For each task, intra- and inter-individual consistency were measured and assessed based on Lin's concordance correlation coefficient ρ_c between ratings given on different blocks of trials [5]. The concordance correlation coefficient ρ_c



Figure 4. Testing interface

is a special case of the Pearson product-moment correlation coefficient that measures departures from the equality lines with slopes $\pm 45^\circ$. As such, ρ_c does not assume linear relationships. For a given participant A, intra-individual consistency was estimated as the average of the ρ_c between ratings of A across all trials. Inter-individual consistency was given by averaging the ρ_c between ratings of A and those of all other participants across all trials. Note that according to this definition, the inter-individual consistency measures for participants A and B would be computed by considering the same set of 9 ρ_c measures between the 3 ratings of participant A and those of participant B. In order to minimize one source of dependence between the inter-individual consistency measures for different participants, correlations were distributed among participants at random (e.g., for participant A the inter-individual consistency measure considered 4 or 5 randomly selected $\rho_c(A, B)$ measures, whereas for participant B it included the other 5 or 4 respectively). However, there is another source of dependence as all correlations come from the same matrix and are therefore linked to each other. As a result, any statistical inferences on inter-individual consistency such as confidence intervals of the mean (see Fig. 5) or *t*-tests should be treated with caution.

For the constrained tasks, the average measure of intra-individual consistency was substantially high for richness, average value = .697, but less so for dynamic range, average value = .472. Concerning the unconstrained tasks, the average measure of intra-individual consistency was relatively high for richness and preference, average value = .443 and .442 respectively, but considerably lower for dynamic range, average value = .292. Inter-individual consistency was generally low for both constrained and unconstrained tasks, $.145 \leq \text{average value} \leq .189$, except for richness-constrained, average value = .305. Considering the unconstrained tasks, no significant differences emerged between the intra-individual consistency measured for the preference task on the one hand, and the richness and dynamic range tasks on the other [paired samples $t(15) \leq 1.87$, $p \geq .081$].

3.2 Constrained vs. unconstrained evaluation

To examine the effect of constrained versus unconstrained task (condition) in the perceptual evaluation of richness and dynamic range (attribute) on self-consistency, a two-way repeated-measures analysis of variance was conducted on the corresponding measures of intra-individual consistency. Following the notable decrease in self-consistency from the constrained to the unconstrained tasks for each of the two attributes as well as from richness to dynamic range in both the constrained and unconstrained tasks, the analysis of variance revealed that both condition and attribute had a significant effect on how self-consistent participants were in their judgements [$F(1, 15) = 8.64$, $p = .01$ and $F(1, 15) = 7.72$, $p = .014$ respectively]. The interaction between attribute and condition fell short of significance [$F(1, 15) = .25$, $p = 0.628$], hence the two factors do not appear to influence each other.

	Within	Between
Experiment 1		
Preference	.62(.09)	.015(.04)
Experiment 2		
Easy to play	.24(.16)	.064(.04)
Response	.328(.1)	.042(.05)
Richness	.389(.16)	.068(.06)
Balance	.203(.12)	-.005(.03)
Dynamic range	.333(.16)	.071(.03)
Preference	.38(.14)	.089(.05)
Experiment 3		
Richness (c)	.697(.17)	.305(.09)
Dynamic range (c)	.472(.16)	.154(.09)
Richness (u)	.443(.2)	.189(.08)
Dynamic range (u)	.292(.17)	.145(.06)
Preference (u)	.442(.18)	.179(.08)

Table 2. Across-participants average intra-individual consistency and inter-individual agreement measures in each of the experiments (c = constrained, u = unconstrained).

3.3 Comparisons with previous studies

The overall measures of intra- and inter-individual consistency for richness and dynamic range were compared with those measured in the respective attribute-rating scales used in the second study, wherein players were instructed to develop their own strategy. Both intra-individual consistency and inter-individual agreement for the evaluation of richness-constrained were notably higher than in the second experiment, average value = .697 and .305, and .389 and .068 respectively. Indeed, the large increase in both intra- and inter-individual consistency was found to be significant [independent samples $t(27) = 2.81$, $p = .009$, and $t(27) = 4.59$, $p < .001$, equal variance]. In the case of dynamic range-constrained, intra- and inter-individual consistency were also higher, albeit to a lesser extent, than in the second study, average value = .472 and .154, and .333 and .071 respectively. Although the relative increase in intra-individual consistency fell short of significance [independent samples $t(27) = 1.32$, $p = .199$, equal variance], the increase in inter-individual agreement was significant [independent samples $t(19.78) = 3.36$, $p = .003$, unequal variance]. Intra- and inter-individual consistency in richness-unconstrained were moderately higher than in the second experiment, average value = .443 and .189, and .389 and .068 respectively. The increase in self-consistency was not significant [independent samples $t(27) = .44$, $p = .665$, equal variance], but the increase in inter-individual agreement was [independent samples $t(27) = 2.33$, $p = .028$, equal variance]. In dynamic range-unconstrained, intra-individual consistency was slightly and not significantly lower than in the second study, average value = .292 and

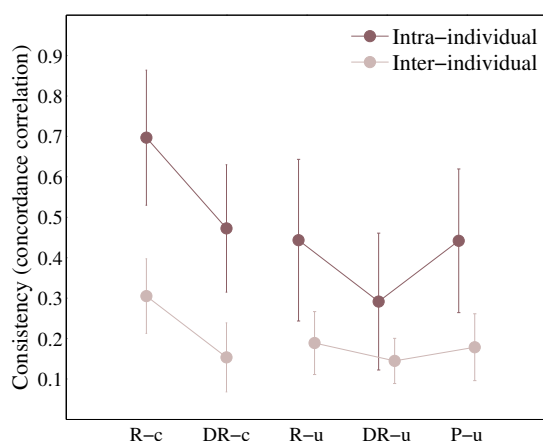


Figure 5. Across-participants average intra- and inter-individual consistency scores for each of the constrained and unconstrained playing tasks (R = richness, DR = dynamic range, P = preference; c = constrained, u = unconstrained; error bar = 95% confidence interval of the mean). See text for details on averaging of concordance correlations.

.333 respectively [independent samples $t(27) = -.38$, $p = .709$, equal variance]; inter-individual agreement was significantly higher than in the second experiment, average value = .145 and .071 respectively [independent samples $t(21.32) = 3.56$, $p = .002$, unequal variance].

The overall measures of intra- and inter-individual consistency collected during the first experiment (i.e., preference judgements) and those measured during the second study for the preference-rating scale were compared with those measured during the current study for the preference task. Intra-individual consistency for the evaluation of preference was higher in the first than in the third experiment, average value = .62 and .442 respectively, but the decrease was not significant [independent samples $t(23.5) = -1.88$, $p = .072$, unequal variance]. In the second study, intra-individual consistency in preference judgements was lower than in the current one, average value = .38 and .442 respectively, but the increase fell short of significance [independent samples $t(27) = .56$, $p = .577$, equal variance]. Inter-individual consistency in preference judgements gradually increased from the first to the second to the third study, average value = .015, .089 and .179 respectively. Despite the increase from the second to the third experiment not being significant [independent samples $t(23.52) = 1.91$, $p = .068$, unequal variance], the overall increase from the first to the third study was found to be significant [independent samples $t(20.38) = 3.79$, $p = .001$, unequal variance].

3.4 Violin ratings

For each of the violins, a task-specific score defined as the across-participants average rating of a violin throughout all trials was computed. The across-participants average violin rating scores for each task are shown in Fig. 6, where we

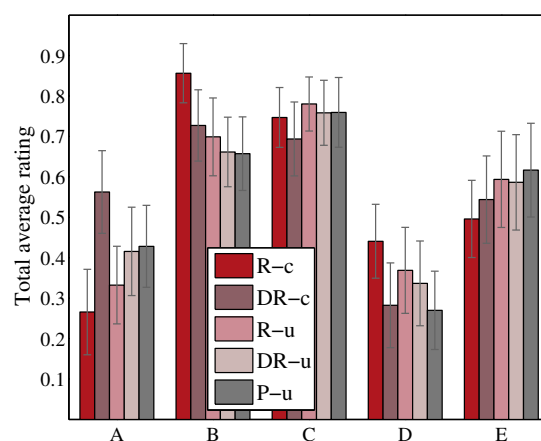


Figure 6. Across-participants across-trials average rating scores for each violin (R = richness, DR = dynamic range, P = preference; c = constrained, u = unconstrained; error bar = 95% confidence interval of the mean).

observe the same grouping pattern for all tasks: violins A and D always alternated between the two lower ranks, violin E was always placed in the middle position and violins B and C alternated between the two higher ranks (in the case of dynamic range-constrained, the grouping was only slightly different as violin E alternated with A). In particular, violin A was consistently chosen as the least rich instrument and violin D was consistently thought to have the narrowest dynamic range. Violin B was characterized as both the most rich and having the broadest dynamic range when evaluated in the constrained tasks; for the unconstrained tasks participants appeared to prefer violin C over B. The difference in the playing range between constrained and unconstrained tasks should be considered here. It may be possible that violin B has a particularly rich low end compared to violin C but the latter “wins” in the middle and high register. Similarly, violin B may allow better control of dynamics for the low and high *B♭* than violin C but the latter’s dynamic range may be better highlighted through playing the Bruch passage.

4. DISCUSSION

The results of this experiment showed that experienced violin players are self-consistent when evaluating different violins by focusing on a specific attribute of the instrument and following prescribed musical material and technique. An analysis of variance revealed that violinists are significantly more self-consistent in well-focused evaluation tasks than in a less restrained setting. Several methodological differences between the two types of tasks could explain this effect. The non-randomized order of the constrained tasks (i.e., first all richness trials followed by all dynamic range trials) gave participants a better opportunity to stabilize their responses than in the unconstrained tasks (where the order of the three tasks was randomized in each of the three blocks of trials). Moreover, the order of the constrained

tasks was recurrent across participants, while the (random) order of the unconstrained tasks was different (i.e., randomized) for each participant. Playing a violin concerto passage that involves a wider range of notes and nuances (unconstrained tasks) entailed a more differentiated evaluation strategy than playing certain notes in a certain way (constrained tasks). Furthermore, as the unconstrained tasks were carried out in the second half of the session, fatigue affected the level of attention in evaluating richness and dynamic range as well as preference.

Participants were considerably more self-consistent in the constrained-playing tasks involved in this experiment than in the respective attribute-rating scales involved in the second study whereby there were no playing constraints. Several methodological differences between the two experimental settings could explain this effect. The rating of richness alongside other attributes (in the second study) did not allow the same level of attention as focusing only on richness. Similarly, the level of attention is increased when the number of violins is reasonably small. Furthermore, being able to compare the various violins to determine ratings is ecologically more valid than rating one violin at a time (as in the second experiment).

Participants were less self-consistent when evaluating preference in this study than in the first experiment. This could be explained by the higher number of repetitions in the first study (10 ranks for each violin across the 2 sessions) than in the current experimental setup (3 ranks for each violin) as well as the presense of two attribute tasks alongside preference. On the other hand, participants appeared slightly more self-consistent in this study than in the second experiment. To a certain extent, these observations seem to suggest that when evaluating a set of violins, comparing all instruments at a time is more meaningful and thus more reliable than assessing each violin individually. Furthermore, no significant differences were observed in this study between the level of intra-individual consistency in the preference ratings and that in the attribute ratings (unconstrained tasks).

More importantly, a higher inter-individual agreement in the playing tasks relative to the previous studies was present. This is further confirmed by the average ratings of the violins, whereby we observe three distinct groups in all tasks but for dynamic range-constrained (though the difference in the respective ordering is relatively minor). On the one hand, this observation seems to support the hypothesis that different violin players may take varying approaches to assess different attributes of the instrument and hence designing focused evaluative tasks may trigger more agreement between individuals. On the other hand, it is possible that participants were able to agree more with each other because they had to evaluate only five violins, a relatively smaller number than in the previous studies.

5. CONCLUSIONS

We reported a study aimed to investigate the effect of playing constraints on the assessment of violins by experienced musicians within the context of better understanding how players evaluate violin quality. We focused on the prefer-

ence for violins as well as the perceptual attributes of richness and dynamic range, which had previously been shown to be strongly associated with preference. We observed that the psychoacoustic judgements of violinists became more reliable as the tasks became more controlled.

While specifying the musical material and technique may improve consensus, there remains the issue of addressing differences in how people play. Different violinists may use different combinations of gestures when playing, each producing a fundamentally different behaviour of the instrument for a certain attribute. For example, player A may use more bow force than player B and thus produce a more *bright* timbre [6]. Further exploration is needed in this direction. That still would not address differences in the semantic interpretation of such verbal tags. To this end, we are currently studying player verbalizations using a linguistic approach to identify the meaning(s) of *richness* [7].

Acknowledgments

This project has been partially funded by the Centre for Interdisciplinary Research in Music Media and Technology and the Natural Sciences and Engineering Research Council of Canada. All violin shoulder rests were generously provided by Kun Shoulder Rest Inc. We thank Stephen McAdams, Jean-François Petiot and Jacques Poitevineau for constructive comments. We are grateful to luthiers Olivier Pérot and Tom Wilder for loaning the violins.

6. REFERENCES

- [1] G. Bissinger, "Structural acoustics of good and bad violins," *J. Acoust. Soc. Am.*, vol. 124, no. 3, pp. 1764–1773, 2008.
- [2] C. Saitis, B. L. Giordano, C. Fritz, and G. P. Scavone, "Perceptual evaluation of violins: A quantitative analysis of preference judgements by experienced players," *J. Acoust. Soc. Am.*, vol. 123, no. 6, pp. 4002–4012, 2012.
- [3] C. Saitis, C. Fritz, C. Guastavino, and G. P. Scavone, "Conceptualizing the "good" violin in preference descriptions by experienced players," in *Proc. Stockholm Music Acoust. Conf.*, Stockholm, Sweden, 2013.
- [4] G. Bissinger and F. Gearhart, "A standardized qualitative violin evaluation procedure?" *Catgut Acoust. Soc. J. (Series II)*, vol. 3, no. 6, pp. 44–45, 1998.
- [5] L. I. Lin, "A concordance correlation coefficient to evaluate reproducibility," *Biometrics*, vol. 45, no. 1, pp. 255–268, 1989.
- [6] E. Schoonderwaldt, K. Guettler, and A. Askenfelt, "An empirical investigation of bow-force limits in the Schelleng diagram," *Acust. Acta Acust.*, vol. 94, pp. 604–622, 2008.
- [7] P. Cheminée, "Est-ce bien *clair* ? Stabilité, instabilité et polysémie d'une forme lexicale en contexte," in *Le sentir et le dire: Concepts et méthodes en psychologie et linguistique cognitives*, D. Dubois, Ed. Paris: L'Harmattan, 2009.

ACOUSTICAL CONSTRAINTS AND INDIVIDUAL PREFERENCE IN THE COORDINATION OF COMPLEX BOWING PATTERNS

Erwin Schoonderwaldt

IMMM,
Hanover Univ. of
Music, Drama and Media,
Hanover, Germany
schoondw@gmail.com

Matthias Demoucron

IPEM,
Ghent University,
Ghent, Belgium
matthias.demoucron@ugent.be

Eckart Altenmüller

IMMM,
Hanover Univ. of
Music, Drama and Media,
Hanover, Germany
altenmueller@hmt-hannover.de

ABSTRACT

A series of recent studies have shed light on coordination in complex bowing gestures involving string crossings and bow changes, both in performance and perception. However, significant individual differences in coordination behavior were found, which puts into question the strictness of acoustical constraints. In this paper the inter- and intra-individual variance will be examined more closely. It is suggested that the acoustical constraints are (within certain limits) not very strict, and leave room for the performer to manipulate subtle quality aspects of the note transitions.

1. INTRODUCTION

1.1 Background

Complex bowing patterns, in which notes are alternately played on adjacent strings are common in violin repertoire (see note examples in Fig. 1a). They are often used to intertwine two musical lines, mostly melody and accompaniment, a compositional technique known as latent polyphony. From the performer it requires a precise coordination of bow changes and string crossings. At fast tempo (e.g. 16th notes) the bowing movements form typical circular patterns, as already demonstrated by Hodgson [1], the relative timing of bow changes and string crossings being inherent in their exact shape.

A series of recent studies provided new insights in the details concerning the performance of such bowing patterns. Measurement of bowing movements in performance revealed a remarkable coordination behavior. It was shown that string crossings anticipated bow changes, which was achieved by a small but consistent phase lead of the string-crossing movement over the to-and-fro movement of the bow responsible for playing the notes, both movement components being approximately sinusoidal [2, 3]. This coordination behavior could be reproduced in a perceptual study, in which participants were asked to optimize the coordination of simulated bowing movements controlling a

virtual violin [4]. The latter study confirmed the perceptual origin of the coordination behavior, providing support for the notion that complex bowing trajectories in performance for an important part emerge from auditory-motor interaction.

The study of complex bowing gestures can be considered as an extension of earlier approaches of studying motor control phenomena in bowed string instruments, which have mainly been focusing on more elementary aspects, such as *detaché* bowing on a single string [5–7]. Some preliminary observations of string crossings were reported in [8] with regard to coordination with finger action. The study reported here will focus on unimanual coordination of bowing movements in two spatial dimensions, combining aspects of auditory perception and motor behavior.

1.2 Aims

The current paper will mainly discuss preliminary outcomes of a recent motion capture study with 22 violinists. The participants belonged to three groups, namely (good) amateurs, violin students majoring in a music university, and established professional violinists.

A preliminary analysis reported in [9] revealed rather large differences in the coordination strategies of individual performers and, unexpectedly, no clear distinction between the three groups. This puts into question the strictness of the acoustical constraints associated with good note transitions. As argued in [4] and elsewhere in these proceedings [10] the optimization of note transitions involves a trade-off in which several acoustical features arising from the finite width of the string crossing (in particular “false” attacks and “after-ringing”) play a role. Individual coordination strategies might therefore be associated with different qualities of note transitions, which might reflect personal preference or distinct expressive purposes.

The aim of the current paper is to shed more light on common and individual aspects of coordination behavior in complex bowing gestures by considering between- and within-performer variability. An attempt will be made to relate the coordination strategies to perceptual qualities of the note transitions, partly building on the results of a simulation study presented elsewhere in these proceedings [10].

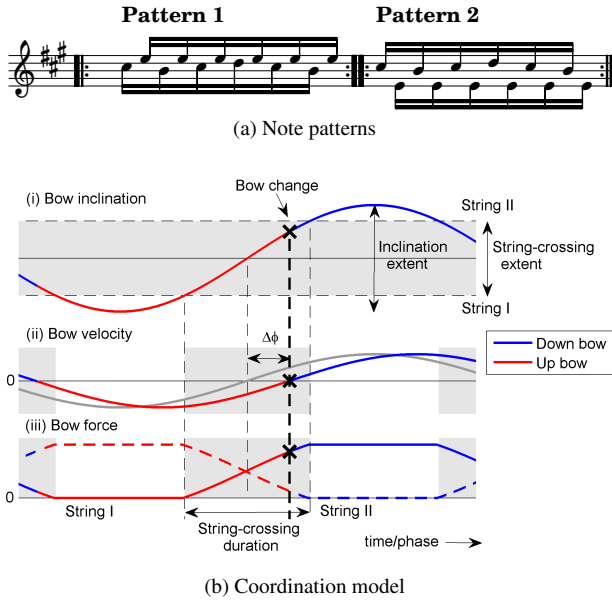


Figure 1. Note patterns and coordination model.

1.3 Coordination model

In this section a simplified coordination model is introduced, which is needed for a general understanding of the complex bowing gestures described here. A more comprehensive description can be found in [4] and [9], as well as elsewhere in these proceedings [10].

The coordination model is illustrated in Fig. 1b. It takes into account the finite width of the string crossing, indicated by the shaded areas. The string crossing is controlled via the inclination of the bow. During the string crossing, bow force is transferred from one string to the other. Bow changes correspond to zero crossings of bow velocity. Both bow inclination and bow velocity are modeled as sinusoidal over time. The main coordination parameters in the model are 1) the relative phase $\Delta\phi$ between the bow velocity and bow inclination signals, and 2) the normalized string-crossing range (normalized range r). The relative phase is here defined positive for a phase lead of bow inclination. The normalized range is obtained by dividing the string-crossing extent by the inclination extent. The timing of the bow change relative to the string crossing is dependent on the combination of relative phase and normalized range, which determine the conditions for stopping the “old” note and starting the “new” note.

The main simplifications with respect to actual performance are that bow force, and thus the width of the string-crossing transition are considered constant, and that the bow inclination movement is symmetric with respect to the center of the string crossing (no offset). Even though the model provides a reasonable description, in actual performance there can be considerable deviations from these simplifications, as can be seen in Fig. 2. Furthermore, bow inclination and bow velocity are usually not perfectly sinusoidal, giving rise to slight fluctuations of relative phase.

2. METHOD

In this paper the results of two complementary studies are discussed. A perceptual study was done using simulated complex bowing gestures controlling a virtual violin to provide insight in the acoustical constraints of note transitions. For a detailed description of the method see [4]. A detailed account of the acoustic features as a function of the main coordination parameters is presented elsewhere in these proceedings [10]. In addition, a motion capture study was done to provide insight into the coordination behavior of complex bowing patterns in actual performance. Additional information about the quality of the individual performances was obtained via expert ratings.

2.1 Perception study

The above coordination model was implemented in Max/MSP to control a virtual violin. The participants could control the parameters in real time. They were asked to adjust the coordination parameters via a simple graphical user interface in order to achieve the best sounding note transitions according to their judgment.

A total of 16 participants was recruited. They were all experienced string players, considered to be able to judge the subtle influence of the coordination parameters on the quality of the transitions. The stimuli consisted of note patterns 1 and 2 (see Fig. 1a) at different string combinations (transposed accordingly) and dynamic levels (different combinations of bow velocity, bow force and bow-bridge distance), yielding a total of eight different stimuli. Each stimulus was presented twice per experimental condition, yielding a total of 16 stimuli per experimental condition. They were played at a tempo of 100 bpm (note rate of 6.67 notes per second). In one of the experimental conditions (the one discussed in this paper) the participants were asked to adjust the coordination parameters via a 2D slider in order to obtain optimal sounding note transitions. The participants were entirely free to choose the combination of $\Delta\phi$ and r that gave the best sounding note transitions according to their judgment. For each stimulus the participants were given as much time as they needed to find the optimal combination of coordination parameters.

2.2 Motion capture study

Motion capture data (frame rate 240 fps) were collected with a 7 camera Qualisys 3D optical motion capture system, together with synchronized analog data from a sensor for measuring bow force, video and audio. Audio was recorded using a DPA-4099 microphone mounted on the violin. Bowing gesture data (including bow velocity, bow force, bow-bridge distance and bowing angles) were obtained using the methods for measuring and extraction of bowing parameters described in [11, 12]. In addition, instantaneous data on the width of the string-crossing areas taking the compliance of the strings and the bow hair into account was calculated using the method described in [13].

A total of 22 violinists participated in the experiment (7 male, 15 female; age: $M=28$; $SD=9.75$). The participants

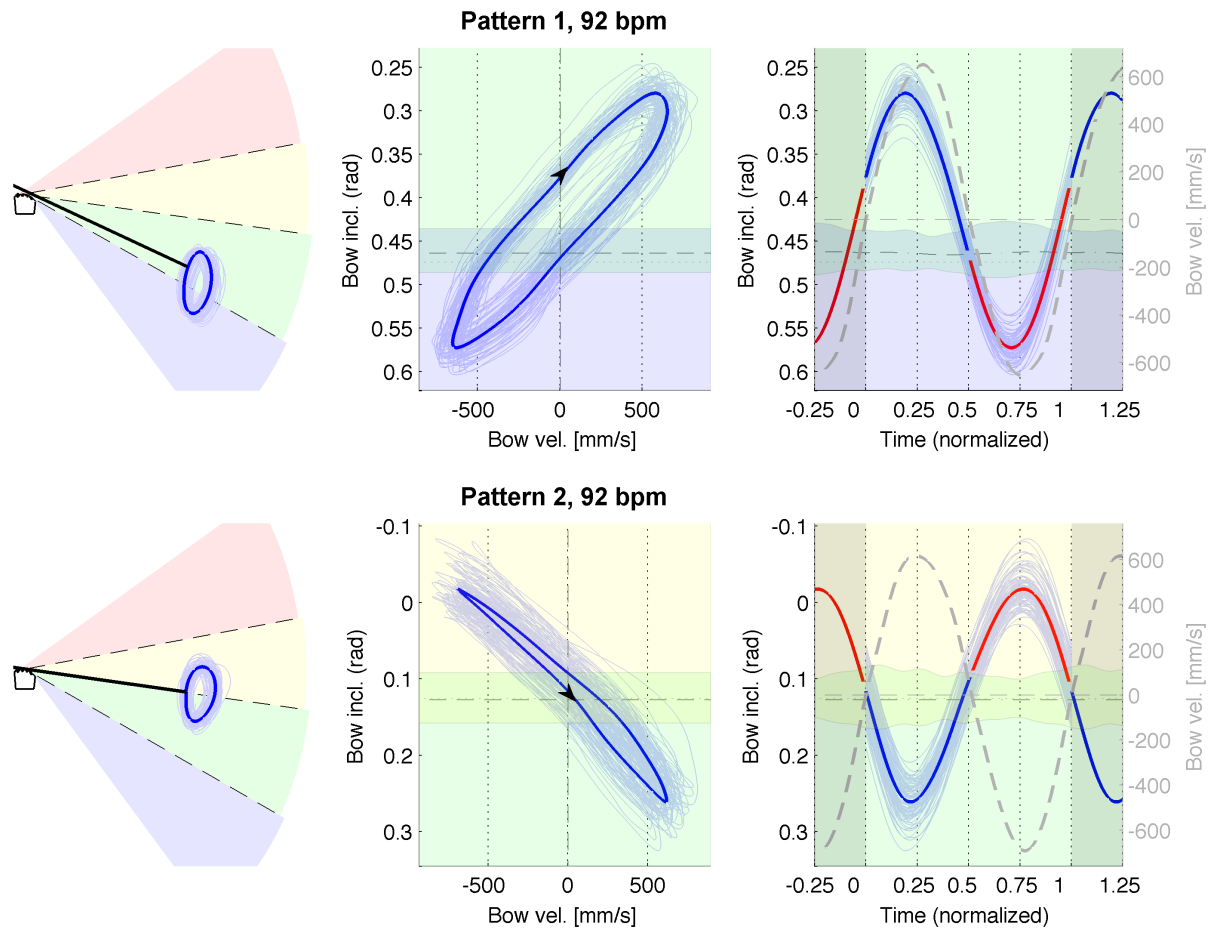


Figure 2. Performances of note patterns 1 and 2 played at 92 bpm by the same performer. Repeated cycles (about 100 per condition) are indicated by the light curves. The thick curves show the averages across repeated cycles. The left panels show the spatial trajectory of the frog from the player's perspective, the middle panels show bow inclination versus bow velocity, and the right panels show inclination (blue=down bow, red=up bow) and bow velocity (dashed gray line) versus time. In the middle and right panels the averaged string-crossing area is indicated by the overlapping colored regions associated with the strings. Bowing patterns 1 and 2 (bowing trajectory, incl vs. bow vel., and incl/vel vs. time).

were categorized in three groups with varying level of expertise: amateurs (8), students with violin as major subject (8), and established professionals (6).

The participants were asked to play simple repetitive musical patterns involving bow changes and string crossings, played on different string combinations. The current analysis is limited to a subset of the tasks involving the two note patterns shown in Fig. 1a played forte without metronome (preferred tempo between about 92-112 bpm), and with metronome at 92 and 112 bpm (note rates 6.1 and 7.5 notes/s, respectively). For each take the pattern was repeated during 6 bars, and each experimental condition was repeated three times, yielding about 100 cycles of the repetitive bowing pattern per experimental condition for each participant.

The focus of the analysis was on the coordination of bow inclination (timing of string crossings) and bow velocity (timing of bow changes). The signals were denoised using a low-pass second-order Butterworth filter (cut-off

30 Hz, applied back and forth to compensate for phase shifts). The instantaneous relative phase between bow velocity and bow inclination was obtained via a Hilbert transforms (both signals were approximately sinusoidal in the conditions included in the current analyses). The data was segmented into single period cycles cut at interpolated zero crossings in bow velocity (i.e. at the bow changes), and the cycles were aligned by resampling them (close to the original sample rate) in order to compute average patterns of the parameters of interest (see Fig. 2 for an example). Finally, features characterizing the timing and other details concerning of the coordination of string crossings and bow changes were extracted from the unsegmented signals, allowing for statistical analysis. All processing was done in Matlab®.

2.3 Expert ratings

Expert ratings of a selection of performances from each participant in the motion capture experiment were obtained

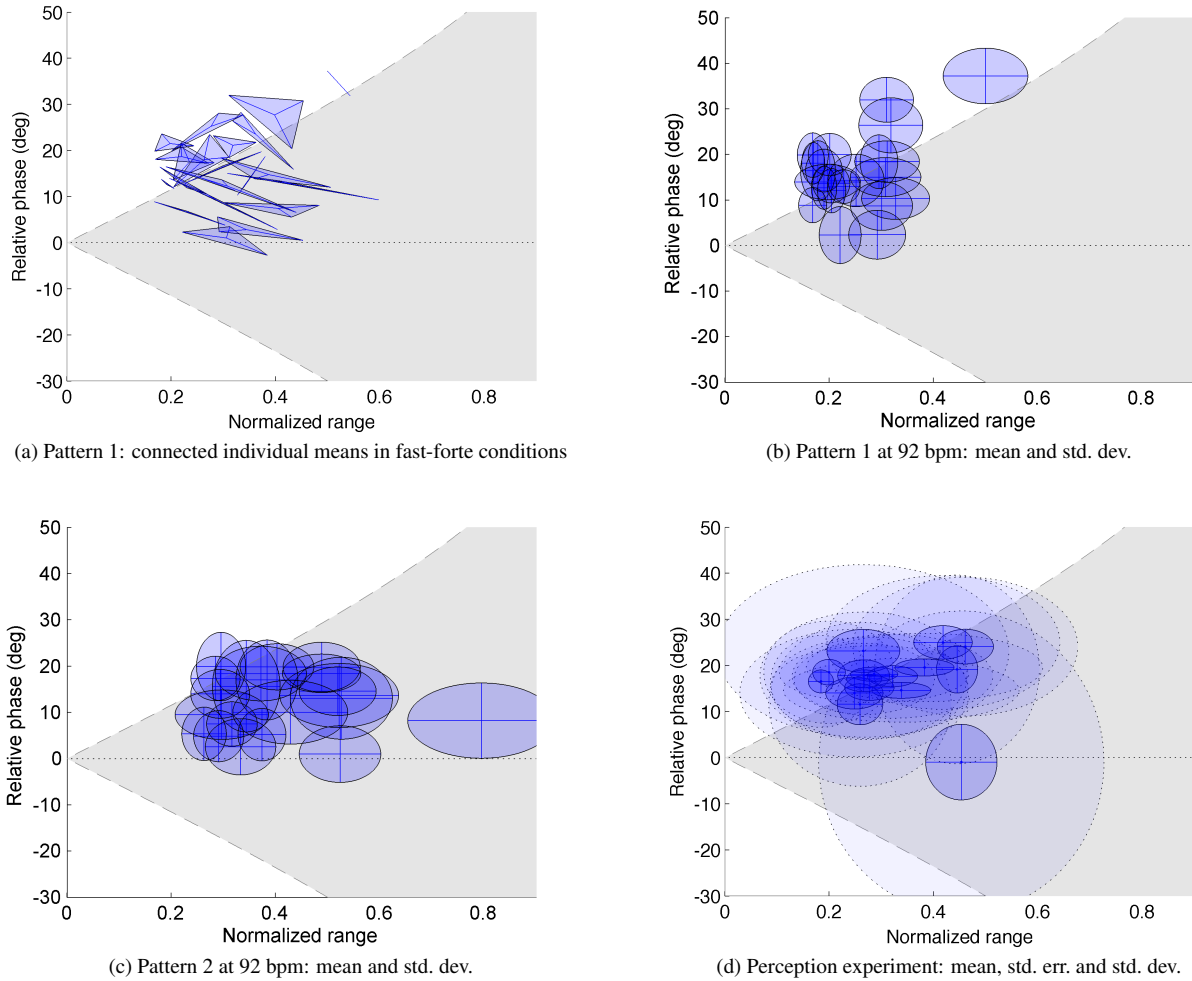


Figure 3. Visualization of between- and within-participant variability in (a) the perception and (b-d) the performance experiment. Individual variances are represented by the semi-transparent blue shapes: (a) connected individual means, (b-c) individual means and standard deviations in two dimensions, (d) individual means, standard errors and standard deviations (standard deviations are indicated very lightly to better distinguish individuals). The gray area indicates combinations of relative phase and normalized range at which bow changes take place within the string-crossing range (i.e. at the moment of the bow change the bow is in contact with both strings); above this area bow changes take place after the string-crossing has been completed. For more details, see the text.

as follows. For each bowing pattern an audio fragment was selected from the 92 bpm condition (the metronome click was not audible in the recording), yielding a total of 66 audio fragments (22 participants x 3 bowing patterns; a third bowing pattern is not discussed here). Five professional violin teachers participated in the rating task, which was implemented in a graphical user interface programmed in Java. The rating task consisted of three blocks, one for each bowing pattern. Each block was preceded by a short training block of five fragments. The order of the blocks and the order of the fragments within the blocks were randomized. The audio fragments were presented by headphones (a Sennheiser HD 595). They had a duration of about five seconds and were presented only once. After each presentation the raters were asked to judge the quality of the performance on three 6-point Likert scales, representing technical quality, sound quality and expression. Thus the obtained scores ranged from 0 to 5 from very bad

to very good (technical quality and sound quality) and not expressive to highly expressive (expression).

3. RESULTS

3.1 Variance within and between individuals

Figure 3 presents an overview of how individual participants used the coordination parameters in both the performance and the perception experiment. Each semi-transparent shape in the panels represents an individual participant. Since there was no obvious difference between the groups of different expertise all participants are included in the graphs without making distinction. Regions where there is a lot of overlap between individuals show up as darker. It should be noted that in case of performance (panels (a)-(c)) the points in the coordination space are based on average values of relative phase and normalized range at the bow changes (both up-to-down and down-

Group	Pattern 1	Pattern 2
Am.	2.6 ± 1.5 [1.4-4.0]	2.2 ± 1.3 [1.0-3.4]
Stud.	2.8 ± 1.2 [1.2-3.6]	2.6 ± 1.6 [1.0-4.2]
Prof.	2.4 ± 1.6 [0.6-4.2]	3.2 ± 1.2 [2.2-3.8]

Table 1. Expert ratings (from 0 to 5) for technical quality of performance ($M \pm SD$ [min-max]) per performer category.

to-up). Possible asymmetries due to inclination offset are ignored in this representation. This asymmetry was most notable in the performance of pattern 1 (panels (a) and (b)), in which changes from up to down bow mostly fell outside the string crossing range, whereas changes from down to up bow fell within the string-crossing range (see [9]).

In panel (a) the triangular shapes represent the connected means per individual of three comparable experimental conditions, namely fast-forte performances of pattern 1 (played at 92 and 112 bpm with metronome, and at preferred tempo without metronome). The shapes were rather small and individual shapes could be clearly distinguished. This indicates that the individual performers were quite consistent in their use of the coordination parameters, but that there was little agreement among the performers.

Panels (b) and (c) show the means and standard deviations across cycles per performer for two selected conditions (pattern 1 and 2, respectively, played at 92 bpm), giving an indication of the cycle-to-cycle consistency of individual performers. The standard deviations (indicated in two dimensions by the ellipsoids) were rather small, indicating a high stability of individual performance. Again, the individual shapes were distributed across a larger region of the parameters space, confirming that there was little agreement among the performers. In addition, there was an interesting difference between pattern 1 and 2. In pattern 2 (panel (c)) the standard deviations per individual were mostly larger, and there was notably less agreement between performers in their use of normalized range compared to pattern 1 (panel (b)). This reflects the practical experience that pattern 2 is more difficult to perform than pattern 1. A possible explanation proposed in [14] (page 81) is that the reversed direction of the string crossing relative to the bowing direction leads to a more complicated relationship of the respective motions of the bow and the arm parts.

For comparison, individual choices of the coordination parameters in the perception study are shown in panel (d). In this study the participants (with some exceptions not the same as the ones in the mocap experiment) could adjust the coordination parameters by a 2D slider. Thus, the choice of the parameters was based on the quality of the (synthesized) sound, and was not influenced by motor conditions. The mean values of the individual participants fell mostly in the same range of the coordination space as those in the performances, providing support for an acoustic origin of the coordination behavior as observed in performance. However, the standard deviations (light shaded areas in panel (d)) were notably larger than those of the

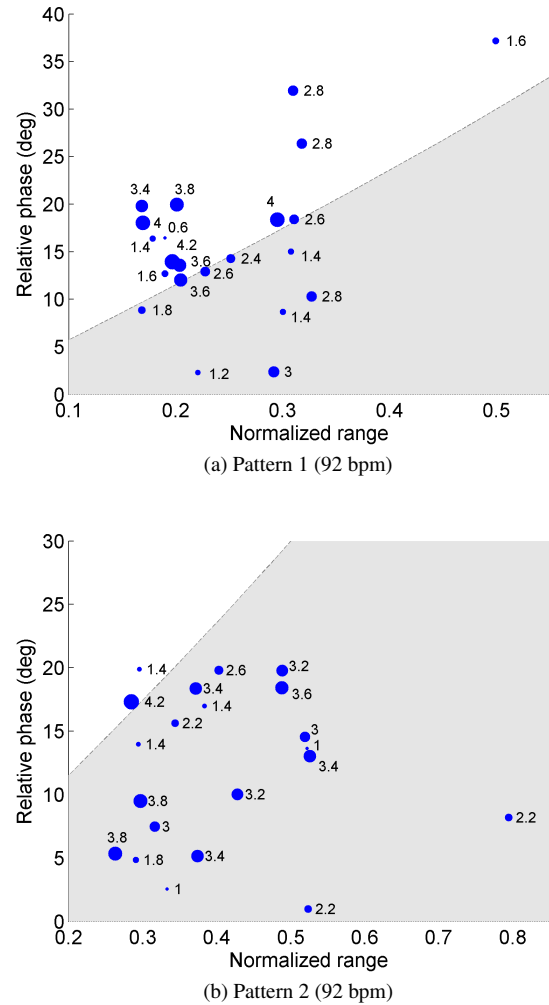


Figure 4. Expert ratings.

performances (shaded areas in panel (b) and (c)), blurring the individual differences. This indicates that the acoustic constraints in the perceptual experiment were not too strict, and cannot fully account for the small standard deviations observed in the performance experiment. Of course, it should be noted that direct comparisons between the results of the perception and the performance experiment have limited validity given the simplifications of the coordination model and the limitations of the sound synthesis in the perception experiment.

3.2 Expert ratings

As already mentioned in the introduction there was no clear distinction in coordination behavior between the three categories of performers (amateurs, students and professionals). It is however quite likely that there were differences in the quality of the individual performances. Even though several measures of performance expertise of the participants were obtained (cumulative practice hours as a function of age [15], and the “Ollen Musical Sophistication Index” [16]), it was decided not to rely on these for classification of the performers, but use expert ratings as a more specific measure of performance quality instead.

The inter-rater reliability among the five raters, estimated

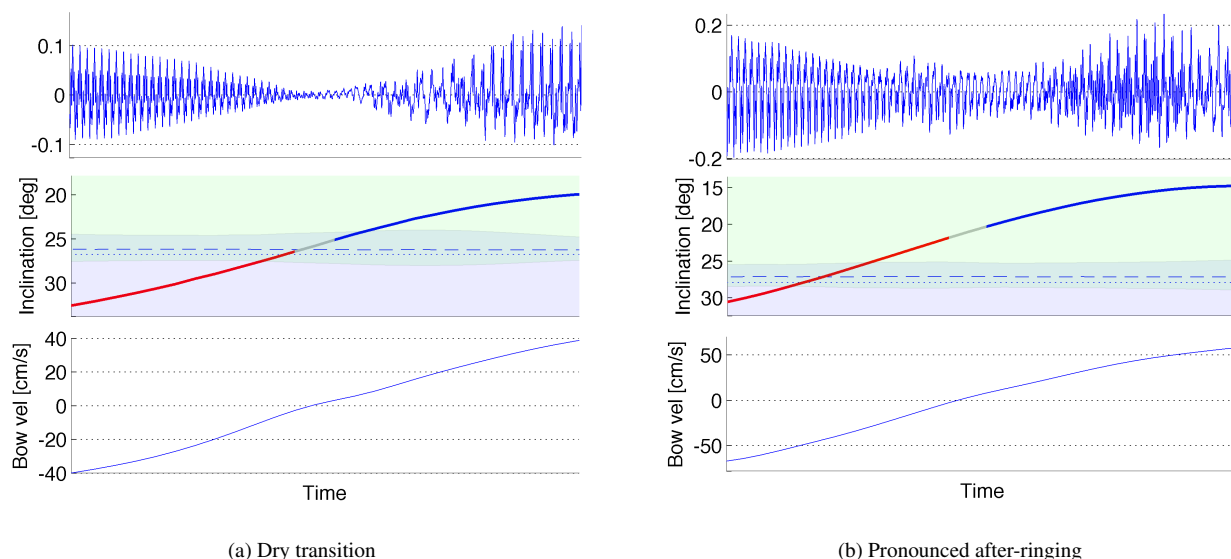


Figure 5. Note transition by two performers (pattern 1 at 92 bpm, transition from open E string to B played on the A string).

by taking intra-class correlations, was rather low, probably due to the difficulty of the rating task. For the combined ratings (technique, sound and expression) the ICC was 0.433 [$F(194,780)=4.81$, $p<0.001$], and for the technique rating alone the ICC was 0.442 [$F(64,260)=4.96$, $p<0.001$]. The correlations between the three rating scales was high ($r>0.91$). In the following only the technical quality rating is considered.

The technical quality ratings per performer category are shown in Tab. 1. Surprisingly, for pattern 1 the professionals were on average rated lower than amateurs and students. For pattern 2 (the more difficult pattern) the order of the group averages was according to expectation. Interestingly, there was a considerable overlap in the ratings between groups. Apparently, the performance quality of these particular note patterns is not so much dependent on the level of expertise as one might expect.

In order to figure out if there was a systematic relationship between performance quality and the used coordination parameters, expert ratings were mapped onto the two-dimensional coordination space. The results for patterns 1 and 2 (performed at 92 bpm) are shown in Fig. 4. Visual inspection does not reveal any trend between rating and the combination of relative phase and normalized range. In fact, in both patterns the highest and lowest ratings are located directly next to each other. The perceived quality of the performances in the mocap experiment seems therefore not to depend on the used combination of coordination parameters.

4. DISCUSSION

To summarize the above results it has been shown that individual performers are rather consistent in their coordination behavior, but that there are considerable inter-individual differences. The part of the coordination pa-

rameter space that is covered by the “ensemble” of performers is roughly similar to that in a complementary perceptual experiment with a virtual violin, where participants could adjust the coordination parameters by a simple slider. Furthermore, the expert ratings of the performances showed no systematic relation with the coordination parameter space, indicating that other factors probably were more important for the perceived quality of the performance. Taken together, it can be concluded that the acoustical constraints with regard to coordination of string crossings and bow changes are (within certain limits) not very strict.

The question remains how to explain the high consistency of individual performers. One possible explanation might be that it is due to motor constraints or habitual motor patterns. Motor constraints certainly play an important role in performer-instrument interaction and might, for example, to a large extent explain differences in coordination behavior between pattern 1 (the “normal” pattern) and pattern 2 (the “reversed” pattern). However, even though there might be individual differences in the physical build of the participating performers, it does not provide a very satisfactory explanation for the individual differences in coordination behavior.

An alternative explanation might be that the coordination parameters influence certain perceivable quality aspects of the note transitions, for example the quality of the attacks or the amount of after-ringing. In that case the differences between individual performers might be explained by differences in personal sound ideals or different expressive intentions. An example of the influence of the relative timing of the bow change relative to the string crossing on the wave form (in this case partly due to a considerable inclination offset) is shown in Fig. 5, where the one shown in panel (a) can be characterized as “dry”, whereas the one in (b) sounds more resonant. According to this idea the

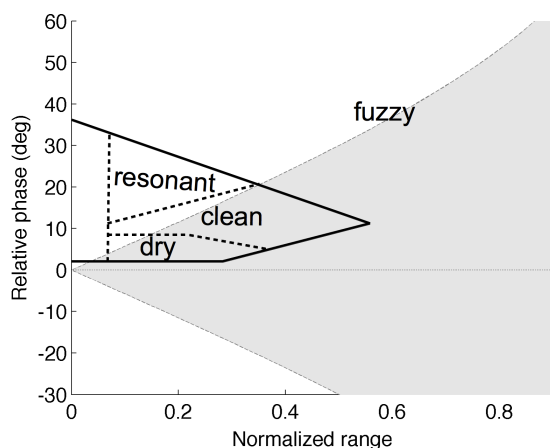


Figure 6. Tentative subdivision of the coordination parameter space in different quality regions.

coordination parameter space might be subdivided in different quality regions. A tentative subdivision is presented in Fig. 6, based on analysis of simulated complex bowing gestures [10] and subjective listening of performances with knowledge of the used coordination parameters.

5. CONCLUSIONS

The results of the perception and the performance studies suggest that coordination of bow changes and string crossings can be partly explained by acoustical constraints that should be respected for good quality note transitions. However, the consistent individual coordination in combination with the relatively large inter-individual differences suggests that the acoustical constraints are not very strict and that they allow for a certain degree of individual freedom. This is in agreement with earlier findings by Young, who showed that individual violinists could be distinguished by means of machine-learning techniques [17].

Based on the presented studies no definitive conclusions can be drawn regarding the possible reasons for the observed inter-individual differences. Further investigation is needed to find out in how far they might be related to motor constraints, habitual motor patterns, or personal preferences for sound quality. A related question is in how far individual performers can (or do) exert control of the coordination parameters in order to achieve a certain sound or expressive effect (e.g. stressing a melody line).

From a theoretical perspective the presented studies touch upon an interesting issue, namely that subtle aspects of sound quality might be integrated in the sensorimotor loop that is responsible for the formation of bowing patterns. This might not come as a surprise to most musicians, among whom it is common knowledge that (great) artists can be distinguished by their particular sound. The current study might give a glimpse into one of the many aspects of how musicians could achieve a personal sound.

Acknowledgments

Many thanks to Britta Westner for conducting the expert

rating study. The first author was supported by the Alexander von Humboldt foundation.

6. REFERENCES

- [1] P. Hodgson, *Motion study and violin bowing*. Urbana, Illinois: American String Teachers Association, June 1958, first published 1934 by J. H. Lavender & Co., London.
- [2] E. Schoonderwaldt, "Control and coordination of complex bowing patterns," in *Proceedings of the Second Vienna Talk*, University of Music and Performing Arts, Vienna, Austria, 2010, pp. 133–136.
- [3] E. Schoonderwaldt and E. Altenmüller, "Mastering the violin: Motor learning in complex bowing skills," in *Proceedings of the International Symposium on Performance Science*, A. Williamon, Ed. Toronto, ON, Canada: AEC, 2011, pp. 649–654.
- [4] E. Schoonderwaldt, M. Demoucron, E. Altenmüller, and M. Leman, "Auditory perception of note transitions in simulated complex bowing patterns," *J. Acoust. Soc. Am.*, vol. 133, no. 6, pp. 4311–4320, 2013.
- [5] H. Winold, E. Thelen, and B. D. Ulrich, "Coordination and control in the bow arm movements of highly skilled cellists," *Ecological Psychology*, vol. 6, no. 1, pp. 1–31, 1994.
- [6] N. Rasamimanana, "Geste instrumentale du violoniste en situation de jeu : Analyse et modélisation," Ph.D. dissertation, Université Pierre et Marie Curie (UPMC), Paris VI, 2008, english transl.: Violin player instrumental gesture: Analysis and modelling.
- [7] J. Verrel, W. Manselle, U. Lindenberger, and M. Woolacott, "Exploiting biomechanical degrees of freedom for fast and accurate changes in movement direction: coordination underlying quick bow reversals during continuous cello bowing," *Frontiers in Human Neuroscience*, vol. 7, p. 157, 2013.
- [8] O. Kazennikov and M. Wiesendanger, "Bimanual coordination of bowing and fingering in violinists-Effects of position changes and string changes," *Motor Control*, vol. 13, pp. 297–309, 2009.
- [9] E. Schoonderwaldt, M. Demoucron, E. Altenmüller, and M. Leman, "Perception and production of complex bowing patterns in violin performance," in *Proceedings of Meetings on Acoustics*, vol. 19, Montreal, 2013, p. 035013.
- [10] M. Demoucron, E. Schoonderwaldt, and M. Leman, "Acoustical features of complex bowing patterns in violin performance," in *Proceedings of SMAC/SMC 2013*, 2013.
- [11] E. Schoonderwaldt and M. Demoucron, "Extraction of bowing parameters from violin performance combining motion capture and sensors," *J. Acoust. Soc. Am.*, vol. 126, no. 5, pp. 2695–2708, 2009.

- [12] M. Demoucron, A. Askenfelt, and R. Caussé, “Measuring bow force in bowed string performance: Theory and implementation of a bow force sensor,” *Acta Acustica united with Acustica*, vol. 95, no. 4, pp. 718–732, 2009.
- [13] E. Schoonderwaldt and M. Demoucron, “An extended geometric model for analysis of string crossings in bowed-string instrument performance,” in *Proceedings of Acoustics 2012*, Nantes, France, 2012.
- [14] O. Szende and M. Nemessuri, *The physiology of violin playing*. London and Wellingborough: Collet’s (Publishers) Ltd., 1971.
- [15] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, “The role of deliberate practice in the acquisition of expert performance,” *Psychological Review*, vol. 100, pp. 363–406, 1993.
- [16] J. E. Ollen, “A criterion-related validity test of selected indicators of musical sophistication using expert ratings,” Ph.D. dissertation, Ohio State University, Columbus, OH, USA, 2006.
- [17] D. Young, “A methodology for investigation of bowed string performance through measurement of violin bowing technique,” Ph.D. dissertation, Massachusetts Institute of Technology, 2007.

CONCEPTUALIZATION OF VIOLIN QUALITY BY EXPERIENCED PERFORMERS

Charalampos Saitis^{1†} Claudia Fritz² Catherine Guastavino¹ Gary P Scavone¹

¹Center for Interdisciplinary Research in Music Media and Technology, McGill University

²Lutherie-Acoustique-Musique, Université Pierre et Marie Curie, UMR CNRS 7190

[†]charalampos.saitis@mail.mcgill.ca

ABSTRACT

This paper explores how violin quality is conceptualized as reflected in spontaneous verbal descriptions by experienced performers collected while playing in a perceptual evaluation experiment. Players were asked to rank the different violins in order of preference and to justify their ranking in free-format responses. The constant comparison analysis from grounded theory was employed to develop a classification scheme of concepts and the attributes that embody them. A quantitative analysis, based on the number of occurrences for each attribute and concept, provided a hierarchy of violin preference criteria/quality concepts: The conceptualization of violin quality encompasses the response of the violin to the various techniques and musical intentions in direct association with the quantity and quality of the produced sound as well as the emotions and values of the player.

1. INTRODUCTION

The prominent nineteenth century Italian cellist Alfredo Piatti once spoke of his Stradivarius cello (from *The Adventures of a Cello* by C. Prieto):

I have at times become enamoured at the sight of a *fine* instrument, have been impressed by its *beauty*, and when I have become its owner I have tried to believe that its tone equalled that of my Stradivari. Time, however, has invariably seen me return to my old friend with a feeling of *satisfaction* difficult to explain. True, the differences of tone between my Stradivari and other recognized *fine* instruments are subtle, but I can only say that I obtain from the former a *depth* and *nobility* of tone which ever affords me a sense of *contentment*; in fact, there is something *unattainable* elsewhere.

What is a “fine” violin? A long-standing goal of violin acoustics has been to identify which vibro-acoustical factors affect the timbre and feel of a particular instrument—for example, its perceived “depth,” thus distinguishing one

violin from another. Most previous research has traditionally attempted to answer this question through mechano-acoustical measurements and/or listening tests. Both approaches seem unsuitable for addressing the critical role of the violin player in determining the quality of an instrument. To this end, recent studies have focused attention to the perceptual and cognitive processes that take place when players assess violins in playing tests [1–6]. Of particular interest is the diverse vocabulary shared by musicians to describe the quality of a violin or its sound, as illustrated in Piatti’s own words, and how these verbalizations can be mapped to acoustical properties of the instrument.

As part of the VIOCADEAS project, a standardized qualitative violin evaluation procedure was proposed [7]. Frequently used English descriptions of violin sound were grouped according to different quality categories: across range (*evenness of tone, evenness of response, problem notes on each string*), overall (*loud, responds easily*), tonal qualities (*mellow vs. strong, gritty vs. smooth, harsh vs. warm, thin vs. deep, complex vs. one-dimensional, tight vs. open, fuzzy vs. clear, bright vs. dark*), and playing qualities (*transient behaviour, notes hard to play very softly or very loudly*). Each description was mapped to an acoustical or spectral property—for example, a *complex* sound “has many overtones and color.”

In another study, sixty-one common English adjective descriptions of desirable and undesirable violin tone qualities were collected and then arranged by violinists on a two-dimensional map, so that words with similar meanings lay close together, and those with different meanings lay far apart [1]. Multidimensional scaling demonstrated consistent use among performers of many words, and highlighted which words are used in similar situations. It was also observed that almost all verbal descriptions of violin sound incorporate an evaluative judgement as being either good or bad qualities. Further, three dimensions for the characterization of violin sound quality emerged (with acoustical and perceptual interpretations): *warm/rich/mellow vs. metallic/cold/harsh* (spectral balance, undesirable qualities associated with excessive high-frequency content or too little low-frequency content); *bright/responsive/lively vs. muted/dull/dead* (“amount of sound” produced by the instrument, particularly in the middle and upper ranges); and *even/soft/light vs. brash/rough/ raspy* (noisy character, i.e., width of distribution of spectral energy).

It is unclear whether the acoustical interpretations of verbal violin sound descriptions suggested in these studies are

reliable or generalizable, primarily because they are based only on a priori knowledge of the respective authors as opposed to emerging concepts grounded in the verbal data. Attempts to find relationships between measurable vibrational properties of violins and their perceived qualities first require a closer look into the ways violinists process and conceptualize the latter. To this end, a recent study examined the differences between preference judgements made by violin players in active playing vs. passive listening situations in conjunction with psycholinguistic analyses of free-format verbal French descriptions of the participants' experience [2]. Two distinct objects under evaluation for the violinist were identified: descriptions refer either to the sound of the violin (e.g., sound is *acide* or with *une certaine chaleur*) and/or to the instrument itself (e.g., the violin is *facile à jouer* or *très égal*). Results suggested that the influence of sound on the overall evaluation of a violin varies between playing and listening conditions. This seems to support the discussion that, concerning the perspective of the player, listening tests are probably not much indicative of the processes that take place when assessing the qualities of a violin; playing-based evaluations afford a higher level of ecological validity [3].

The present study aimed at identifying the different concepts and situations of violin quality relevant to the player and how they link to each other: what is meant by "depth" of tone and how this relates to the "beauty" of Piatti's cello. An open-ended questionnaire was given to experienced violin players during an experiment for the perceptual evaluation of violins. In the experiment, musicians were asked to play and rank a set of different violins in terms of preference and subsequently justify their choices through answering open-ended questions.

2. METHOD

2.1 Participants

Twenty skilled string players took part in this experiment (8 females, 12 males; average age = 34 yrs, SD = 13 yrs, range = 20–65 yrs; 11 native English speakers, 3 native French speakers, 6 other). They had at least 15 years of violin experience (average years of violin training = 26 yrs, SD = 12 yrs, range = 15–60 yrs; average hours of violin practice per week = 15 hrs, SD = 9 hrs, range = 9–30 hrs), owned violins with estimated prices ranging from less than \$1K to \$30K, and were paid for their participation. Thirteen participants described themselves as professional musicians, and 8 had higher-level degrees in music performance (MMus, MA, DMus, DMA). They reported playing a wide range of musical styles [classical (95%), folk (47%), baroque (37%), jazz/pop (10%), and contemporary (5%)] and in various types of ensembles [chamber music (70%), symphonic orchestra (70%), solo (55%), and folk/jazz band (40%)].

2.2 Preference ranking task

Participants freely played 8 violins of different make and age and ranked them from least to most preferred in 5 identical trials. Participants returned for a second, identical

session 3–7 days later. Violins of different periods were used, varying from student to performance level. Low light conditions and dark sunglasses were used to hide the identity of the instruments as much as possible. Considering the bow as an extension of the player, violinists carried out the task using their own bow (see [3] for a detailed discussion on the control of certain experimental conditions).

2.3 Questionnaire and procedure

Taking into account the lingual diversity of Quebec, we compiled a bilingual questionnaire in English and French and invited participants to respond in that language they felt most comfortable with. To avoid confining the responses into pre-existing categories, we formed very general, open-ended questions with input from an expert in the psycholinguistic evaluation of sound quality. The same questionnaire was used in both experimental sessions. Upon completing the first trial, participants provided spontaneous verbal (written) responses to the questions:

- QA1. *How and based on which criteria did you make your ranking? / Avec quels critères avez-vous effectué votre classement et de quelle façon les avez-vous utilisés ?*
- QA2. *Considering the violin that you ranked as "most preferred," can you say why? / A propos du violon que vous avez classé comme votre préféré : pourriez-vous nous dire pourquoi ?*
- QA3. *Considering the violin that you ranked as "least preferred," can you say why? / A propos du violon que vous avez classé en dernier : pourriez-vous nous dire pourquoi ?*

At the end of each subsequent trial, they were given the opportunity to modify their initial response if they so wanted. Upon completing the last trial, participants responded to the question:

- QB. *More generally, what is a very good violin for you? / En général, comment définissez-vous personnellement un très bon violon ?*

2.4 Analysis

All answers across the four questions were consolidated in a single data set as all questions were directly related to violin preference and quality descriptions. In each of the sessions, all participants answered questions QA1–QA3 in up to 4 trials as well as question QB (one time only). In total, 680 phrasings (34 phrasings per respondent on average, SD = 12) were extracted from the data. Of the phrasings, 61% came from professional musicians answers and 39% from amateur violin players answers. In total, 5 participants answered in French and we chose not to translate the phrasing extracted from their answer.

We used the constant comparison technique from grounded theory [8] to extract emergent concepts and attributes from the free-format verbal responses. Contrary to the typical approach of beginning with a hypothesis, grounded theory provides a systematic way of formulating a theory that is

THEME	CONCEPT	Property	Classification scheme	#	%
HANDLING	RESPONSE	<i>Ease</i>	ease of playing; liberty; flexibility; ease of response; playability	76	11
		<i>Speed & Articulation</i>	responsive; successive notes do not blend together; blurry; muddy; clarity; transients; articulates well; missing of the tuning; playability	47	7
	DESIGN & COMFORT		size; shape; weight; curvature; comfort; feel of the instrument: bulk, lightness	42	6
SOUND	CAPACITY	<i>Resonance</i>	resonant; ringing; vibrant; present; open; ample; muffled; éteint; tight; dormant; singing; muted; brilliance; brilliant; bright; nasillard; liveliness; sonority	74	11
		<i>Projection</i>	projection; ability of the sound to fill the room; ability of the sound to travel; to carry in a hall; focus; dry	41	6
		<i>Power & Volume</i>	powerful sound; a violin that carries a lot of sound; big; small; mince; weak; strong; thick; thin; petit	39	6
	TIMBRE	<i>Texture</i>	rough; raw; grossiere; soft; smooth; sweet; mellow; velvety; silky; golden; warm; cushioned; round; harsh; tinny; shrill; strident; stringy; acide	72	11
		<i>Richness</i>	rich; deep; hollow; has weight; flat; rich in/with a lot of harmonics/overtones; full, range/palette of colors/timbres; dark; complex; simple; colorless	63	9
		<i>Timbre-abstract</i>	tone quality; sound quality; timbre; color; color of sound; sound color	13	2
	CLARITY		pure; clean; direct; straightforward; no wolf tones; buzzing; scratches; whistles; (doesn't) speak well; blurry; muddy	48	7
	SOUND-GENERIC		toujours en écoutant le son du violon; avec le registre les plus bas, et le registre le plus haut; based on the sound	7	1
	BALANCE ACROSS STRINGS		well adjusted and balanced from G-string till E-string; the tone was very even over the range of the instrument; string differentials; consistency across the range of the instrument	55	8
RELEVANCE	AFFECTIVE REACTIONS		interesting; beautiful; fascinating; irritating; overbearing; pleasant; pleasing; fun to play; enjoyable	72	11
	MUSICAL & EMOTIVE POTENTIAL		can respond emotionally and dramatically to my playing; can do anything you want it to; does not require me to work too hard to overcome its personality but lets me play my own; possibility to vary my vibrato and bow pressure for my musical needs	40	6

Table 1. Classification scheme for the conceptualization of violin quality in player verbal descriptions. Number (#) and percentage (%) of occurrences across all four questions for each class are shown in the two rightmost columns.

grounded in data. One component of grounded theory is the constant comparison technique, whereby a theory is generated through contrasting emergent concepts at every level of analysis.

Linguistic devices constructed on the same stem (e.g., “rich,” “richness”) were grouped together. We also grouped together lexical devices that were semantically related (e.g., “balance” and “evenness”). To better illustrate the relationships between different concepts, we allowed the same phrasing to be coded into more than one categories (i.e., the derived concepts are not mutually exclusive).

3. RESULTS

The inductive analysis principle of grounded theory generates groupings starting from low levels to reach, a posteriori, more abstract themes. But we will instead discuss these themes from the more generic to the more specific for

the sake of argumentation. A typographic-style scheme is used to differentiate these different levels of categorization: highest-level themes are displayed in LARGE CAPITAL LETTERS; high-level concepts in SMALL CAPITAL LETTERS; and low-level properties in *Italics*.

At a first level of analysis, three underlying themes of evaluation emerged from the data: the HANDLING of the instrument, the produced SOUND, and the RELEVANCE to the player. A second level of analysis revealed eight concepts of violin quality, each situated within one of the three themes: {DESIGN & COMFORT, RESPONSE}, {TIMBRE, CAPACITY, CLARITY, SOUND-GENERIC}, and {AFFECTIVE REACTIONS, MUSICAL & EMOTIVE POTENTIAL} respectively. A ninth, autonomous concept also emerged: BALANCE ACROSS STRINGS. A third level of analysis led to a structure of properties for RESPONSE, TIMBRE, and CAPACITY: {*Ease*, *Speed & Articulation*},

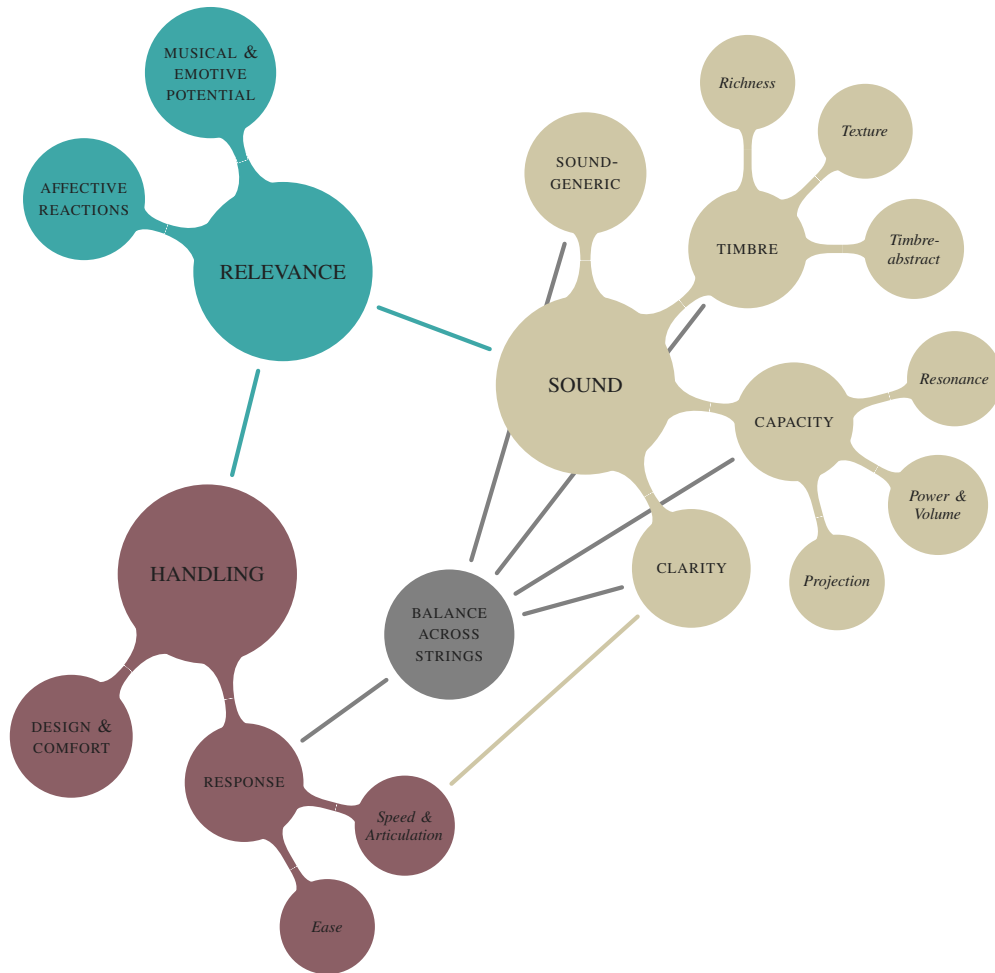


Figure 1. Concept map of emerging CONCEPTS, their *Properties* and underlying THEMES in player verbal descriptions of violin quality evaluation. The size of the circles corresponds to the different levels of categorization; lines indicate how different concepts link to each other (lengths are arbitrary).

{*Richness, Texture, Timbre-abstract*}, and {*Resonance, Power & Volume, Projection*} respectively. The classification scheme is outlined in Table 1. The emerged themes, concepts and properties, and how they link to each other are illustrated in Fig. 1. Definitions are described in the following section.

3.1 Concepts, properties and themes

HANDLING refers to the ergonomic aspects of the violin-musician system and relates to such concepts as responsiveness, comfort and flexibility of playing.

- DESIGN & COMFORT addresses how comfortable it feels to hold the instrument in relation to its size and curvature.
- RESPONSE describes how the instrument behaves when played, how it responds to the actions of the performer. We identified two properties: ease of response to different bowing gestures, and speed of response, which relates to note articulation.
 - *Ease* denotes how easy and flexible it is for the violinist to interact with the instrument and control the played sound.

- *Speed & Articulation* refers to how quickly and readily the violin responds to the different bowing techniques in terms of transients, dynamics and fast passages.

SOUND comprises descriptions about the quality, quantity and spatiality of the produced sound.

- TIMBRE specifies perceptual attributes of the violin sound related to harmonic content, in particular to spectral density and spread across registers.
 - *Richness* describes a certain quality of full-bodied sound (e.g., “full/fullness”) that appears related to harmonic density, particularly in the middle and low frequency regions of violin notes.
 - *Texture* pertains to descriptions of sound semantically associated with touch (e.g., “soft/softness”) and taste (e.g., “sweet/sweetness”), and is thus related to the perceived across-range spread of harmonics present in a played note. Similarly to the first dimension of violin quality identified in [1], undesirable qualities such as “strident” or “stringy” appear to be associated with excessive high-frequency content or too little low-frequency content.

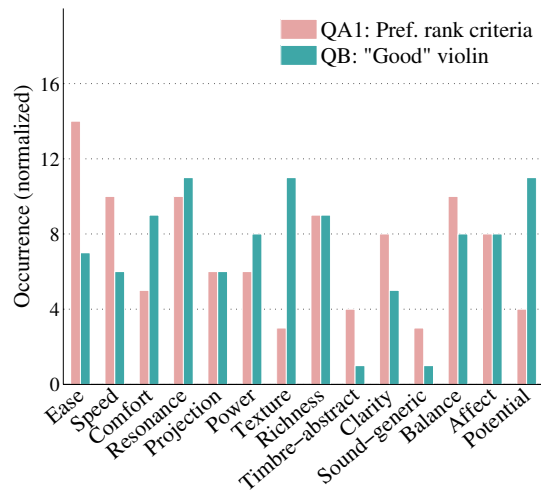


Figure 2. Comparison between preference ranking criteria and characteristics of the “very good” violin (normalized occurrence). Concepts are ordered as in Table 1.

- *Timbre-abstract* includes abstract allusions to the concept of timbre, such as “color” or “quality” of the sound.
- *CAPACITY* refers to descriptions of the instrument’s ability for substantial sound delivery: a sound that is present (i.e., it *resonates*), has *power* and *projects* well in the performance space.
- *Resonance* refers to the duration and quality of the sustained part of the sound. It is not related to the physical resonances of the violin body but rather to the perceived presence of a “ringing” sound.
- *Power & Volume* refers to the perceived intensity of the sound “under the ear.” It includes descriptions associated with the semantic field of size/volume (e.g., “big”).
- *Projection* relates to the performance space and concerns the quality and quantity of the played sound at different distances from the musician.
- *CLARITY* mainly refers to the presence of extraneous noise in the sound, such as wolf tones, “whistles” or “scratches.” In this context, “clear” or “clean” is used to describe a sound that is free from audible artifacts. We further identified a second situation, wherein *CLARITY* is used to describe articulation (i.e., successive notes do not blend together). Hence, the concepts of *CLARITY* and *RESPONSE* are linked via the latter’s *Speed* property.
- *SOUND-GENERIC* includes context-free references to the “sound” of the violin (i.e., it was not possible to identify associated concepts).

BALANCE ACROSS STRINGS describes the lack of pronounced differences in the response of the violin across the four strings (e.g., one or several strings being harder to play or slower to respond to varying gestures) as well as the quality of the produced sound across the different

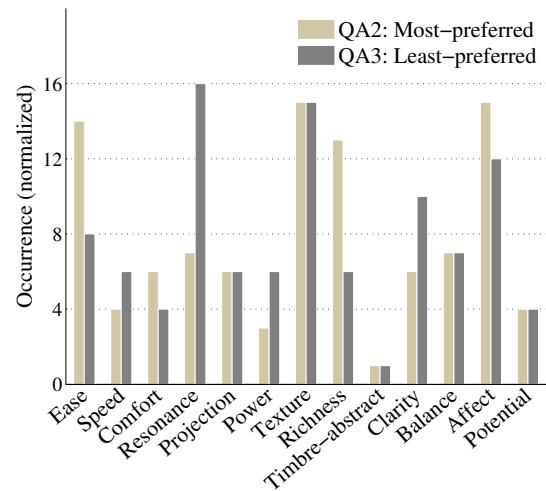


Figure 3. Comparing preference ranking criteria that determined the most- and least-preferred violins (normalized occurrence). Concepts are ordered as in Table 1.

registers (e.g., certain notes having too much or too little harmonic content or audible artifacts). It is therefore situated within both *HANDLING* and *SOUND* through *RESPONSE* and {*TIMBRE*, *CAPACITY*, *CLARITY*, *SOUND-GENERIC*} respectively.

RELEVANCE refers to quality judgements based on the musical, cultural and emotional involvement of the violinist.

- *AFFECTIVE REACTIONS* includes subjective, emotional and value responses to the sound of the violin as well as the playing experience [9].
- *MUSICAL & EMOTIVE POTENTIAL* denotes the ability of the violin to convey the musical and affective intentions of the player in varying situations.

3.2 Content analysis and discussion

The distribution of concepts was similar for the two experimental sessions, so we collapsed occurrences across sessions in Table 1. Note that results are reported in terms of number of occurrences of individual phrasings rather than percentages across the respondents as one original response can include several phrasings coded into the same or different concepts. For the conceptualization of the characteristics of the “good” violin, 22% of the phrasings refer to the *CAPACITY* of the instrument for substantial sound production, 22% to the *TIMBRE* of the played sound, 18% to the *RESPONSE* of the violin to the actions of the player, 11% to *AFFECTIVE REACTIONS* of the violinist to the produced sound and playing experience, 8% to the *BALANCE ACROSS STRINGS* of response and sound quality, 7% to *CLARITY* in the played note, 6% to the *DESIGN* of the instrument and thus the *COMFORT* of playing, and 6% to the *MUSICAL & EMOTIVE POTENTIAL* of the violin in performance and personal contexts.

We compared questions QA1, whereby violin preference was described in direct relevance to the experimental setting

(i.e., preference rankings of given instruments), and QB, whereby respondents provided context-free descriptions of violin quality (see Fig. 2). Whereas RESPONSE prevailed when violinists described their preference ranking criteria, it was considerably less present in the general descriptions of the “good” violin. Similarly, violinists called upon *Texture*, POTENTIAL and, to a lesser extent, COMFORT more often in question QB than in QA1. It thus appears plausible that context, or the lack thereof, influences the level of abstraction in the conceptualization of violin quality.

The proportions of concepts within the different discriminating situations of describing the most- vs. the least- preferred violin (from the answers to questions QA2 vs. QA3 respectively) are contrasted in Fig. 3 (SOUND-GENERIC is excluded as we found no related phrasings in the responses to either question). A possible explanation for the differences in the distribution of concepts between the most- and least-preferred violin descriptions is that violin players use different verbalizations to describe instrument qualities they prefer from those they do not.

4. CONCLUSIONS

When evaluating a violin or its sound, musicians call upon a wide diversity of linguistic forms (e.g., nouns, adjectives, expressions, metaphors, etc.) to describe the perceptual qualities of the sound or the instrument. Notably, no previous study has investigated how violinists conceptualize these perceptual qualities. From verbal responses of experienced violinists collected in a preference ranking experiment for the perceptual evaluation of violins, a classification scheme emerged that illustrates the complex links between the different player-typical concepts (e.g., RESPONSE, CLARITY, BALANCE), properties (e.g., *Ease*, *Richness*, *Projection*), and underlying themes (HANDLING, SOUND and their RELEVANCE to the individual).

The analysis of the verbal data identified two distinct objects under evaluation for the violin player: the sound of the violin and the instrument itself. This confirms previous findings in violin quality evaluation [2]. To describe the timbre of a particular violin, violinists appear to focus on spectral density (conceptualized in the perceptual attribute of *richness*) and spread (conceptualized in the perceptual attribute of *texture*) across the low, middle and high registers. However, a “good” sound is dependent on the amount of effort required to obtain it, with different musical or subjective situations leading to different degrees of compromise between sound quality and playability. This is illustrated in the following response by one of the participants: “A good violin for me is one that combines an even, resonant, singing tone with good sound production. I often play fiddle and rock music, and although a good sound is always important, I also need to be able to play loudly.”

“Rich” (or “richness”) was the most frequently quoted description of sound in the data, indicating a strong, widely-shared concept of violin quality. This observation is in agreement with results from a previous, more rudimentary analysis of the verbal responses to answer QA1 [3]. In fact, an analogy may be drawn between the importance of richness in violin sound quality and that of brightness

in brass instrument sound quality. We are currently analyzing a different set of violin player verbalizations using a linguistic approach to tease apart the different semantic interpretations of richness [10].

Acknowledgments

This project has been partially funded by the Centre for Interdisciplinary Research in Music Media and Technology and the Natural Sciences and Engineering Research Council of Canada. We thank Danièle Dubois for constructive comments and fruitful discussions.

5. REFERENCES

- [1] C. Fritz, A. F. Blackwell, I. Cross, J. Woodhouse, and B. C. J. Moore, “Exploring violin sound quality: Investigating English timbre descriptors and correlating resynthesized acoustical modifications with perceptual properties,” *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp. 783–794, 2012.
- [2] C. Fritz, A. Muslewski, and D. Dubois, “A situated and cognitive approach of violin quality,” in *Proc. 2010 Int. Symp. Music. Acoust.*, Sydney and Katoomba, Australia, 2010.
- [3] C. Saitis, B. L. Giordano, C. Fritz, and G. P. Scavone, “Perceptual evaluation of violins: A quantitative analysis of preference judgements by experienced players,” *J. Acoust. Soc. Am.*, vol. 123, no. 6, pp. 4002–4012, 2012.
- [4] C. Saitis, C. Fritz, C. Guastavino, B. L. Giordano, and G. P. Scavone, “Investigating consistency in verbal descriptions of violin preference by experienced players,” in *Proc. 12th Int. Conf. Music Percept. Cogn.*, Thessaloniki, Greece, 2012.
- [5] I. Wollman, C. Fritz, and S. McAdams, “The role of auditory and tactile modalities in violin quality evaluation,” in *Proc. 12th Int. Conf. Music Percept. Cogn.*, Thessaloniki, Greece, 2012, p. 1135.
- [6] C. Fritz, J. Curtin, J. Poitevineau, P. Morrel-Samuels, and F.-C. Tao, “Player preferences among new and old violins,” *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 3, pp. 760–763, 2012.
- [7] G. Bissinger and F. Gearhart, “A standardized qualitative violin evaluation procedure?” *Catgut Acoust. Soc. J. (Series II)*, vol. 3, no. 6, pp. 44–45, 1998.
- [8] B. G. Glaser and A. L. Strauss, *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine Publishing, 1967.
- [9] D. J. Hargreaves and A. M. Colman, “The dimensions of aesthetic reactions to music,” *Psychology of Music*, vol. 9, no. 1, pp. 15–20, 1981.
- [10] P. Cheminée, “Est-ce bien *clair* ? Stabilité, instabilité et polysémie d’une forme lexicale en contexte,” in *Le sentir et le dire: Concepts et méthodes en psychologie et linguistique cognitives*, D. Dubois, Ed. Paris: L’Harmattan, 2009.

THE INFLUENCE OF PLATE ARCHING AND THICKNESS ON THE SECOND AND FIFTH MODE ON VIOLIN TOPS

Mats Tinnsten

Dept of Engineering and Sustainable Development
Mid Sweden University
mats.tinnsten@miun.se

Peter Carlsson

Dept of Engineering and Sustainable Development
Mid Sweden University
peter.carlsson@miun.se

ABSTRACT

The objective of the work presented is to numerically investigate the influence of different arching and thickness on the resonance frequencies and mode shapes of the top plate. Here the focus has been on the second and fifth resonance frequency, the so called X and ring mode. The use of numerical analyses makes it possible to vary only one parameter while all other are held constant. The geometrical model was generated by use of the ideas of Robert Zuger and the general 3D-modelling software Rhinoceros. For each arch-height value, a new geometric model has been generated and exported to the finite element program Abaqus. The geometric model was generated in the same systematic way in all analyses. The results show surprising linearity both when varying the thickness and the arching.

1. INTRODUCTION

The top is considered to be one of the most important parts of the violin when it comes to sound production and it has been extensively studied both experimentally, *e.g.* [1, 2], and numerically, *e.g.* [3, 4]. It is possible to change the vibration properties, such as mode shapes and resonance frequencies, and the characteristics of the sound emanating from a vibrating structure by changing structural design variables such as geometric dimensions, shell thickness, and material parameters [5]. Of course, changes to one or more of these variables will result in changes to other structural characteristics.

A lot of work is spent on the top in order to make it perfect. But in addition to the practical experience of violin making, numerical calculations might also give hints on how to make good tops from carefully chosen blanks. In this study, the influence of arching and thickness for the X and ring mode on the violin top was studied in particular (*i.e.* mode 2 and 5).

2. NUMERICAL MODEL

The numerical model of the spruce top had a total length of 366 mm. A top with an arch height of 14 mm and a thickness of 3.2 mm had a weight of 87.4 g (with a density of $\rho = 482 \text{ kg/m}^3$). The geometry was generated according to the ideas of Robert Zuger [6], a model that is well adapted for transformation to a computer model. When studying the influence of arching, a new geometric model had to be generated for each height of the arch. Every model was generated in the same systematic way, the only difference being arch height. When studying the influence of thickness the same geometric model was used in all analyses, the only difference being shell thickness.

2.1 Shell model in Rhino

The top was modeled with and without *f*-holes with good precision in the general 3D-modeling program Rhinoceros, see Figure 1. Several tops with varying arch height were developed from the same base geometry. Smooth transitions have been used from the flat boards to the curved inner parts.

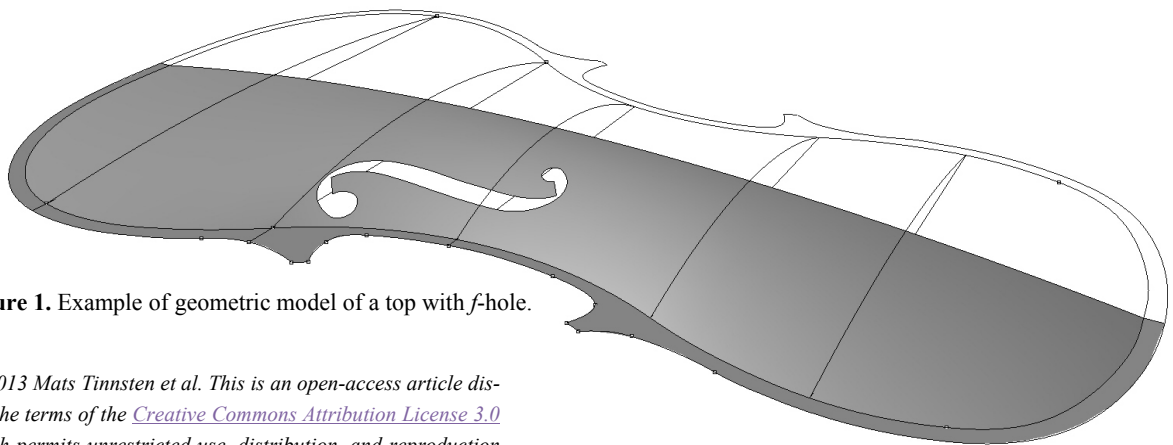


Figure 1. Example of geometric model of a top with *f*-hole.

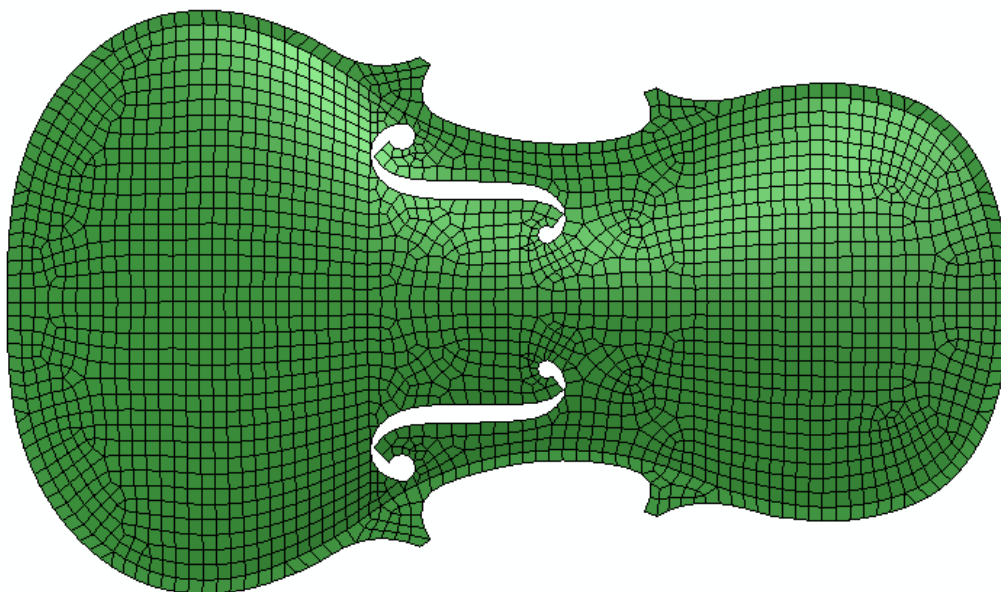


Figure 2. Example of finite element model of the top.

2.2 FEM model

The Rhino model was imported to the finite element program Abaqus, meshed as a half top and mirrored to a whole, symmetrical shell model of the top (without bass bar). The wood material of the top was spread equally on both sides of the shell with a constant thickness all over the shell. The top was meshed with 2454 mainly linear quadrilateral shell elements, see Figure 2. The dense mesh used in the FEM model gives reliable values for frequencies and mode shapes up to several thousand Hz, *i.e.* far beyond the results presented here. Lanzos solver was used for the modal analysis and boundary conditions were set to free-free. A disadvantage with the chosen shell model of the top, compared to a solid model, was that the fiber direction of the wood follows the curvature of the shell.

2.3 Material model

The wood material was modeled as an anisotropic material with main directions following along (1) and across (2) the centre line of the top. Chosen values for the Young's modulus were $E_1 = 14.5$ and $E_2 = 1.0$ GPa. The shear modulus was set to $G_{12} = G_{13} = G_{23} = 0.7$ GPa, Poissons ratio $\nu = 0.019$ and density $\rho = 482$ kg/m³. All according to reasonable values for common spruce used for violin tops.

3. NUMERICAL CALCULATIONS

Calculations were performed on tops with varying arching and thickness. Resonance frequencies and mode shapes for the first six modes were studied, but in this paper, the influence of arching and thickness on these properties is only presented for mode 2 and 5.

3.1 Varying arching

In order to study the influence of arching on resonance frequencies and mode shapes, tops with base geometry from Zuger were built up from 10 mm arching up to 18 mm in steps of 1 mm. Density and other material parameters were held constant during all calculations and the thickness was set to 3.2 mm.

3.2 Varying thickness

To study the influence of thickness variations on resonance frequencies and mode shapes, tops with base geometry from Zuger were modeled with a constant arching of 13 mm and with thicknesses from 1.8 to 3.8 mm. Density and other material parameters were held constant during all calculations.

4. RESULTS

Results from the numerical calculations are shown in Figure 3 and 4. In the diagrams, relations between resonance frequencies and arching and thickness are presented together with mode shapes for all calculated combinations.

The small pictures of mode shapes are placed directly above or beneath the corresponding resonance frequency point in both diagrams. In the mode figures, areas with blue color have no movements, areas with green or yellow color moderate movements, and areas with red color great movements. A linear connection for mode 5 is given in both diagrams.

All mode shapes are in good agreement with observed shapes in practical experiments. Also the calculated frequencies are in reasonable order of magnitudes.

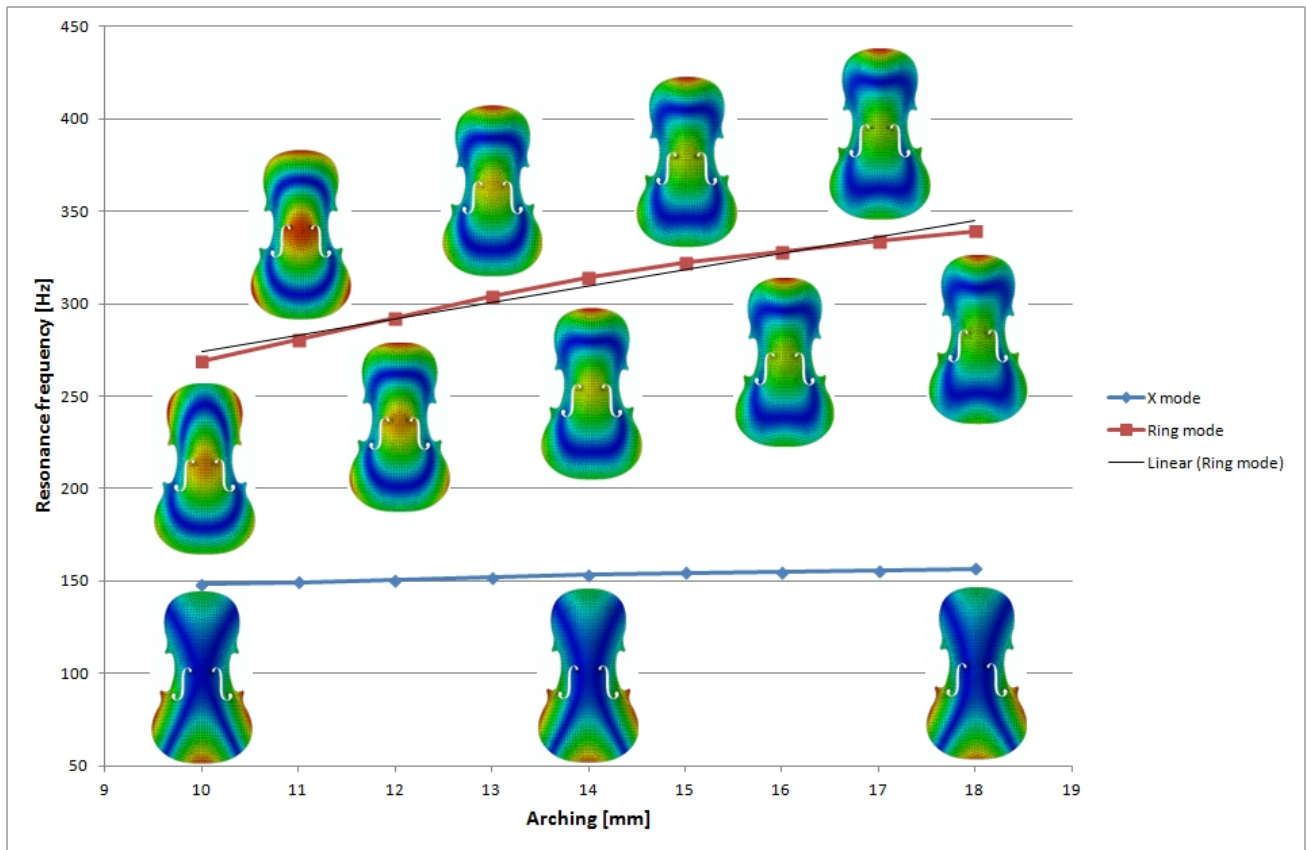


Figure 3. Relations between arching and resonance frequency for mode 2 (X mode) and 5 (ring mode). Pictures showing mode shapes are included in the same diagram.

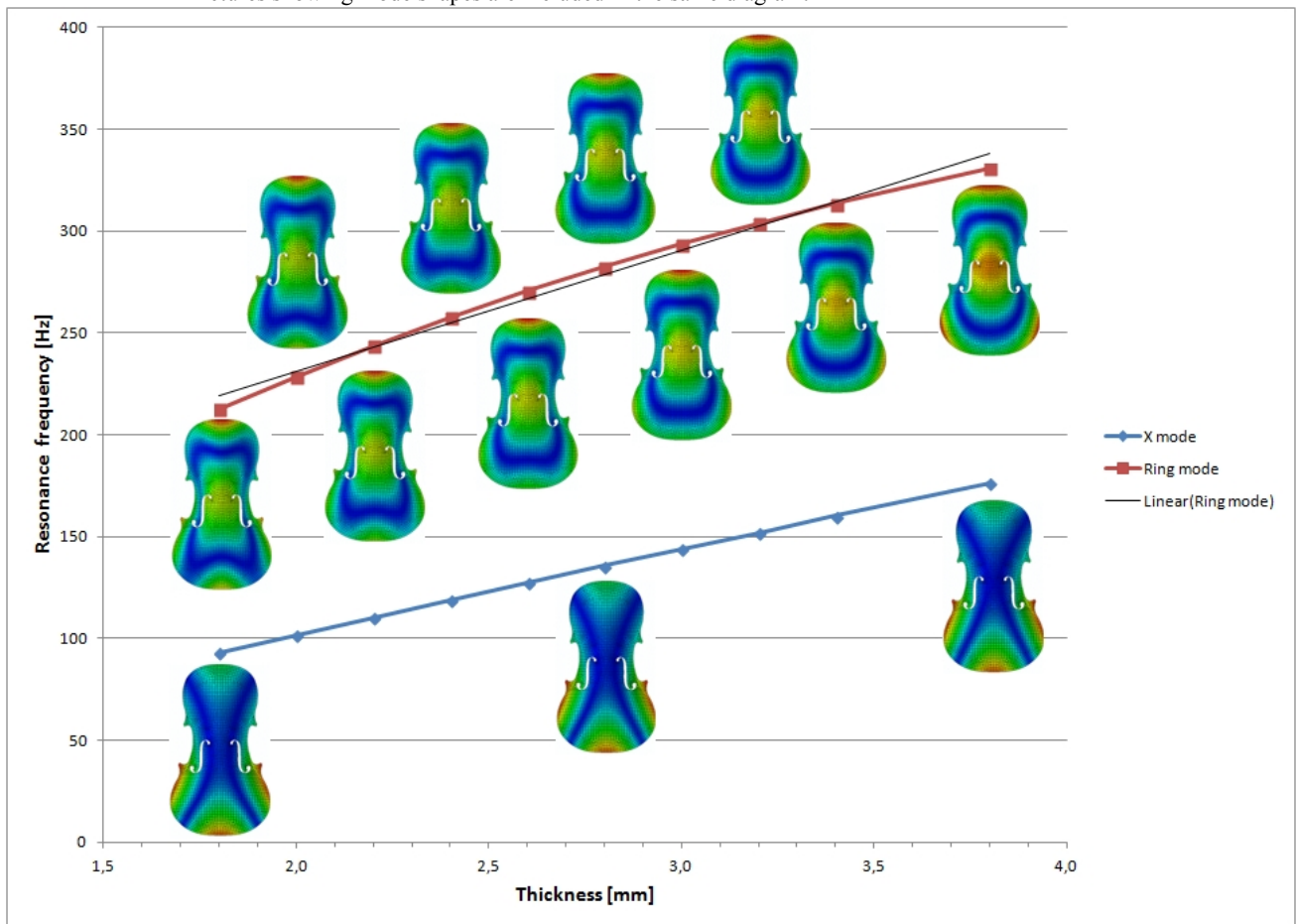


Figure 4. Relations between thickness and resonance frequency for mode 2 (X mode) and 5 (ring mode). Pictures showing mode shapes are included in the same diagram.

5. CONCLUSIONS

The relations between resonance frequencies and arching and thickness are almost linear for both mode 2 and 5. Changes in arching, with constant thickness, results in a raise of the resonance frequency with increasing arch height for mode 5. For mode 2, there are almost no changes in the resonance frequency within the same arching region.

An analogous change in thickness, this time with constant arching, has a different effect. The resonance frequencies for both mode 2 and 5 increases linearly with thicker tops and the gradient of the lines are almost the same.

A conclusion of the above described effects is that if a certain relation between the resonance frequencies of mode 2 and 5 (for example that they should be tuned in octaves) is desired, it is more important to start with a proper arching of the top than a suitable thickness.

According to the movement pattern of mode 2 and 5, it seems likely that mode 2 is mostly dependent of the stiffness of the top across the fibers and mode 5 along the fibers. This is probably why mode 2 is almost independent of the arching as long as the thickness is constant. When it comes to tops with constant arching, both mode 2 and 5 will have increased stiffness with greater thickness since this affects the stiffness in both directions.

An interesting observation is that both mode 2 and 5 have red areas near the tail gut and near the neck respectively. A consequence of this is that it is possible to tune the resonance frequencies 2 and 5 for the free top plate in octaves. Furthermore, since the top has great movements on both ends of the strings at these frequencies, the sound producing bridge will in its turn be strongly influenced by these movements. Here red areas are only observed in mode 2 and 5 for relatively low resonance frequencies on the top.

Acknowledgments

This work was partly funded by Sparbanksstiftelsen, Östersund, Sweden.

6. REFERENCES

- [1] A. Runnemalm, and N. E. Molin, "On operating deflection shapes of the violin body including in-plane motions," *J. Acoust. Soc. Am.*, 107, pp. 3452-3459, 2000.
- [2] E. Jansson, N. E. Molin, and H. Sundin, "Resonances of a violin body studied by hologram interferometry and acoustical methods," *Phys. Scr.*, 2, pp. 243-256, 1970.
- [3] N. E. Molin, M. Tinnsten, U. Wiklund, and E. Jansson, "A violinmaker's practical test of wood properties suggested from FEM-analysis of an orthotropic shell," *J. Catgut Acoust. Soc.*, 46, pp. 24-29, 1984.
- [4] J. Bretos Linaza, C. Santamaria, and J. Alonso Moral, "Vibrational patterns and frequency responses of the free plates and box of a violin obtained by finite element analysis," *J. Acoust. Soc. Am.*, 105, pp. 1942-1950, 1999.
- [5] M. Tinnsten, and P. Carlsson, "Numerical optimization of violin top plates," *Acta acustica united with Acustica*, 88, pp. 278-285, 2002.
- [6] www.zuger.se

PROBABILISTIC MODELING OF BOWING GESTURES FOR GESTURE-BASED VIOLIN SOUND SYNTHESIS

Akshaya Thippur¹

Anders Askenfelt²

Hedvig Kjellström¹

¹Computer Vision and Active Perception Lab, KTH, Stockholm, Sweden akshaya, hedvig@kth.se

²Department of Speech, Music and Hearing, KTH, Stockholm, Sweden andersa@speech.kth.se

ABSTRACT

We present a probabilistic approach to modeling violin bowing gestures, for the purpose of synthesizing violin sound from a musical score. The gesture models are based on Gaussian processes, a principled probabilistic framework. Models for bow velocity, bow-bridge distance and bow force during a stroke are learned from training data of recorded bowing motion. From the models of bow motion during a stroke, slightly novel bow motion can be synthesized, varying in a random manner along the main modes of variation learned from the data. Such synthesized bow strokes can be stitched together to form a continuous bowing motion, which can drive a physical violin model, producing naturalistic violin sound. Listening tests show that the sound produced from the synthetic bowing motion is perceived as very similar to sound produced from real bowing motion, recorded with motion capture. Even more importantly, the Gaussian process framework allows modeling short and long range temporal dependencies, as well as learning latent style parameters from the training data in an unsupervised manner.

1. INTRODUCTION

The aim of the current study is to develop natural sounding violin sound synthesis, which includes the characteristics of human performance. Our thesis is that this is accomplished by modeling the music-production process as accurately as possible: The player reads the musical score and interprets the piece as a sequence of events linked by the musical structure. The interpretation involves planning a sequence of control gestures, each producing a single note or a short sequence of notes.

Two aspects on sequences of sound-producing gestures can be noted.

- I) The exact shape of the control gestures depend on the individual interpretation of the musician, based on the knowledge of the style of the composition. It follows that it is desirable to be able to control performance style in a synthesized performance (e.g., from baroque to romantic violin playing style).

- II) Each control gesture, corresponding to a single note or a short sequence of notes, depends on a range of other gestures preceding and following after the current gesture.

Both these aspects are captured by a probabilistic framework which can represent a set of generic bow motion gestures which together define a performance of a piece of music as well as important modes of variation. This is further discussed in Sec. 2.

As a basis for this framework, we propose to use Gaussian processes (GP) [1], see Sec. 3. In this paper we present a pre-study where GPs are trained with recorded bow motions in the same manner as Bezier curves in related work [2–4]. The results indicate the GPs have the same descriptive power as the Bezier curves. A listening test presented in Sec. 4 shows that the violin sound produced from synthetic bow motions is perceived as very similar to the sound produced from real bow motion, recorded with motion capture.

Furthermore, we suggest in Sec. 5 that GP provides a solid mathematical framework for addressing the issues of individual performances and the style of playing in a principled manner. Our thesis is that a GP framework will make the best use of recorded bow motion gestures. Such dependencies reflect variations in player interpretation and modes of performance based on composition style.

2. BACKGROUND

Bow motion studies. Recently, Demoucron and Schoonderwaldt have studied bow motion in violin performance using motion capture methods and a robust method for sensing bow force [5–7]. Their work resulted in a large database of calibrated motion capture measurements of main and secondary bowing parameters of the performances of professional violinists (see Figure 1). Major results of their analyses were to establish the bow control strategies that players use when changing the dynamic level and timbre, and playing on different strings [8, 9].

Demoucron also developed a parametric description of bouncing bowing patterns (spiccato) based on the recorded data. Bow control data generated by his analytical models were used to drive a physical model of the bowed string [5]. The results showed that the realism of the synthesis increases significantly when the variation in control parameters reflect real violin performance. Again the coordination of the variations in the control parameters is a key to realistic violin sound. The conclusion drawn was that modeling

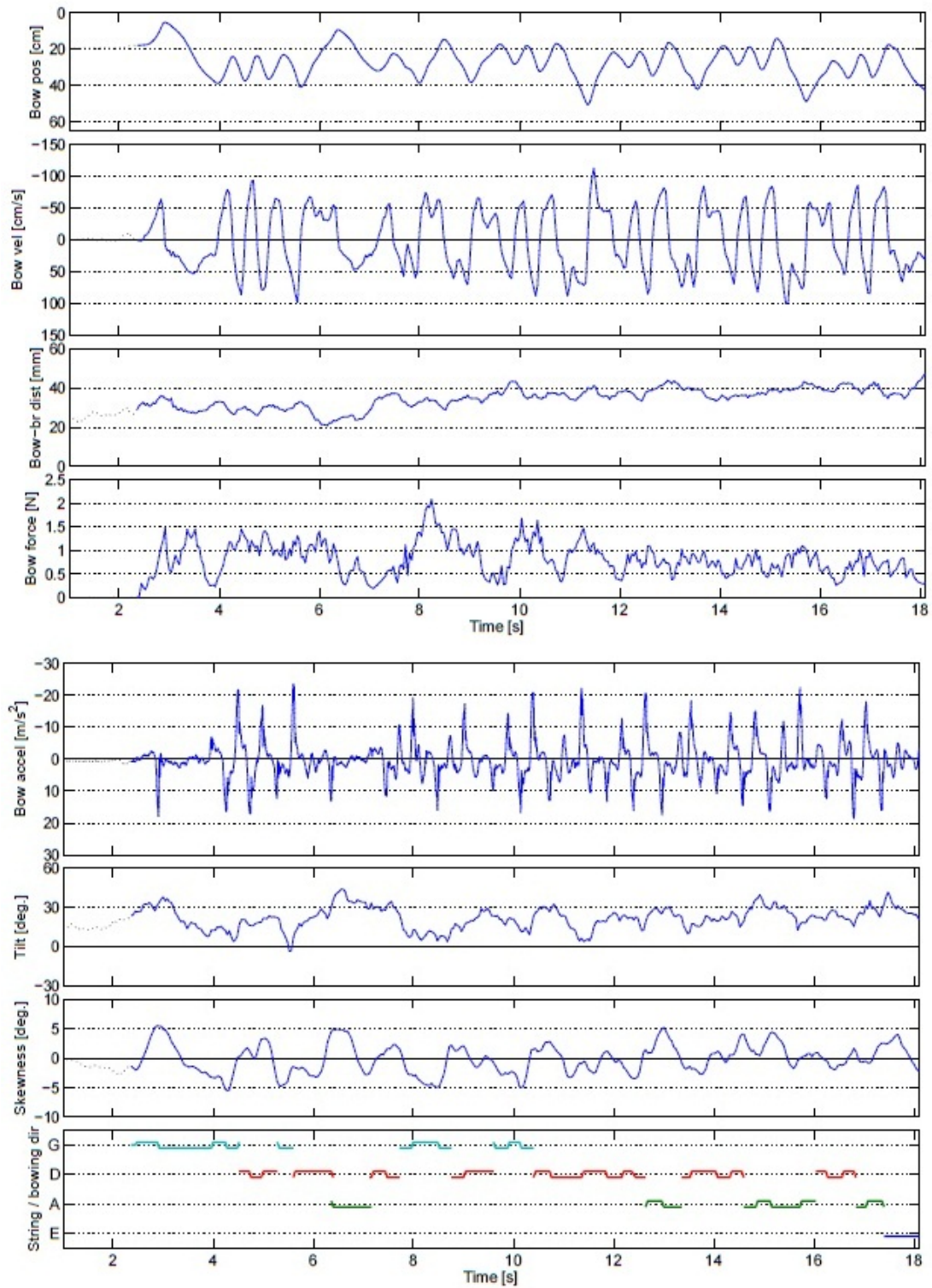


Figure 1. Example of a full set of bow control gestures for the first 18 notes of a Bach partita. From top: Transversal bow position, bow velocity, bow-bridge distance, bow force, bow acceleration, tilt, inclination, skewness, string played/bowing direction. From [6].

instrument control parameters is as important as modeling the instrument itself.

Violin synthesis from score information. A next step is using the acquired knowledge to learn computer models that produce violin sound (or rather, violin bowing gestures) given a musical score. Two approaches can be found in the literature: Reinforcement learning, where the computer model learns to perform bowing gestures (i.e., produce time sequences of bow motion parameters) under supervision of a human teacher, and supervised learning, where the computer model is trained with a set of recorded gestures, correlated with a musical score.

A reinforcement learning approach has been reported recently [10], where a generative model of bow motion is trained much in the same way as children learn to play according to the Suzuki school: The bow velocity and bow-bridge distance are preset using plain score information, while the range of bow forces producing a successful sound is learned using discriminative classifiers with human feedback judging the tone quality. Although a very interesting approach – the model can after four hours of training play like a Suzuki student with one year of experience – this is not a feasible approach to professional level violin sound synthesis.

A more time-efficient alternative is thus to directly show the model successful gesture examples, using a supervised learning approach. A recent, very ambitious collection of works from the Music Technology Group at Universitat Pompeu Fabra, Barcelona, deals with the task of automatically generating naturally-sounding violin performances from an annotated score [2–4]. Their approach is to retrieve samples of control gesture parameters from a database and concatenate them according to the score including instrument-specific annotations. The database is obtained from motion capture recordings of real violin performances, which have been segmented and classified into different groups for single notes with specific features (bowing style, dynamic level, context). All bow control gesture samples are parametrized using Bezier curve segments. For synthesis, the task of selecting the proper gesture sample in the database for each note in the score is based on an elaborated cost function which takes into account the deviations in duration and dynamic level between the stored samples and the specifications for the note in the score. The selected samples are stretched and concatenated and used to drive a simple physical model of the violin.

The obtained degree of realism in sound demonstrates the potential of gesture control of violin synthesis – it is indeed possible to simulate this extremely complex dynamical process. In the work presented here, we continue this path, using a supervised learning approach.

However, two aspects remain unaddressed in the Barcelona work, corresponding to aspects *I* and *II* discussed above.

- I*) No means exist to steer the performance style during synthesis. The grouping of the gesture examples according to bowing style (legato, staccato etc.), dynamic level, and context give some possibility of style control, but more ubiquitous style variations

(e.g., baroque vs. romantic playing style) are not captured – the model simply generates the “mean” performance in the database, given a certain score. This is however possible to accommodate in a probabilistic framework, such as the one proposed in this paper. The GPs capture the entirety of the numerous training curves comprehensively. It not only captures the average curve shape but also captures the different modes of variation in the training set. From a trained GP, slightly novel instances of synthetic bow motion can be generated, preserving the general shape and variances in the training set.

- II*) No means exist to model long-term gesture correlations. The curves are stitched together so that a coherent and physically plausible motion is generated. However, there is no way of selecting curves depending on gestures more than one time-step away. This is however possible to accommodate by adding extensions to our proposed framework. This is further discussed in Sec. 5.

3. MODELING BOWING GESTURES

Figure 2 gives an overview of the sound generating process. The score is viewed as a sequence of notes, each belonging to a note class defined by note value (duration), pitch, dynamic level, articulation, and bowing style. The strategy is to transform the notes in the score to a sequence of generic bow control gestures, each representing one note. The control gestures are then concatenated and used to drive a physical model of the bowed violin.

Gesture modeling using Gaussian processes. We model mapping 2 in Figure 2 using GPs in a manner similar to how Maestre et al. [3] use Bezier curves. The added value of the GPs is that not only the mean curves are captured by the model, but also the typical variations among training examples. This enables learning of style parameters, as discussed above. Furthermore, the GP is non-parametric, meaning that no analytical form is imposed on the data – in other words, we do not risk introducing erroneous assumptions in the model [1].

The models are trained with motion capture recordings from the database of Schoonderwaldt and Demoucron [7]. We use the bow velocity (V), bow-bridge distance (D), and bow force (F) data from two different professional violinists playing sequences of detached notes in forte (f), mezzo-forte (mf), and piano (p), on each of the four strings. The recorded sequences are segmented into strokes (by detecting bow velocity zero crossings), and all segments are re-sampled to a length of $n = 125$ points, equal to 2.08 s with a sampling frequency of 60 Hz. Figure 3, left graph in each subfigure, show segmented and length-normalized mf curves corresponding to down-bow and up-bow, respectively.

In total, 6 models are learned: f down-bow, f up-bow, mf down-bow, mf up-bow, p down-bow, and p up-bow. There are $m = 16$ training examples for each model. Each model

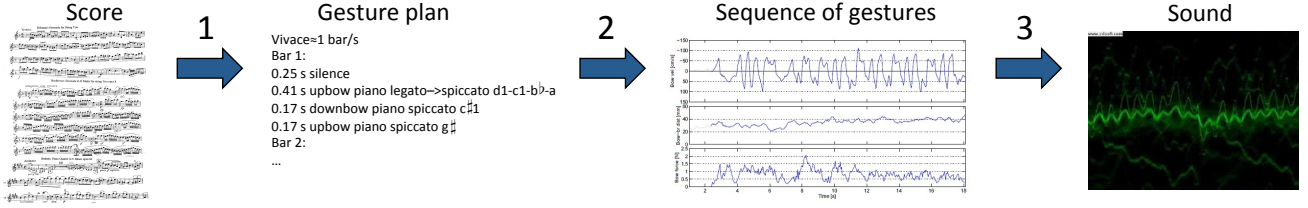


Figure 2. The violin sound generating process. Mapping 1 is well defined, the musical score is simply a coding scheme for the sequence of gestures in the plan. For mapping 3, we use the system developed by Demoucron [5]. The focus of the work here is mapping 2, the generative process of continuous bow motion from the sequence of gestures extracted from the score.

has three dimensions (V, D, F) which are modeled independently from each other – in practice, there are three separate GPs for each model; for V, D, and F, respectively (for *mf* examples, see Figure 3(a,c,e) or (b,d,f)).

A GP is defined as follows (for a more detailed explanation, see [1]): View the training curves for a certain GP (for examples, see the left graph in each subfigure of Figure 3) as an array of tuples $[(x_i, t_i)]$ where $i \in [1, mn]$ (i.e., one tuple for each point on each training curve). Assume the data to be spatio-temporally Gaussian distributed:

$$[x_1, \dots, x_{mn}] \in \mathcal{N}(\mu([t_1, \dots, t_{mn}]), k([t_1, \dots, t_{mn}], [t_1, \dots, t_{mn}])) \quad (1)$$

where $\mu(t) = \frac{1}{m} \sum_{j:t_j=t} x_j$, the mean value of x at time step t , and $k(t, t')$, the covariance between values x at timesteps t and t' . We use a stationary squared exponential covariance function:

$$k(t, t') = \exp\left(-\frac{\|t - t'\|}{2\sigma^2}\right) \quad (2)$$

where σ is the characteristic time dependency length in the GP $x(t)$, and is learned from the data. A natural extension would be to use a non-stationary function, with a time-varying characteristic length $\sigma(t)$. For example, the velocity is by definition 0 at the beginning and end of each bow stroke; this would be learned with a time-varying $\sigma(t)$. We use the GP implementation by Lawrence [11]. Further extensions are discussed in Sec. 5.

Figure 3, right graph in each subfigure, show the GPs learned from the training data in the same subfigure. From these GPs, novel curves with the same characteristics, but with some stochasticity, can be sampled from the learned mean $\mu(t)$ and covariance $k(t, t')$.

The output of the mapping 2 model is a sequence of synthetic bow control curves. The choice of dynamics ($f - mf - p$) and the bowing, are selected according to the notation in the sheet. One V, D, F curve is then sampled for each note (or sequence of notes played with one bow) in the musical sheet, and stretched to the right duration as indicated in the sheet. The curves are then stitched together, forming a coherent synthetic bow motion.

4. LISTENING TESTS

The naturalness of the curves generated from the Gaussian processes was evaluated using a listening test. Violin notes were synthesized using real bow control gestures from the

database [7], and artificial gestures from the Gaussian processes, respectively, and compared to check if they were perceived as significantly different. The focus of the evaluation was not on the realism of the generated sounds as such, rather on the naturalness of the underlying bow motion. This aspect required listeners with extensive own experience of string playing. In order to make a fair comparison, all violin sound stimuli were synthesized in an identical manner (see Figure 2, mapping 3), using the bowed-string model developed by Demoucron [5]. The model, which is implemented using modal synthesis, gives a realistic bowed string sound when controlled by calibrated bow control gestures.

Stimuli. Bow control gestures from the Gaussian processes were compiled for pair of half notes played *detaché* down-bow-up-bow (or up-bow-down-bow), and the artificial V, F and D curves were fed into the bowed-string model. The length of the stimuli were $2 \times 2.08 = 4.16$ s. These stimuli were compared with sounds generated by feeding real motion capture recordings of V, F, D sequences of half notes, also of length 4.16 s, from the database into the same model. Two pitches were used, C4 and G5, played on the G and D string, respectively, combined with two dynamic levels (*mf* and *f*). No vibrato was included. Two independent samples of bowing gestures for each of the four cases were synthesized. A corresponding set of stimuli was generated played up-bow-down-bow. In all, 16 stimuli were generated from the GPs, and 32 from the database by including recordings of two players.

A selection of four down-bow-up-bow cases (and corresponding four up-bow-down-bow cases) from the Gaussian process stimuli was made after selective listening. The purpose of the selection was to limit the size of the listening test, and to include stimuli with different qualities of the attack which normally occur in playing; perfect, choked (prolonged periods) and multi-slip attacks. The 2×4 stimuli from the Gaussian processes were combined with the corresponding cases from two players. The listeners judged each of the 3×8 stimuli three times, in all 72 responses. The stimuli were presented in random order, different for each listener.

Procedure. Eight string players participated in the test; one professional and seven advanced amateurs. Their musical training as violin players ranged between 7 and 19 years, and they had between 12 and 32 years of experience of playing string instruments in general.

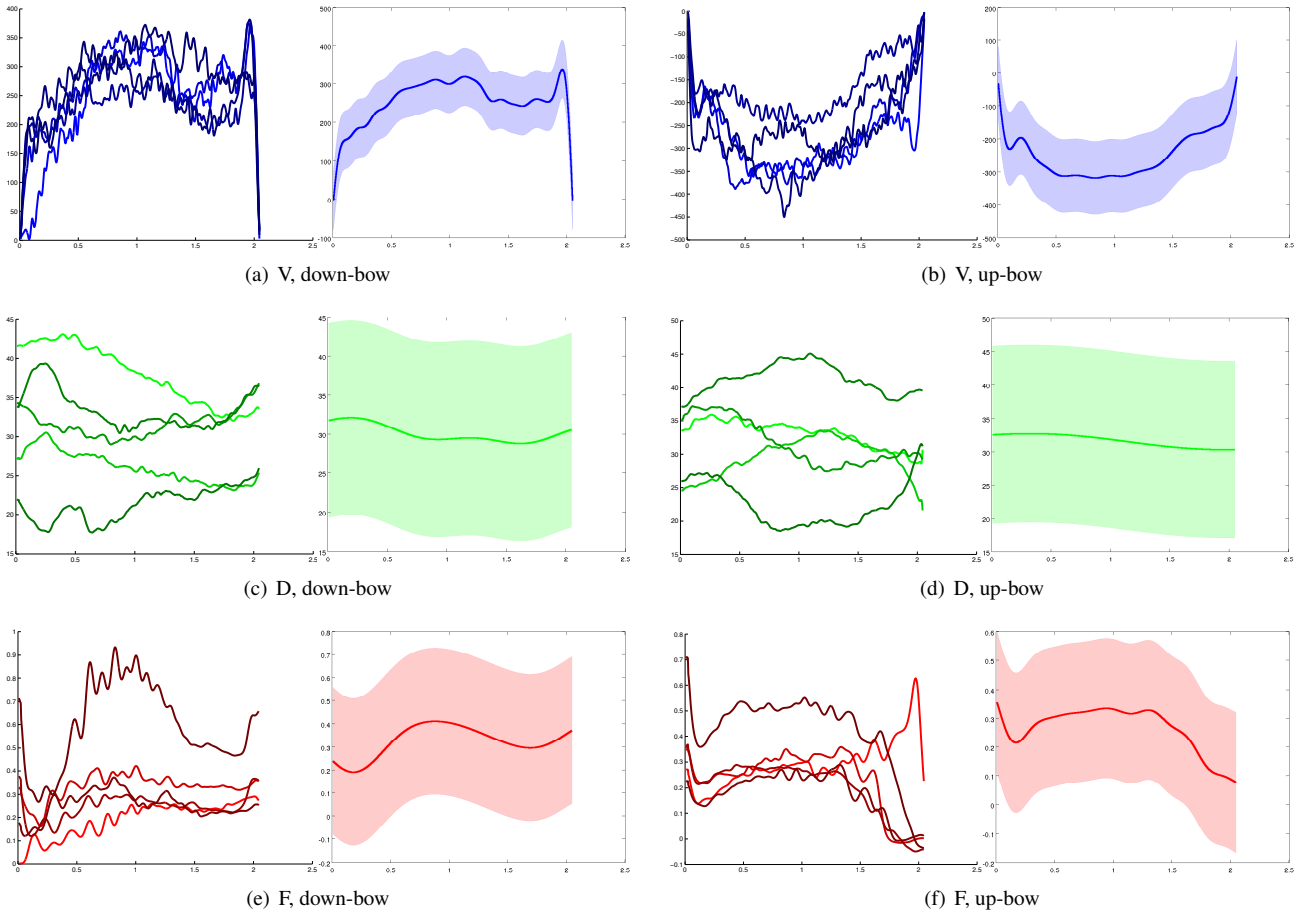


Figure 3. GPs trained with motion capture recordings of half notes played *detaché* from the database in [6]. A bow stroke consists of bow velocity (V), bow-bridge distance (D), and bow force (F) curves. The V, D, F curves are modeled independently from each other in three GPs. There are separate models for up- and down-bow, and for forte (f), mezzoforte (mf), and piano (p). The figures show models for mf: (a,c,e) The three GPs for down-bow. (b,d,f) The three GPs for up-bow. Left in each subfigure: Examples of training data. Right in each subfigure: The GP learned from this training data. The shaded region indicates the standard deviation $\sqrt{k(t, t)}$. Note that the covariance is stationary (time independent). A natural extension would be to use a non-stationary function, with a time-varying characteristic length $\sigma(t)$.

The task of the listeners was to rate the naturalness of the bow motion. They were explicitly instructed to not pay attention to the general quality of the violin sound, but to focus on the underlying bow motion by responding to the question “How natural is the bow motion that produced the notes you heard?” The response was given on a scale from 0 (“artificial”) to 1.0 (“like a human player”) using a slider on a computer screen. The stimuli could be repeated as many times as desired, but that feature was rarely used. A short familiarization with the task including four stimuli initiated each test session. The listeners were informed about that the sounds could contain attacks of different quality and other noises which normally occur in playing. They were neither informed about the origin of the stimuli, nor about the purpose of the test.

Results. The results are summarized in Figure 4, showing average ratings across all 72 stimuli for each of the eight listeners. It is clear that the listeners had different opinions about the general level of naturalness of the bow gestures. Most listeners, however, gave an average response midway between “artificial” (0) and “like a human”

(1.0), with a notable exception for Listener 7. The important result, however, is that the bow gestures generated by the Gaussian processes were judged to be more natural than the real gestures from the database by all but two listeners (5 and 7). For Listeners 1 and 2, the preference for the Gaussian processes was quite marked. The consistency and repeatability in judgements appeared satisfactory as indicated by the error bars.

A conservative interpretation of the results is that six out of eight experienced string players did not hear any difference between synthesized violin notes generated by bow gestures from Gaussian processes and real performances, respectively. Two listeners perceived the Gaussian processes bow gestures as more natural than the corresponding real ones.

5. CONCLUSIONS

We presented a probabilistic approach to modeling violin bowing gestures, for the purpose of synthesizing violin sound from a musical score. The gesture models were based on GP, a principled probabilistic framework. Models

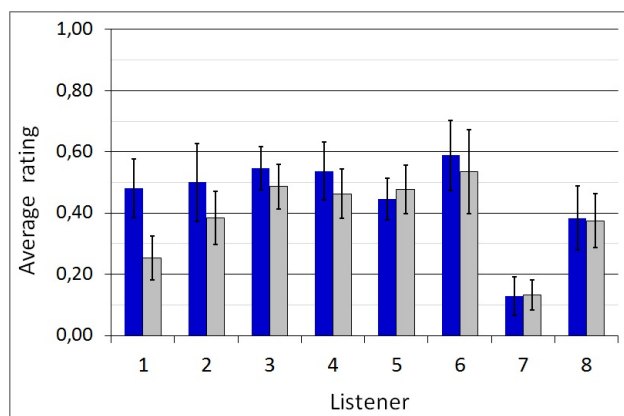


Figure 4. Result of the listening test. Average scores for eight listeners across all stimuli generated by bow gestures from the Gaussian processes (dark blue) and from real bow gestures in the data base (light grey). Error bars correspond to ± 0.5 standard deviation.

for bow velocity, bow-bridge distance and bow force during a stroke were learned from training data of recorded bowing motion. From the models of bow motion during a stroke, slightly novel bow motion could be synthesized, varying in a random manner along the main modes of variation learned from the data. Such synthesized bow strokes could be stitched together to form a continuous bowing motion, which was used to drive a physical violin model, producing naturalistic violin sound. Listening tests showed that the sound produced from the synthetic bowing motion was perceived as very similar to sound produced from real bowing motion, recorded with motion capture.

5.1 Future Work

The proposed framework built on GP allows for principled extensions to address aspects *I* and *II* in Sec. 2. Capturing aspect *I* requires models where style modes can be learned from data. We propose to use Gaussian process latent variable models (GPLVM) [11] which are an extension of GP, and have been used extensively for modeling human behavior. Long-term dependencies (aspect *II*) can be modeled using grammar-like models. We propose to use dynamic Bayesian networks (DBN) [12] where each node is a GPLVM representing a gesture.

Fewer, parameterized, gesture classes. The number of conceivable note classes is very large due to the combinations of score characteristics: duration, pitch, dynamic level, articulation, and bowing style. An example of a note class would be [A4, quarter note, forte-level, sforzando (accented), staccato (short), preceded by a long note at piano-level and followed by a rest (silence)]. One could also add instrument-specific information, e.g., that the note should be played in a down-bow on the D string. A combinatorial explosion of cases will emerge; the task of assigning a set of bow control gestures to each note class will not be scalable, when going from a basic division into few note classes based on a couple of broad characteristics (e.g., high pitch, long note, loud) to a more detailed

description as in the example above.

In [3], 102 note classes were used even without including pitch and duration among the note class properties. These were handled later in the selection of a suitable sample of bowing gestures from the generic gestures. A particular concern in music performance is the strict timing imposed by the note values and tempo given in score. The attack and termination of a note cannot be stretched or compressed much without changing the perceived quality [13].

We propose to use the experience of expert players to investigate to what extent the number of note classes can be restricted. Bowing styles like *detaché*, *legato*, *spiccato* are examples of note characteristics which definitively define different note classes. Pitch, duration, dynamic level are examples of characteristics which are possible to encode as latent parameters in the GPLVM models. The context dependence – which notes come before and after – may also be possible to handle to a certain extent by controlling the end constraints when sampling from the processes.

Learning ubiquitous style modes. The linear and well-defined modes of variation described above are possible to train in a supervised manner, since the training examples could be labeled with objective measures of volume (dB), duration (s), pitch (Hz). However, style variations such as baroque vs. romantic violin playing style are not apparently observable in recorded bowing parameters. As discussed in aspect *I* above, a highly desirable property of a violin synthesizer is the possibility to control high-level performance style parameters.

It is however possible to learn unobservable latent parameters from data using GPLVM [11]. Any level of supervision can also be included if such is available; for example, a mode of variation corresponding to music style could be learned from the data given that the examples were labeled with baroque – Viennese classic – romantic – jazz etc. It will be necessary to collect a wider range of data examples.

Learning phrases, bow planning, and other long time-range dependencies. Addressing aspect *II* above, we will then proceed to modeling dependencies between gestures that are separated in time. This is necessary in order to be able to represent phrase-based music interpretation (see Sec. 2). Moreover, on a slightly shorter time scale, the finite length of the bow needs to be taken into account. This will require a preplanning which takes many notes ahead into account so that bow changes can take place at musically motivated instances, and that notes are played using a "natural" bowing direction (down-bow/up-bow). Related to this question is the modeling of sound feedback in the gesture production [5, 14]. Sound feedback is very important for small modulations in bowing motion, e.g., during *spiccato*.

To represent hierarchical dependencies and dependencies between a whole sequence of gestures – a gestural "grammar" – we will employ Dynamic Bayesian Networks (DBN) [12] which is the mathematically principled way of representing probabilistic dependencies between data segments over time.

6. REFERENCES

- [1] C. Rasmussen, *Gaussian Processes in Machine Learning*. Springer, 2004.
- [2] A. Perez, *Enhancing Spectral Synthesis Techniques with Performance Gestures using the Violin as a Case Study*. Universitat Pompeu Fabra, Spain: PhD Thesis, 2009.
- [3] E. Maestre, *Modeling Instrumental Gestures: An Analysis/Synthesis Framework for Violin Bowing*. Universitat Pompeu Fabra, Spain: PhD Thesis, 2009.
- [4] E. Maestre, M. Blaauw, J. Bonada, E. Guaus, and A. Perez, “Statistical modeling of bowing control applied to violin sound synthesis,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 4, pp. 855–871, 2010.
- [5] M. Demoucron, *On the Control of Virtual Violins: Physical Modelling and Control of Bowed String Instruments*. KTH, Sweden: PhD Thesis, 2008.
- [6] E. Schoonderwaldt, *Mechanics and Acoustics of Violin Bowing: Freedom, Constraints and Control in Performance*. KTH, Sweden: PhD Thesis, 2009.
- [7] E. Schoonderwaldt and M. Demoucron, “Extraction of bowing parameters from violin performance combining motion capture and sensors,” *Journal of the Acoustical Society of America*, vol. 126, no. 5, pp. 2695–2708, 2009.
- [8] E. Schoonderwaldt, “The violinist’s sound palette: Spectral centroid, pitch flattening and anomalous frequencies,” *Acta Acustica united with Acustica*, vol. 95, no. 5, pp. 901–914, 2009.
- [9] —, “The player and the bowed string: Coordination and control of violin bowing in violin and viola performance,” *Journal of the Acoustical Society of America*, vol. 126, no. 5, pp. 2709–2720, 2009.
- [10] G. Percival, N. Bailey, and G. Tzanetakis, “Physical modeling meets machine learning: Teaching bow control to a virtual violinist,” in *Sound and Music Computing Conference*, 2011.
- [11] N. D. Lawrence, “Probabilistic non-linear principal component analysis with Gaussian process latent variable models,” *Journal of Machine Learning Research*, vol. 6, pp. 1783–1816, 2004.
- [12] K. P. Murphy, *Dynamic Bayesian networks: Representation, Inference and Learning*. University of California at Berkeley, USA: PhD Thesis, 2002.
- [13] K. Guettler and A. Askenfelt, “Acceptance limits for the duration of pre-Helmholtz transients in bowed string attacks,” *Journal of the Acoustical Society of America*, vol. 101, pp. 2903–2913, 1997.
- [14] —, “On the kinematics of spiccato and ricochet bowing,” *Catgut Acoustical Society Journal*, vol. 3, no. 6, pp. 9–15, 1998.

VIOLIN QUALITY EVALUATION: EXAMINING THE ROLE OF AUDITORY AND VIBROTACTILE FEEDBACKS

Indiana Wollman
LAM, IJLRDA, Université
Paris 6
woll-
man@lam.jussieu.fr

Claudia Fritz
LAM, IJLRDA, Université
Paris 6
fritz@lam.jussieu.
fr

Jacques Poitevineau
LAM, IJLRDA, Université
Paris 6
jacques.poitevineau
u48@orange.fr

Stephen McAdams
Schulich School of Music,
McGill University
smc@music.mcgill.c
a

ABSTRACT

The present paper explores the role of auditory and vibrotactile feedback involved in violin playing and evaluation. Twenty professional violinists took part in a perceptual experiment employing a near-blind violin evaluation task under three different playing conditions: i) normal, ii) with auditory masking, and iii) with vibrotactile masking. Under each condition, the violinists evaluated five violins according to criteria related to violin playing and sound characteristics and rated their overall quality and relative preference. Both auditory and tactile feedback appear important in the violinists' evaluations; if on average, auditory masking had a greater effect on criteria evaluation and preference than did tactile masking, their relative importance was found to depend on the violinist, the violin and the type of evaluation (criteria or preference). The overall quality ratings were accurately predicted by the rating criteria, which also proved to be perceptually relevant to violinists, but were poorly correlated with the preference ratings, suggesting that the two types of ratings may stem from different strategic decision processes.

1. INTRODUCTION

As in the performance of many musical instruments, violinists are in intimate contact with the violin, the instrument being held very close to the player: the chin and the shoulder hold the violin body, the left hand the neck, and the right hand the bow. In addition to the sound, the musicians receive vibrotactile feedback of the violin as an implicit complement of the music produced.

In the violin literature, the notion of "feel" of a violin is commonly associated with vibrotactile feedback that arrives through the left hand [1, 2, 3]. However, all the studies which have dealt with the "feel" of a violin are limited to the study of the acoustical features of violin structures, from an acoustical point of view, but lack perceptual validation of their claims. A study of "feel" in a perceptual experiment is thus needed to quantify the extent to which the tactile modality alone can contribute to the perception of the overall quality of a violin. Eventually, this examination could help better correlate acoustical properties of the instrument with perceptual characterization of violins by deepening knowledge of the vibrational modes that influence the feel of violins more

than of what determines the tonal quality of the instrument. Thus an initial investigation of how important the tactile sense is in violin playing and evaluation is required, particularly in relation to the contribution of the auditory modality.

1.1 The role of Auditory and Vibrotactile feedback in musical instrument performance

Auditory and tactile modalities are regarded as important feedback mechanisms that enable musicians to control and experience their playing.

Research on musical instrument performance has demonstrated that an absence of auditory feedback does not impair or disrupt performance [4, 5, 6] and only slightly affects expressive performance [7]. One study [8] conducted with pianists who were blindfolded and "deafened" by way of headphones with white noise showed that musician experts could discriminate the quality between three pianos primarily based on their feeling of the mechanical response of the instrument. This research needs to be extended to the performance of other musical instruments.

Askenfelt and Jansson [9] suggested that the vibrations of the violin that can be felt by performers play an important role in ensemble playing, where hearing one's own auditory feedback is not always possible, as well as in timing. Goebel and Palmer [10] extended this finding by highlighting the role of tactile feedback in performance timing accuracy for pianists. However, it should be noted that information concerning the role of tactile cues in musical performance and instrument evaluation is still scarce in the literature.

1.2 Aim and Research Questions

This study aims to investigate the respective contribution of two sensory modalities on violinists' evaluations of criteria related to the sound and the "feel" of instruments by exploring how much auditory masking and tactile masking affect the violinists' evaluation of violin qualities. Indeed, sensory masking is a way to uncouple the two modalities and so to study each sensory modality separately.

The perceptual experiment presented in this paper thus investigates whether violinists still have sufficient cues to evaluate violins under playing conditions with sensory

masking and if so to determine on which cues they rely to compensate for either type of sensory masking. Furthermore, the paper explores whether these cues lead to consistent evaluation. A list of evaluation criteria related to common specific attributes of violin playing and sound characteristics was chosen for the evaluation process. The investigation was designed to assess how these standard criteria relate to each sensory modality to better understand their relative significance for professional violinists.

The first part of the paper presents the perceptual experiment and investigates the extent to which auditory and tactile feedback can be eliminated without impairing violinists' playing or their ability to evaluate violins. In the second part of the paper, we test the effects of the playing conditions with different sources of sensory feedback on the answers given by violinists to the same set of questions related to quality evaluation.

2. METHOD

2.1 Participants

Twenty classical professional or semi-professional violinists took part in the experiment: eleven females and nine males with a mean age of 28 years old (range: 18-54). None of them reported having auditory or tactile deficits. Subjects were paid for their participation.

2.2 Violins

Five violins of different make and age — made between the early 18th and early 21st centuries — and ranging from \$20,000 to \$30,000 were assembled for this study. The five violins will be referred to as VA,VB,VC,VD, and VE. The instruments were chosen to be geometrically similar, so it was not easy to differentiate them by merely holding them or seeing them with seriously reduced visual information. In the same way, identical shoulder rests (Kun Original model) were used for all five violins.

2.3 Procedure

The experiment employs a blind evaluation task in which the violinists played and rated violins under different playing conditions: normal (N), with auditory masking (noA) and with vibrotactile masking (noT). The order of the three conditions was counterbalanced across participants. Within each playing condition, the order of presentation of violins was randomized, and the participants were instructed to complete a series of evaluation tasks for one violin at a time:

1. For each violin, they were first presented with eight different perceptual descriptors of violin playing and sound characteristics (see 2.4), appearing one at a time in random order and presented in the form “the violin is X”. They were asked to assess the magnitude and importance of each criterion by moving an on-screen slider along a “strongly disagree”/“totally agree” scale for the magni-

tude and by moving an on-screen slider along a “not important”/“very important” scale for the importance. They could choose not to answer the questions if they judged that the criterion was not relevant; they were asked to indicate the reasons by ticking as many boxes as appropriate, namely, “I cannot evaluate this criterion on this violin”, “I cannot evaluate this criterion in this playing condition”, “I cannot evaluate this criterion because it does not mean anything to me”.

2. Ratings of the overall quality of each violin were made by moving an on-screen slider along a “bad quality”/“good quality” scale.

3. Finally, participants were able to replay all the violins and were asked to rank the five violins by preference by moving five on-screen sliders along the same “least-preferred”/“most preferred” scale. They were instructed to use the entire slider range in their ranking. The positions of the five sliders along the preference axis thus constitute scores between 0 and 1 regarding violin preference.

Figure 1 illustrates the evaluation procedure.

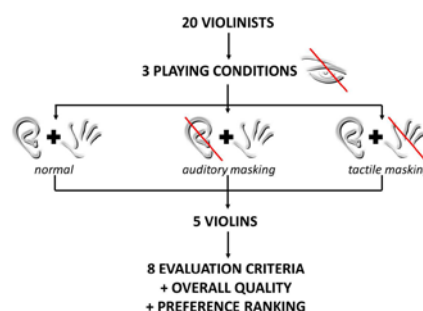


Figure 1. Experimental procedure

At the beginning of the session, the violinists were given 20 min to play and freely explore the five violins under normal conditions to familiarize themselves with the violins. Also, they had 10 min before each condition to familiarize themselves with the experimental playing situation by performing on their own violin taken as a reference.

No constraints were imposed on the musical style and repertoire used by the violinists during the experiment. The experiment lasted approximately two and a half hours.

2.4 Evaluation Criteria

The selection of criteria used in the experiment was based on previous studies investigating violin quality [11, 12]. The criteria were assumed a priori to be evaluable through one or both sensory modalities under study, so that all criteria were presented under all sensory feedback conditions.

EASE OF PLAYING: the violin is easy to play, speaks easily = the violin requires minimal effort to produce sound, it is easy to avoid wolf tones and easy to “get around” on the instrument.

LIVELINESS: the violin is lively.

RESPONSIVENESS: the violin is responsive, speaks quickly = the violin has an immediate response to pressure and speed, there is no delay in response (almost no time between touching the string and the desired sound being produced).

SOUND RICHNESS: the violin has a rich and full sound = the violin produces a sound that is rich in harmonics and overtones.

DYNAMIC: the violin has a broad dynamic range, from piano to forte

LOUDNESS/POWER: the violin is loud and powerful.

EVENNESS: the violin is well-balanced across the strings.

SOUND PALETTE: the violin has a broad sound palette with different colors.

2.5 Control of sensory loss

2.5.1 Visual Feedback Control

The experiment took place under low lighting conditions, and dark sunglasses were worn by the violinists to prevent detailed visual feedback that could help them recognize the instruments (such as by the color and texture of their wood and varnish). Nevertheless, they could see enough to comfortably play and complete the evaluation task on the computer screen.

2.5.2 Auditory Feedback Masking

The combination of earmuffs in addition to in-ear monitors playing white noise with a bandwidth of 20-20,000 Hz (see [13]) was chosen to prevent the musician from hearing the airborne sound produced by the violin.

A volunteer violinist was recruited to record sound levels that typically exist in violin playing. He was equipped with head microphones (Soundman OKM II Classic/Studio A3) and earmuffs (Bilsom Lightning L3, SNR 34 dB). He was asked to play forte on the G string of a violin considered as the most powerful string. The level of the masking noise for violin auditory feedback was chosen to correspond to 6 dB above the violin sound level recorded at the left ear.

In the end, the combination of earmuffs and in-ear monitors playing white noise at 90 dB SPL masked the violinists' auditory feedback. Finally, passive anti-vibration material (Vibra Block® Sound Deadening Material) was added to the chin rest to damp the sound that could be perceived via bone conduction through the jaw.

During the experiment, participants reported not being able to recognize the pitch of the notes being played, being able to slightly detect that they were playing, ensuring a minimum playing comfort. It should be noted that all violinists but one could evaluate violins under the auditory masking condition. Although the one exception had "naturally" performed on all five violins during the familiarization phase, she considered during the evaluation phase that no criterion was relevant to her because of

the playing condition and thus decided to skip this particular condition.

2.5.3 Tactile Feedback Masking

As discussed previously, different contact points exist between the violinist's body and the violin in a playing situation. We decided to primarily mask the vibrations that can be felt by violinists through the left hand, which has been claimed to be an important cue to the perception of how a violin feels [1, 2, 3], and to simply attenuate the vibrations that can be perceived through the jaw and shoulder.

Three vibrating rings to be worn on the thumb, index and ring fingers of the violinist's left hand were constructed. Each ring consisted of a small vibrator (Dayton Audio DAEX13 Mini Exciter Pair 13mm) held tightly between the finger and an elastic band so that the vibrating surface was in direct contact with the skin of the finger. When vibrating, the rings transmitted the vibrations to the finger flesh and bone. We decided to use a tactile masking noise with a bandwidth of 10-1,000 Hz to feed the three vibrating rings as this corresponds to the skin sensitivity range [14].

The masking threshold was determined in a preliminary experiment employing a psychophysical adaptive staircase procedure. In the end, the tactile masking noise level was set to 9.5 dBV in the rings; vibrotactile attenuation at the jaw and shoulder was achieved by adding passive anti-vibration materials to the chin and shoulder rests.

Participants were equipped with the three rings at the beginning of the experiment and had to wear them during the whole session. However, the rings vibrated only during the tactile masking condition.

3. RESULTS

3.1 Violin Preference under different sensory feedback conditions

3.1.1 General Variability between conditions

Saitis et al. [11] showed that violinists are self-consistent in a preference ranking task under normal feedback playing conditions. For this reason, the normal condition is considered as the baseline situation in the following analyses.

It is interesting to compare our results with those of Saitis et al.. Therefore, all participants' relative ratings of preference were rank transformed. In our experiment, the mean Spearman rank correlations between rankings made under conditions N and noA (19 subjects) and conditions N and noT (20 subjects) are both equal to 0.35 (SD = 0.39 and 0.41 respectively). In both cases, the difference with Saitis et al.'s mean is 0.27. In comparison with N-

and-noA: 95% CI [0.02;0.51], $t(37) = 2.210$, $p = 0.017$ (one-sided), $Pr^*[\text{population difference} > 0] = 0.98^1$.

In comparison with N-and-noT: 95% CI [0.02; 0.51], $t(38) = 2.214$, $p = 0.016$ (one-sided), $Pr^*[\text{population difference} > 0] = 0.98$.

These results suggest that the differences observed in the preference ratings and rankings of the sensory masking conditions regarding the normal playing condition are due to the deprivation of either auditory or vibrotactile feedback or both sources of feedback and not from a lack of consistency in the musicians' ratings.

We computed the Pearson correlation coefficients by participant to study the impact of sensory masking on the preference ratings made in the normal condition. The aim was to investigate whether auditory masking was more disruptive than the tactile masking in a preference rating task (i.e., the correlation between N and noA was lower than the correlation between N and noT) as auditory feedback is the most important source of sensory feedback for violinists in playing.

Sbj #	N/noA	N/noT	Sbj #	N/noA	N/noT
1	0.2	0.4	12	0.3	0.6
2	0.6	0.9	13	0.8	0.1
3	0.1	0.5	14	0.0	0.5
4	0.7	0.1	15	-0.4	0.6
5	1.0	1.0	16	1.0	0.9
7	0.3	-0.6	17	0.6	0.5
8	0.7	0.9	18	0.2	0.6
9	0.3	0.4	19	-0.3	0.3
10	0.5	0.8	20	0.2	0.9
11	0.2	-0.2			
			Mean	0.4	0.5
			(SD)	(0.4)	(0.4)

Table 1. Spearman correlation coefficients between preference rankings made under the normal condition and each sensory masking condition. Violinist # 6 did not complete the auditory masking condition.

Table 1 does not show any overall trend across participants as regards the levels of correlation between preference ratings made under the normal condition and each sensory masking condition. Indeed, it can be seen that some participants have higher correlations between N and noA than between N and noT, whereas the opposite holds for some others. And for still others, another trend is observed. With our data, the mean difference between N-and-noA and N-and-noT is -0.12 , 95% CI $[-0.37; 0.13]$, $t(18) = -1.02$, $p = 0.16$ (one-sided), $Pr^*[\text{true difference} < 0] = 0.84$. So, although the observed difference is negative as expected, the possibility of a positive (or null) population difference cannot be ruled out with sufficient confidence to infer the same conclusion at the population

level. Consequently, with our data, it is not possible to conclude that auditory masking is more disruptive in the preference rating task than is tactile masking; it depends on the violinist.

3.1.2 Auditory vs. Tactile feedback in Violin Preference

Figure 2 shows the mean preference ratings of the five violins under the three playing conditions.

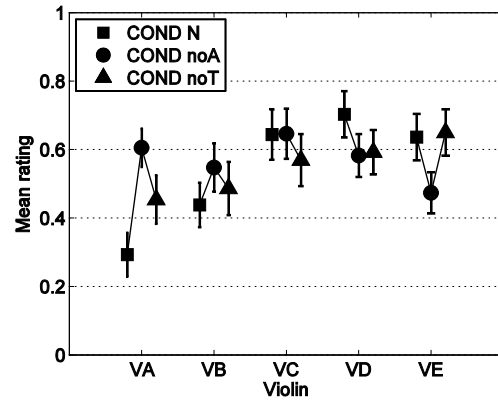


Figure 2. Mean preference ratings under the three playing conditions. The vertical bars represent the standard errors of the mean.

There is no clear consensus on the most preferred or least preferred violins in any of the three playing conditions, confirming that violin preference is highly individual [11]. The variability of preference ratings across participants does not depend on the violin, because the bars representing the standard errors of the mean are of relatively equal magnitude across the five violins.

Under the normal playing condition (squares in Figure 2), violins VA and VB differentiate themselves as the least-preferred ones on average. Under conditions of sensory masking (circles and triangles in Figure 2), the mean preference ratings are of relatively equal magnitude among violins, so less different among themselves than in the normal reference condition. This means that there are fewer perceptual differences between the instruments when eliminating not only auditory cues, but also, more interestingly, the vibrotactile information.

Furthermore, inspection of Figure 2 provides important new data on the role of auditory feedback. Under the auditory masking condition, the mean ratings of violins VA and VE vary relative to the normal-feedback condition, whereas the ratings barely change for violin VC. Moreover, the direction of variation of the ratings between the normal and auditory masking conditions depends on the violin. For VA the mean rating is indeed higher when not heard, whereas the opposite occurs for VE. From this observation, it can be hypothesized that there are three types of violins. Violins of the first type are preferred for their sound, becoming more ordinary, thus less preferred when not heard. Violins of the second type are not preferred for their sound and are thus more preferred when not heard. Lastly, some violins have

¹ Throughout the rest of the article, Bayesian statements will be denoted by Pr^* . They are based upon a non-informative prior (no particular hypothesis is favored a priori). Naturally they are conditional on the data, but, for sake of brevity, the conditional notation is omitted.

something special that does not relate to their sound characteristics. These observations would suggest that the contribution of audition to the evaluation of violins with respect to preference depends on the violin.

In order to test the effect of the factor “violin” on the ratings, the preference ratings in the normal condition were subtracted from those in the noA condition for the 19 participants who completed the task. The overall raw effect of violin, measured as Root Mean Square (RMS) of the pairwise mean differences, is $RMS = 0.26$. A repeated-measures ANOVA on the difference between ratings shows a significant effect of violin, $F(3.66, 65.95) = 5.42$, $\epsilon = 0.75$, $p < 0.01$, which means that the effect of auditory masking on preference ratings depends on the violin ($Pr^*[population\ RMS > 0.18] = 0.95$). Post-hoc analyses (with Bonferroni correction) show that violin VA is significantly different from violins VC ($p = 0.03$), VD ($p = 0.01$) and VE ($p = 0.02$).

3.2 Criteria Evaluation and Overall Quality Ratings

3.2.1 Within-participants Consistency

Under each condition and for each violin, the weighted sum of criteria was calculated as the sum of the products of each criterion’s magnitude rating and its importance rating, divided by the total sum of the importance ratings. The weighted sum of the ratings is thus another way to give a violin a score for a given condition.

	Mean correlation coefficient	Standard error of the mean (sample size)	95% Confidence Interval
COND N	0.85	0.05 (20)	[0.73; 0.96]
COND noA	0.67	0.11 (17*)	[0.43; 0.91]
COND noT	0.89	0.04 (20)	[0.81; 0.97]

* Three participants had missing data for at least three violins in the noA condition, in which case the correlation was not computed.

Table 2. Mean Pearson correlation between the weighted sum of the criteria ratings and the overall quality ratings for the five violins under the three playing conditions

Table 2 reveals that the weighted sums of the ratings are very well correlated with the overall quality ratings of the violins under all conditions. The participants thus appear to be very consistent in their evaluations.

3.2.2 Effect of Auditory and Tactile Masking on Criteria Evaluation

Because the instructions allowed violinists to decide whether a given criterion was relevant or not, the ratings are consistent neither across violins nor across conditions. Thus the data on criterion assessments were not subjected to inferential statistical analyses. The ranges of the num-

ber of times each criterion was found to be relevant to the violinists across violins are given in Table 3. The data are presented by criterion and condition.

Sample size	COND N 20	CONDnoA 19	CONDnoT 20
Ease of Playing	[20 - 20]	[16 - 18]	[18 - 19]
Liveliness	[18 - 20]	[11 - 13]	[18 - 19]
Responsiveness	[20 - 20]	[15 - 17]	[19 - 20]
Dynamics	[20 - 20]	[10 - 12]	[20 - 20]
Loudness	[19 - 20]	[12 - 13]	[19 - 20]
Evenness	[20 - 20]	[11 - 13]	[19 - 20]
Sound Richness	[20 - 20]	[9 - 11]	[20 - 20]
Sound Palette	[20 - 20]	[6 - 9]	[20 - 20]

Table 3. Ranges of the number of times each criterion was found to be relevant to the violinists across violins.

Considering Table 3, there is tentative evidence of three separate groups of criteria within the list, concerning the sensory modality to which each criterion is related. One group consists of criteria mainly related to violin sound, namely “Sound Richness” and “Sound Palette”. A second group consists of four criteria that relate to both auditory and tactile modalities. These are “Liveliness”, “Dynamics”, “Loudness/Power” and “Evenness”. Incidentally, one could infer that “Dynamics” and “Evenness” are more associated with auditory cues than the other two considering the number of violinists who found them to be relevant in the absence of auditory feedback. Finally, the group consisting of “Responsiveness” and “Ease of Playing” includes criteria that depend to a large extent on tactile cues.

Note that for the total number of times a criterion was found to be irrelevant, the “playing condition” was primarily responsible at 90.7%. The “violin” accounted for 8.7%. Finally, “because the criterion did not mean anything” accounted for only 0.6%. Hence, the list of selected criteria is validated by this latter result, that is to say, the list made sense to the violinists.

3.2.3 Overall Quality vs. Preference

Table 4 reveals some interesting facts concerning the violinists’ ratings given the nature of the evaluation task.

	Mean correlation coefficient	Standard error of the mean (sample size)	95% Confidence Interval
COND N	0.11	0.12 (20)	[-0.14; 0.36]
COND noA	0.12	0.13 (19)	[-0.15; 0.39]
COND noT	0.32	0.12 (20)	[0.07; 0.57]

Table 4. Mean correlations between preference ratings and overall quality ratings by condition

Preference ratings are poorly correlated with the overall quality in all three conditions. In other words, one can give a strong, even the best, rating to a violin in an “objectivized” evaluation task — overall quality rating — and then poorly rate that same violin when it comes to preference. What is also remarkable is that the same phenomenon happens under all conditions, suggesting that a violin preference judgment is commonly unrelated to violin “value judgment”. Incidentally, no criterion ratings give better correlation with preference ratings than did the overall quality ratings.

4. CONCLUSIONS

The present paper has investigated the role of auditory and vibrotactile feedback in the left hand in violin quality evaluation.

It was assumed that the more a given type of sensory masking is disruptive in the evaluation task compared to the normal feedback condition, the more this sensory modality is important in violin playing and evaluation. Even though the trend of criterion evaluation being harder was apparent in the auditory masking condition (a much smaller number of violinists found some criteria to be more relevant in the auditory masking condition than in the tactile masking condition), it must be remembered that this was not necessarily the case at an individual level, because all criteria were still relevant to a third of the participants. Similarly, on average across participants, auditory masking seemed to have greater effect on preference (considering the mean preference ratings). However, this was not necessarily the case at an individual level, because there was no overall trend across violinists concerning correlations between preference ratings made under the normal condition and each sensory masking condition. Therefore, whereas anecdotal evidence suggests that auditory feedback is of major importance in the assessment of violin quality, its relative importance compared to tactile feedback was shown to differ among violinists in both types of evaluations.

However, in the noA condition, the tactile information not only consisted of the vibrotactile feedback at the left hand but consisted also, and certainly to a large extent, of tactile feedback at the bow hand. Since there was no masking of the right hand, it cannot be excluded that it did not play a role. Players can for instance feel the difference between Helmholtz motion and multiple slipping, which might give important cues for the evaluation of some criteria, like the responsiveness or the ease of playing. Moreover, it should be noted that if we can observe that something passes through the tactile sense as evaluation was not completely disrupted by auditory masking, we cannot quantify the extent to which the tactile information would be used in real life. Therefore, it cannot be ruled out that participants changed their evaluation strategy in the sensory masking conditions compared to the normal condition considered as the baseline situation, close to what is done in practice. Nevertheless, considering the complexity of the task, it is highly unlikely that

participants were able to learn the task so quickly if they did not use tactile information, even without being conscious of it, in actual practice.

A list of evaluation criteria related to common specific attributes of violin playing and sound characteristics was chosen for the evaluation process. Participants were asked to assess the magnitude and importance of each criterion in the overall evaluation of each violin. Under all sensory feedback conditions, a large correlation between criterion ratings and overall quality ratings was found, demonstrating an impressive consistency in the way violinists carried out the evaluation task. It can thus be argued that violinists relied on cues that led to consistent evaluation of the violins.

The investigation intended to establish a correspondence between the different attributes of a violin and the sensory modality with which they are associated to better understand their significance for professional violinists. Three separate groups of criteria were suggested. One group consisted of criteria associated nearly exclusively with the auditory modality, another group included criteria associated to a large extent with the tactile modality, and a third group included criteria associated with both auditory and tactile modalities in a more balanced proportion. However, the fact that the criteria Sound Richness and Sound Palette were found to be relevant by about a third of the participants under auditory masking conditions may suggest that these terms do not relate unequivocally to sound properties, but could be based on multi-sensory information processing too. Therefore, if Sound Richness and Palette can be evaluated without hearing the sound of the violin, how can acoustic signal properties only — such as spectral centroid — account for the perceptual criterion ratings? This may explain why correlating physical properties with perceptual properties has been challenging so far.

Last but not least, and very surprisingly, weak correlations were found between how violinists objectively assessed the overall quality of the violin and how they rated and ranked the violins in terms of preference, under any of the sensory feedback playing conditions. One might suppose that in this experiment, what drove the participants’ overall quality ratings came from the stages through which the evaluation proceeded. Although the order of presentation of criteria was randomized across conditions, we cannot exclude the possibility that the participants unconsciously based the overall quality rating on an overall mean of criterion ratings. Nevertheless, considering the complexity of the task, it is more likely that, in fact, the two types of ratings are almost unrelated for the musicians at a psychological level. Preference would then be based on other criteria, whose weighted combination is not identified at this stage.

The results of the present experiment indicate that the type of rating — overall quality vs. preference — should be carefully considered in designing an experiment that aims to study instrument evaluation by players. It can be

hypothesized that acousticians probably care more about absolute ratings of instruments in their attempt to correlate physical and mechanical properties of instruments with perceptual instrument quality, whereas instrument makers may find that preference rating is of crucial importance.

Acknowledgments

This work was supported in part by grants from the Canadian Natural Sciences and Engineering Research Council and the Canada Research Chairs Program to SM. The authors would like to thank all the violinists who participated in this study, and to acknowledge the luthier Olivier Pérot for his kind loan of violins.

5. REFERENCES

- [1] K.D. Marshall, "Modal analysis of a violin," in *J. Acoust. Soc. Am.*, 1985, 77 (2), 695–709.
- [2] C.M. Hutchins, "Effect of an air-body coupling on the tone and playing qualities of violins," in *J. of the Catgut Acoust. Soc.*, 1985, 44, 12–15.
- [3] J. Woodhouse, "The acoustics of "A0-B0 mode matching" in the violin," in *Acustica - Acta Acustica*, 1998, 84, 947–956.
- [4] A. Gates, & J.L. Bradshaw, "Effects of auditory feedback on musical performance task," in *Perception and Psychophysics*, 1974, 16, 105–109.
- [5] L.J. Banton, "The role of visual and auditory feedback during sight-reading of music," in *Psychology of Music*, 1995, 23, 3–16.
- [6] S.A. Finney, "Auditory feedback and musical keyboard performance," in *Music Perception*, 1997, 15(2), 153–174.
- [7] B.H. Repp, "Effects of auditory feedback deprivation on expressive piano performance," in *Music Perception*, 1999, 16 (4), 409–438.
- [8] A. Galembo, "Perception of musical instrument by performer and listener (with application to the piano)," in *Proceedings of the International Workshop on Human Supervision and Control in Engineering and Music*, 2001, pp. 257–266
- [9] A. Askenfelt, & E.V. Jansson, "Vibration Sensation in Stringed Instrument playing," in *Music Perception*, 1992, 9 (3), 311–350.
- [10] W. Goebel, & C. Palmer "Tactile feedback and timing accuracy in piano performance," in *Experimental Brain Research*, 2008, 186, 471–479.
- [11] C. Fritz, J. Curtin, J. Poitevineau, P. Morrel-Samuels, & F.-C. Tao, « Players preferences among new and old violins," in *Proc. of the National Academy of Sciences of the USA*, 2012, 109(3), 760–763
- [12] C. Saitis, B.L. Giordano, C. Fritz & G.P. Scavone "Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players," in *J. Acoust. Soc. Am.*, 2012, 132(6), 4002–4012.
- [13] H.F. Shliesser, & R.O. Coleman, "Effectiveness of certain procedures for alteration of auditory and oral sensation for speech," in *Perceptual and Motor Skills*, 1968, 26, 275–281
- [14] R. T. Verrillo, "Vibration sensation in humans," in *Music Perception*, 1992, 9 (3), 281–302.

The influence of different driving conditions on the frequency response of bowed-string instruments

Ailin Zhang

Jim Woodhouse

Department of Engineering
University of Cambridge
Trumpington street
Cambridge, CB2 1PZ, UK

az304@cam.ac.uk

jw12@cam.ac.uk

ABSTRACT

A well-established measurement on the bodies of bowed-string instruments is the input admittance at the bridge. The commonest method for measuring the input admittance is hammer testing on the bridge corner. However, this method has been questioned, due especially to differences between human bowing and hammer impact. In this paper an attempt has been made to survey the influence of different driving methods on the frequency response of a bowed-string instrument, as well as confirm the reliability of the classic hammer method. A series of experiments are carried out with three different driving conditions in the case of a cello: hammer, normal bowing of a string, and step excitation by a breaking wire. The results suggest that there is nothing fundamentally different about the hammer method, compared to other kinds of excitation methods. Some possible differences between the hammer method and normal bowing are also discussed.

1. INTRODUCTION

The input admittance at the bridge (also called “drive point mobility”) has been used by many experimenters to characterize bowed-string instruments. It is a good choice for several reasons. For all the bowed string instruments, the string is the source of energy and the bridge is the main connection “point” to the body. There is always weak coupling between the strings and the instrument body because of their significant impedance difference at this point of contact. Although no direct information related to the radiation of sound is contained in the input admittance, it includes information about energy transfer between the two coupled systems. Second, while the stick-slip string motion associated with bowing is strongly nonlinear, it is generally thought that the body vibration is well approximated as a linear dynamic system. A frequency response function is the usual way to characterise a linear system. Input admittance is a

good choice to capture the essential properties of the instrument as a whole and to relate the vibration characteristics to the performance of the instrument.

The most common approach is to measure the mechanical input admittance using a miniature force hammer tapping the bridge. That naturally invites the question of whether using this method is significantly different from exciting the violin by playing with a bow in the usual manner. The main aim of this paper is to test the accuracy of the hammer method and explore how the frequency response of a cello is influenced by three different driving conditions: controlled input force is applied to a cello separately by an instrumented hammer, a breaking wire, and regular bowing.

Following some background presented in Section 2, descriptions of the set up used in measurements and experimental procedures are outlined in Section 3. Experimental results after calibration are then provided and compared in Section 4. Conclusions are given in the final section.

2. BACKGROUND

We consider first a standard measurement procedure for mechanical input admittance [1]: a cello is driven by means of a miniature impulse hammer at one corner of the bridge while the resulting motion of the bridge is measured on the other corner of the bridge by a laser vibrometer or a small accelerometer. The input admittance can be found by Fourier analysis of the force and output signals. Both the driving force F and the response velocity V are set to be in the bowing direction on the associated nearby string, and are related by the scalar transfer function Y as follows:

$$V = YF \quad (1)$$

In practice, it might be hard to keep the direction of the force and measure the velocity in a single direction or at a single mechanical point due to the limitations of the experimental setup. Idealization of the point force as having one degree of freedom might therefore be inappropriate; in general it might be necessary to consider two or even all three components of the input force and of the resulting velocity, requiring a matrix of transfer functions.

A lot of previous literature has shed the light on the theoretical considerations and measurements of the input admittance. In 1975, the technique for measuring input

admittance in one direction was developed in the laboratory of Jansson [2]. The simple driving system Jansson presented, consisting of an impedance head and strong magnets, was improved further by Firth [3] so as to be widely applied to the measurement of many instruments. Jansson [4] then increased the range of this measurement up to 10 kHz to cover the full frequency range of interest by replacing the impedance head with a magnet-accelerometer-transducer. Taking advantage of it, he mapped out several eigenmodes of different stringed instruments [5, 6]. In an alternative approach, both Eggers [7] and Trott [8] used a rod linking the shaker to the violin bridge. The first study of acoustical properties of the violin performed with an impact hammer was made by Jansson, Bork and Meyer in 1986 [9]. They showed that impulse response measurements can produce fast and sufficiently detailed mechanical admittance. Henceforth, the hammer method was applied widely in research into the violin admittance measurements.

In addition to certain earlier optical investigations, Jansson and Molin introduced the Laser Doppler Vibrometer system to measure the vibration of the violin instead of the accelerometer. By combining it with the impulse hammer, the input admittance can be measured without any extra mass adding to the instrument.

However, all of these measurements only deal with the mobility in one direction of force and one direction of motion. Lambourg and Chaigne [11] aimed at measurements of a 2×2 admittance matrix. This method was still subject to some problems according to Boutillon et.al [12, 13]. Boutillon and Weinreich [14] presented the theoretical background of a 3-dimensional admittance matrix and proposed a new method to realize the measurement. Three ultralight accelerometers were used to pick up vibration of the bridge by comparing the motion under different external loading conditions. Further investigation was reported by Yoder [15] using a one-dimensional mass-loading technique, which decomposed the three dimensional motions to three independent problems.

3. EXPERIMENTS

3.1 Experimental Setup

In the experiments reported here the one-dimensional input admittance is obtained using a cello supported on soft foam pieces and set on a table. A cello of moderate quality but unknown maker was used for the measurements. All measurements were conducted in the same room.

The response acceleration at the C-string corner in all experiments was measured by an accelerometer which is mounted on the bridge with superglue. A lightweight accelerometer (DJB type M2222C) was used in experiments to minimize mass-loading effects. The accelerometer was orientated so that its central line is parallel to the primary vibration direction of the C-string, namely the bowing direction.

It would in principle have been better to use a laser vibrometer rather than an accelerometer. For tests using the impulse hammer for force input this is easy to do, but

the other forcing approaches used here pose problems. Particularly for excitation by bowing as will be described shortly, it is very hard to maintain a good laser signal in the presence of the disturbances a player inevitably imposes on the cello body during bowing and fingering the C string. It was judged better for a clean comparison of the three forcing methods without any uncertainty in the output measurement to use an accelerometer throughout.

For measuring the input signals, two approaches were used. In the first approach, a miniature force hammer (PCB 086D80) fitted with a piezoelectric force transducer was employed to excite the cello bridge at the A-string corner with a wide-bandwidth force. A hard steel impact tip was preferred since it can provide the broadest frequency range. The hammer is attached to a pendulum fixture so that the same point can be hit every time. It also allows the user to control the impact amplitude very easily.

An alternative approach is to measure the force exerted by the vibrating C string on the bridge using a bridge pick-up system. This pick-up system is shown as applied to a standard cello bridge in Figure 1. It consists of two force sensors made from small pieces of piezoelectric crystal mounted immediately beneath the C string-notch on the bridge and producing voltages, after buffering with suitable charge amplifiers, that can be combined to give a signal proportional to transverse force at the string notch in the bowing direction. The whole sensor is embedded in the bridge so that it does not influence the transverse motion of the string. The same sensor can also be used in the wire-break excitation method. The breaking wire is looped around the C string, up against the face of the bridge. The force step as the wire breaks is thus registered by the bridge sensor.

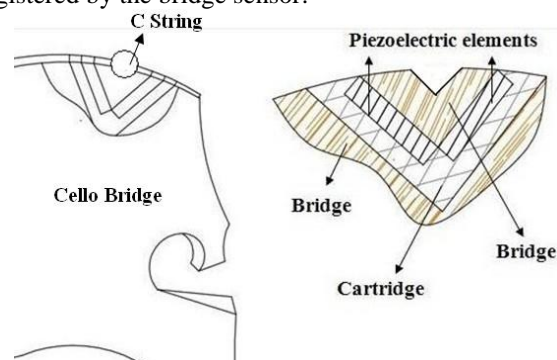


Figure 1. Construction of a cello bridge pick-up system.

3.2 Hammer Method

As already stated above, the body response is not usually measured using the string to excite it, but a mechanical force is applied by a light instrumented hammer hitting the corner of the bridge. However, there are several shortcomings which are hard to overcome. For instance, one cannot be sure of the precise direction of the force, since the steel tip of the hammer is hitting relatively soft wood with local inhomogeneity from the grain structure. Consequently, this approach makes it difficult to attain accurate reproducibility. Another drawback is the fact that the location of the hammer force is at the side of the

bridge, not the same as or even very close to the string notch where force is applied under playing conditions. The final problem arises from the limitation of the frequency range. When we are calculating the frequency response on the computer, the energy in the high frequency range may be very low, making the results at higher frequencies noisy and unreliable.

3.2.1 Measurement procedure

For this method, all the strings are thoroughly damped by a piece of paper woven between them: the test of adequate damping is that no musical pitch can be heard when the strings are plucked. The purpose of damping is to explore the vibration characteristics of the body without the effect of string resonance. However, the influences coming from the strings between the bridge and the tailpiece cannot be removed. Before the measurement, the strings are tuned to their usual playing tensions so that the bridge, body and soundpost are supported and stressed exactly as in normal playing conditions. Also, the influence of the axial stiffness of the strings on the body modes is maintained.

The input impulse is imposed on the C-string corner of the bridge by the impact hammer. Its excitation direction is parallel to the direction of bowing. To obtain the best signal-to-noise ratio, it is necessary to generate multiple impulses and compute an average value. The acceleration signal at the A-string corner is collected by the accelerometer. Both the input and output signals are fed through appropriate charge amplifiers and finally processed by a computer, using software written in MATLAB. The averaged transfer functions together with the associated coherence function are computed in the usual way. The typical experimental set-up is shown in Figure 2.

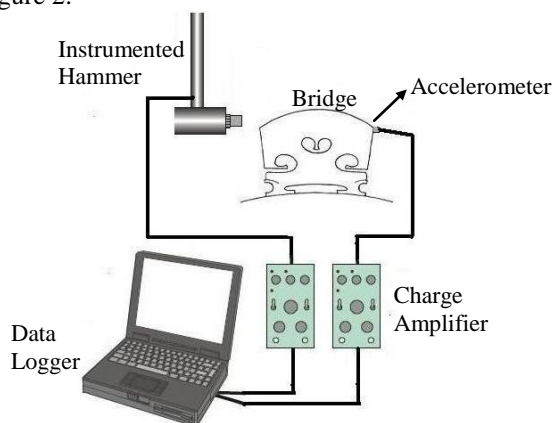


Figure 2. Data logging setup for the hammer method.

3.2.2 Coordinate system

Illustration of the measurement in progress is shown below. Two sets of rectangular axes are shown: the inclined force and response can be resolved into rectangular components. The directions X and X' indicated on Figure 3 are the “directions of bowing” separately for the hammer and response. Y and Y' are the directions perpendicular to bowing, in the bridge plane.

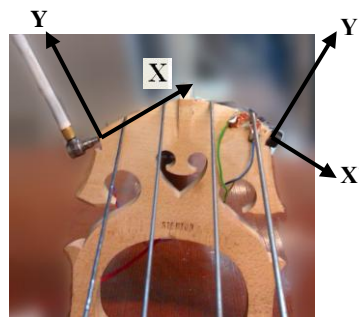


Figure 3. Coordinate system used for the hammer method.

Impulses are carefully applied to the driving point along the X direction; similarly the output signal measured by the sensor is aligned in direction X' at the other side of the bridge. But this raises an important question about whether this is the “best” way to measure a single admittance representing the cello body response to driving by all four strings. This question will be considered further below.

3.3 Wire-breaking Method

The next approach is the wire-breaking method. For obtaining a more controlled direction and location of force, compared to the somewhat uncertain hammer force, a length of fine copper wire is gently pulled until it breaks to impart a step function of excitation to the string just at the contact point with the bridge. The force can be exerted at a precise angle, and the waveform is registered by the bridge sensor. The measurement is easily reproducible and reliable using this method since some of the problems of the hammer method have been tackled. Furthermore, since the using of wire-breaking allows for input signals with a precise direction, the wire-breaking method can be used to investigate the exact effective angles of the force sensors used for measuring the input signals and to provide the corresponding calibration parameters.

However, this method gives relatively low energy input and so might be prone to noise pick up, resulting in the inaccuracy of measurements. The effect of such noise can be reduced by averaging measurements to a certain degree. Another problem is the fiddliness of the wire-breaking: it is much more time-consuming than the hammer method. The wire must be broken off carefully aligned along the desired direction and kept just in the plane of the bridge.

3.3.1 Measurement Procedure

As for the hammer method, the measurements are always made with all strings damped to give results without the effects of string resonances. All four strings are again tuned to normal playing condition.

In the measurement, the string is pulled aside with a piece of copper wire at the C string next to the bridge in the desired direction. The force signal is collected by the pick-up system described in Section 3.1 and then electrically connected to a charge amplifier. The transfer function will be derived from the signals measured by the

force sensors in conjunction with the accelerometer. The experimental set-up can be seen in Figure 4.

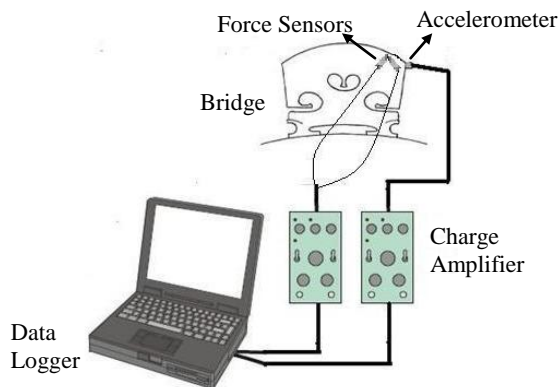


Figure 4. Data logging setup for the wire-breaking method.

3.3.2 Coordinate system

Illustration of the coordinate system is shown below. The directions A, B, X, Y, X' and Y' indicated in Figure 5 are consistent throughout the wire-breaking and bowing method. The direction A and B are the directions of two components of the input force determined according to the arrangement of the two force sensors: they are at approximately $\pm 45^\circ$ to the X direction, but the exact angles are determined by a calibration procedure. The axis X is along the direction of bowing in the plane of the bridge and parallel to the axis X'. The directions Y and Y' are perpendicular to the bowing direction X and X' in the same plane.

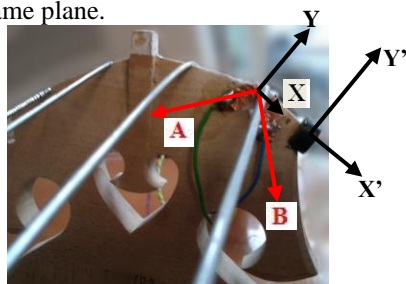


Figure 5. Coordinate system used for the hammer method.

Any force applied in the plane of the bridge can be detected by the two force sensors in the A-B coordinate system. The output from the sensors can be combined to resolve into two perpendicular components, one along the X direction and the other one along the Y direction. This can address the problems of uncertainty associated with the hammer force. During the calibration, wire-breaks are applied at a range of angles from 0° to 90° , easily measured and controlled by means of a piece of paper used as a protractor, mounted to align with the X-Y coordinates. The actual breaking force of the copper wire can be determined by tests using suspended weights to break the wire: the breaking force for copper is very consistent for different samples of wire from the same reel. By fitting calibration factors to these measurements with the logger software, the input force can be resolved accurately and also the magnitude of transfer functions can be converted into meaningful physical units. In the

response measured by the accelerometer, the output signal is determined along the X direction only since a single-axis accelerometer is used here. The response along the Y' direction is ignored.

3.4 Bowing Method

Although the miniature force hammer is so commonly used to obtain the transfer function, in fact bowing from the player can be used to supply response curves under rather natural conditions. By bowing, there is no doubt that the “right kind of force” is applied to the body, located in the right place. The component of force at the bridge in the direction X is measured by the force sensors, just as in the wire-break method, and used as the input signal in the calculation of the admittance.

Although such bowing from a player is the closest to the normal playing condition that can deliver credible measurements, there are some intrinsic disadvantages: Firstly, the tested instrument needs to be held firmly enough to bow. But a firm supporting fixture like a clamping arrangement might change the vibration properties of the instruments greatly. However, the mode frequencies are little affected by supporting the cello on its back on blocks of soft foam, and this support method was used here and found to be firm enough for careful bowing. Furthermore, it is likely that not all components of either force or vibration excited by bowing will be detected by the sensors since only one axis of force input is being recorded, and the accelerometer used here is a single-axis one.

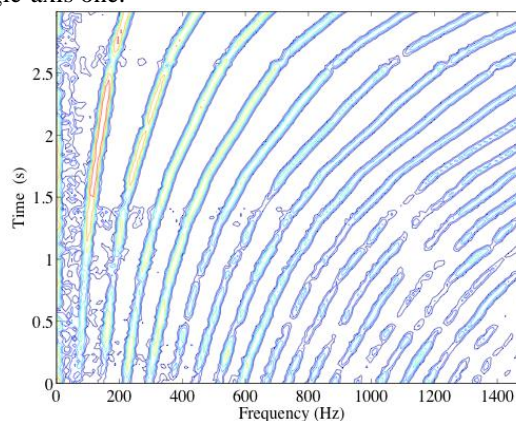


Figure 6. Measured sonogram of a glissando on the C string of a cello.

The measurement is performed with the C string of the cello being bowed, and thus obviously undamped, while the other strings are still damped. Bowing a single note would only provide input force at individual frequencies at the harmonics of the played note, but by playing a one-octave glissando, which is a glide on the string with the finger of the player from one pitch to another, the frequency range can be filled in. A measured sonogram of the sound made by a glissando is shown in Figure 6. It shows the fundamental frequency sweeping upwards in frequency over the played range, and a rich set of harmonics tracking in proportion. By covering at least an octave, the entire frequency range is filled in by energy being supplied by one or more harmonics at some stage during the measurement.

The bowing parameters are controlled to maintain the Helmholtz motion in the C string. The experimental set-up is the same as that of wire-breaking method in Figure 4.

4. RESULTS

The resulting transfer functions are obtained, along the bowing direction as described, by repeating each measurement at least 15 times without interruption, under each of the three driving conditions. All the mean input admittances after calibration in this section are plotted on a log-log scale. A frequency range of 65-4000 Hz is shown, starting from the lowest played note of the cello and covering most of the interesting dynamic response of the body. In addition, the coherences of the three different methods were found to be close to 1 over this frequency range, which suggests the relation between the input and output should be linear in the tested frequency region. For clarity of discussion, we group the input admittances in a sequence of two-way comparisons.

Figure 7 shows two typical results of admittance measured by the hammer method, after dismantling and then reassembling the measurement set-up while aiming for identical conditions. The two curves show good general agreement, but individual deviations are apparent up to about 4 dB. This comparison sets the standard of comparison when looking at the other excitation methods: differences need to be bigger than those seen here before they should be attributed to systematic effects relating to the methods, rather than simply to the limitations on repeatability inherent in tests of this kind.

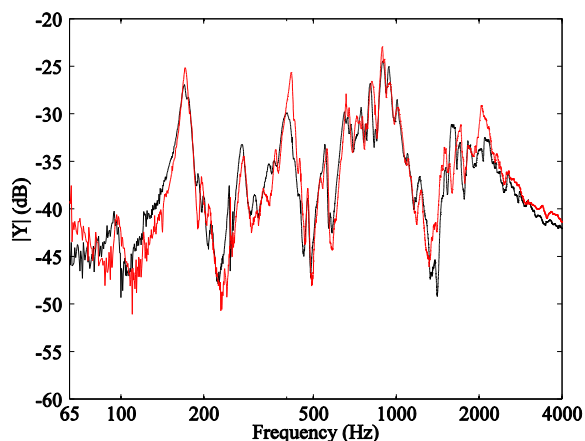


Figure 7. Input admittances of the cello showing variability of the results of hammer testing under nominally identical conditions.

A comparison between transfer functions measured by the wire-breaking method and bowing method is given in Figure 8. The black curve denotes the mean input admittance of the bowing method while the red one indicates that acquiring from the wire-breaking method. The results in this figure show a striking level of similarity between the two admittance curves. Note that the wire-breaking and bowing methods use exactly the same sensors for both input and output, so differences between the curves here can only be attributed to

systematic differences between the forcing methods, or to random effects of non-reproducibility. In fact, the deviations between the two curves are scarcely more significant than was seen in Fig. 7 from nominally identical hammer measurements. At high frequencies, both transfer functions show clear effects of noise pickup, especially the wire-breaking result, but in general these results suggest that there is no major difference between exciting the cello by bowing a string in the normal way, or by applying a controlled external force artificially.

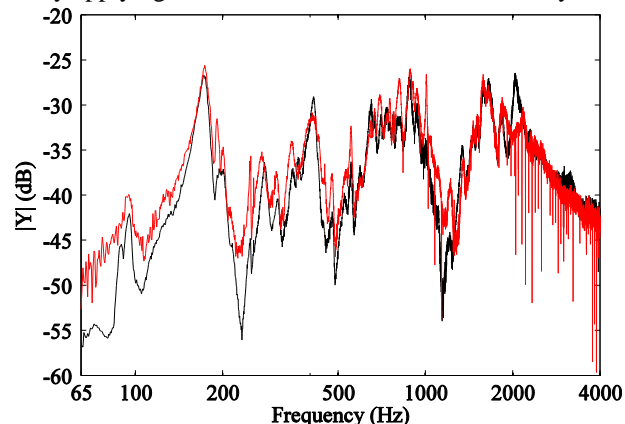


Figure 8. Input admittances measured by the bowing (black) and wire-breaking (red) methods.

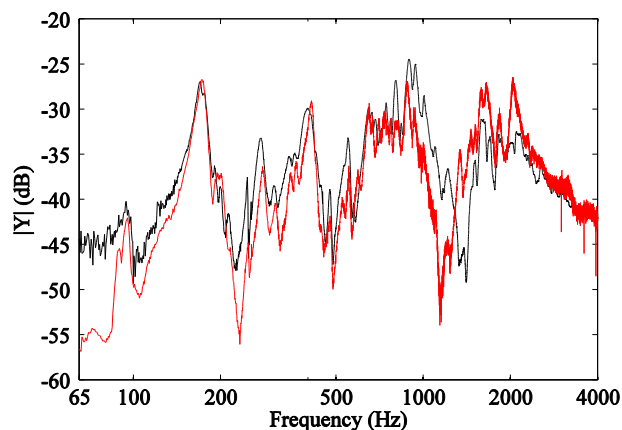


Figure 9. Input admittances measured by the hammer (black) and bowing (red) methods.

Figure 9 shows a comparison of the input admittance by the hammer method and the bowing method. Differences can be seen, perhaps a little greater than those seen in either of the previous comparisons. Of course, the hammer method is expected to be somewhat different from the wire-break method because the force is applied in a different place and with less accurate control over the forcing direction. But the strong similarity among the results in this plot suggests that there is nothing fundamentally wrong with hammer testing by the familiar method. Given its advantages in simplicity, and crucially the fact that it does not require the force transducer embedded in the instrument bridge, it may be concluded that the method is accurate enough for measuring vibration behaviour of bowed-string instruments. Similar findings can be observed when plotting the result of hammer and wire-breaking method on top of each other.

5. DISCUSSION AND CONCLUSIONS

This paper reports an investigation into the influence of different driving conditions on the input admittance of a cello. This study gives the first direct demonstration of the reliability of the classic hammer method for measuring the vibration behaviour of bowed string instruments, in comparison with excitation by normal bowing. Overall we find only slight differences between three different methods of applying force at the bridge of a cello. The results show that the bowing method can be used as a new tool for exploring the characteristics of instruments, but with the drawback that it requires a force sensor on the bridge to measure the input force. The existing sensor is intrusive and not suitable for use on a wide range of instruments, but it is possible that a less intrusive equivalent sensor could be developed, perhaps using PVDF film.

Looking in more detail, these experimental results show some relatively minor differences between the hammer method and bowing method, the explanation for which may lie in the following aspects:

Firstly, the vibration of the instrument body might not be fully described in terms of linear systems theory. If the instrument body shows some nonlinear behaviour, the principle of superposition used in linear systems can no longer be applied to the body vibration and no transfer function can capture the whole behaviour of the body.

Secondly, a transversely vibrating string can have two directions of lateral motion, i.e. two polarizations. For this reason, a 2×2 admittance matrix rather than a single transfer function is required because the input from the string to the body should have two degrees of freedom due to this dual-polarization.

Thirdly, there might be axial forces coming from the string, and also both axial and transverse forces at the other end of the string at the fingerboard. There may also be axial force from the other side of the bridge at the tailpiece: even with damped strings, longitudinal vibration in the string may propagate across the bridge. All these additional input forces from bowing are ignored in the single transfer function discussed here.

The last difference between the bowing and hammer impact concerns the direction and location of the hammer force. The hammer cannot in practice be applied exactly at the string notch, so the force input is inevitably a little different. The direction of the force also raises questions: hitting a wooden bridge in the bowing direction for the nearest string means hitting the bridge corner not quite perpendicular to its cut face, and it is hard to be sure precisely what force is in reality applied.

However, the results shown in this paper suggest that none of these factors has such a strong effect as to render hammer measurements seriously in error. The hammer results match the wire-break and bowing results almost as closely as repeat measurements with the hammer system match each other. The discussion in the previous paragraph may go some way to accounting for the difficulty of obtaining better reproducibility with the hammer method.

6. REFERENCES

- [1] C. Fritz, I. Cross, J. Woodhouse, and B.C.J. Moore, "Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin", *J. Acoust. Soc. Amer.* vol. 122, pp. 3640–3650, 2007.
- [2] I.M. Firth, "Small mechanical impedance head for use with musical instruments," *Acustica* vol. 35, pp. 348–349, 1976.
- [3] I.M. Firth, "Small device for measuring mechanical admittance," *Catgut Acoust. Soc. Newsletter* vol. 25, pp. 19–20, 1976.
- [4] E.V. Jansson, "Tone characteristics of the violin," *Swed. J. of Musicology* vol. 60, pp. 92–94, 1978.
- [5] E.V. Jansson, "Eigenmodes, input admittance and the function of the violin," *Acustica* vol. 50, pp. 329–337, 1982.
- [6] E.V. Jansson, "Admittance Measurements of 25 High Quality Violins," *Acustica* vol. 83, pp. 337–341, 1997.
- [7] F. Eggers, "Untersuchung von Corpus-Schwingungen am Violoncello," *Acustica* vol. 9, pp. 453–465, 1959.
- [8] W.J. Trott, "The violin and its bridge," *J. Acoust. Soc. Amer.* vol. 81, pp. 1948–1954, 1987.
- [9] E.V. Jansson, I. Bork, and J. Meyer, "Investigations into the acoustical properties of the violin," *Acustica* vol. 62, pp. 1–15, 1986.
- [10] E.V. Jansson, and N.E. Molin, "Laser vibrometry measurements of vibration and sound fields of a bowed violin," *Meas. Sci. Technol.* vol. 17, pp. 635–644, 2006.
- [11] C. Lambourg, and A. Chaigne, "Measurements and modeling of the admittance matrix at the bridge in guitars," *Proc. Stockholm Music Acoust Conference 1993*, Royal Swedish Academy of Music, Stockholm, pp. 448–453, 1994.
- [12] G. Weinreich, and L. Yoder, "New method of measuring input admittance of stringed instruments," *J. Acoust. Soc. Amer.* vol. 81, pp. 83, 1987.
- [13] X. Boutillon, G. Weinreich, and N. R. Michael, "Experimental developments for the measurement of violin bridge admittance," *J. Acoust. Soc. Amer.* vol. 84, pp. 179, 1988.
- [14] X. Boutillon, and G. Weinreich, "Three-dimensional mechanical admittance: Theory and new measurement method applied to the violin bridge," *J. Acoust. Soc. Amer.* vol. 105, pp. 3524–3533, 1999.
- [15] L. Yoder, "Violin admittance measurements using a one-dimensional mass-loading technique," *J. Acoust. Soc. Amer.* vol. 89, pp. 1878, 1991.

Guitars and other plucked instruments

Analysis of the harpsichord plectrum-string interaction

D. Chadeaux, J-L. Le Carrou

LAM / d'Alembert,
UMR CNRS 7190,
UPMC Univ Paris 06

S. Le Conte

Laboratoire de Recherche
et Restauration
du Musée de la Musique

M. Castellengo

LAM / d'Alembert,
UMR CNRS 7190,
UPMC Univ Paris 06

ABSTRACT

This paper describes a study of string plucking for the harpsichord. Its aim is to provide an experimentally-based analysis of the plectrum-string interaction and to propose some refinements of an existing model. An experimental setup has been designed using a high-speed camera combined with a laser doppler vibrometer and classical audio recordings. This provides accurate estimations of jack and plectrum motion throughout the harpsichord plucking in a realistic musical context. The set of descriptors extracted from these measurements provides typical orders of magnitude of plucking parameters required to feed and validate the investigated model. Results highlight estimations of the instrumentalist's control parameters as well as of the intrinsic plucking parameters according to the performed sequence *tempo*. Besides, a model of the plectrum-string interaction, which takes into account the section variation at the plectrum tip, gives results close to the experimental ones. This model will be of great interest to enquire about harpsichord plectrum voicing.

1. INTRODUCTION

Among all elements of a harpsichord, the voicing process, which directly deals with the excitatory mechanism tuning, has to be investigated. The harpsichord plucking action mostly consists in a piece of wood moving vertically, called the jack, while the instrumentalist's finger is depressing the key. It conveys the plectrum attached to the jack to pluck the string. Although traditionally made of quill [1], the plectrum is nowadays usually made of delrin. During the voicing process, the maker tunes the plectra to obtain the best equilibrium among the different choirs (8 and 4ft), the bass and the trebles. Practically, it implies to carve each plectrum to adjust the plectrum-string interaction according to the instrument's sound and the instrumentalist's touch. Several kinds of voicing can be outlined in the Musée de la musique's collection; the three main ones are: hard, soft and medium.

The physics of harpsichord has already been investigated in several studies, focusing on its elements [1, 2] as well as on its modal behavior [3, 4] or on its dynamics [5]. Further, the question of the plucking action itself has mostly

been addressed through modeling [6–8]. Besides, the influence of plectrum geometrical parameters on the produced sound has been theoretically [9] and experimentally [10] analyzed for the harpsichord and the classical guitar, respectively. Although few studies have enhanced the harpsichord plucking modeling by measurements [11, 12], no investigation in a musical context has been carried out as for instance for the concert harp [13, 14].

The present paper describes preliminary results from an ongoing exploratory study of the plectrum voicing. We investigate harpsichord plucking action, within the plectrum-string interaction analysis and modeling. For this purpose, we first present the experiment carried out in order to measure plectrum and string motions during playing. Then, we provide a experimentally-based description of the plucking action. Finally, a plectrum-string interaction model is confronted to experimental data.

2. EXPERIMENTAL PROCEDURE

2.1 Experimental protocole

A set of measurements is fulfilled in order to investigate the harpsichord plucking action. The experimental procedure illustrated in Fig. 1 consists in capturing the plectrum-string interaction during a harpsichord performance. As observations indicate that no motion occurs along the string axis (denoted \vec{e}_z in Fig. 2), the analysis is performed in the (\vec{e}_x, \vec{e}_y) plane. For this purpose, plectrum motion is measured through a high-speed camera set at 10000 frames per second focusing on the plectrum. A laser doppler vibrometer is also focusing on the associated key to measure its velocity within the performance. The laser beam is focused at about 3 cm from the tip of the key, avoiding obstruction from the musician hand while depressing the key. Simultaneously, acoustical signals are recorded with a microphone, allowing the synchronization of the database.

A skilled harpsichord player has been asked to perform the sequence presented in Fig. 3 with two various *tempi* which have been estimated at about 230 bpm and 440 bpm. Besides, as the high-speed camera sampling rate induces a restricted resolution of 208px×200px representing a 12.3mm×11.8mm area, we were able to only investigate one note. Based on the harpsichord player advices, G4, which fundamental frequency is 392 Hz, is chosen. This process has been carried out with a plectrum in delrin shown in Fig. 4.

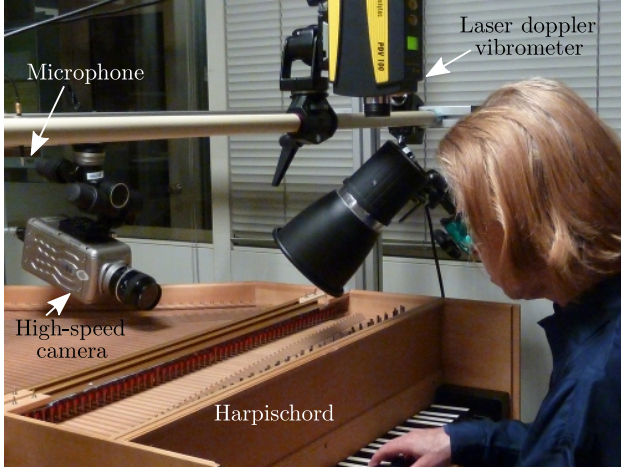


Figure 1. Experimental setup.

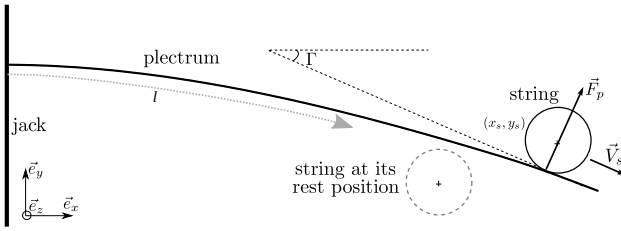


Figure 2. Scheme of the plectrum-string interaction and definition of the parameters.

2.2 Data processing

Fig. 5 shows an example of image obtained through the high-speed camera. The entire sequence is processed to determine a set of plectrum-string interaction descriptors as follow. The desktop background defined at the plectrum rest position is first subtracted from all images. Then, the area of interest containing the plectrum is selected by the user. This template is recursively searched in the entire sequence through a block-matching algorithm model [15]. The string cannot be distinguished in the image because of its contrast. The latter is then refined to obtain a black and white image where only the plectrum appears in white. Hence, based on the framing projection relatively to the measurement plane (\vec{e}_x , \vec{e}_y , see Fig. 2), the trajectory of the string's plucking point (x_s, y_s) in the latter plane can be extracted from the sequence through its shadow on the plectrum.

Further, the synchronization of acoustical and key velocity signals is needed to point out the instrumentalist control. For this purpose, the onset of each sound event of



Figure 3. Score of the performed sequence.

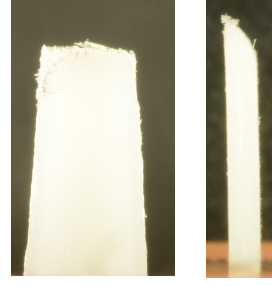


Figure 4. Investigated plectrum: top and side views.

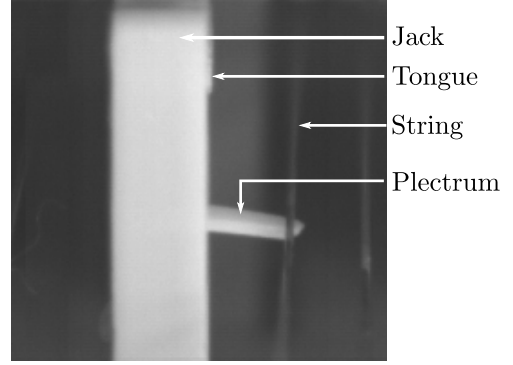


Figure 5. Example of image obtained through the high-speed camera.

the score presented in Fig. 3 is highlighted in the acoustical signal, through a standard onset detection algorithm [16, 17]. These instants indicate the strings' release by the plectrum.

3. HARPSICHORD PLUCKING DESCRIPTION

3.1 Temporal phases

The plectrum is directly governed by the instrumentalist to interact with the string through the harpsichord mechanism. Thus, its motion reveals a part of the musician's control while playing as well as determines the initial condition of the string vibration. The laser doppler vibrometer focused on the pressed key directly conveys the velocity pattern of the plectrum at its connection point with the tongue. This investigation combined with the observation of the plectrum motion through the high-speed camera lead to a description into four temporal phases. Based on a harp plucking analysis [13], it can be decomposed as follow:

- The preparatory movement: the plectrum raises from its rest position and approaches the string, $\forall t \in [t_i, t_c]$. Usually, the player first pushes the key until reaching the string to feel its contact with the plectrum. The plectrum is then kept just below the string.
- The sticking phase: the plectrum moves the string in the vertical direction from its rest position, $\forall t \in [t_c, t_s]$;
- The slipping phase: the string slips on the plectrum surface towards its free end, $\forall t \in [t_s, t_r]$;

- The free oscillations of the string, $\forall t > t_r$.

The determination of each instant t_c , t_s and t_r is based on high-speed camera films. The sticking instant t_c is automatically detected at the first instant where the string is moved up by the plectrum. Then, the beginning of the slipping phase t_s is defined when the string's displacement gets a horizontal component. Finally, the release instant is obvious to determine since it corresponds to the beginning of the plectrum vibrations. Note that t_r is manually estimated. Although the detection of t_c and t_r are straightforward, t_s is more complicated to highlight.

In the present case, we estimate the entire duration of the harpsichord plucking action at 12 ± 2 ms and at 7.9 ± 0.6 ms for the normal and the fast *tempi*, respectively. The measurements reported in the literature are about 80 ms [11] and 25 ms [12]. Differences between these orders of magnitude can be explained by the various musical contexts considered. Further, the slipping phase duration has not been pointed out in these previous studies. However, based on the plucking analysis, we measure that the slipping phase lasts about 36 ± 3 % and 26 ± 4 % of the entire plectrum-string interaction duration for the normal and the fast *tempi*, respectively. The small uncertainties, computed over each played note with a 95 % confidence interval, tend to indicate that the measurement protocol is reliable. Eventually, remark that the first note has not been taken into account since it corresponds to the instrumentalist adaptation to the plectrum, which implies a longer interaction: 26 ms, i.e. about two times longer than the following plucking actions.

3.2 Jack velocity

The jack motion is directly measured through the high-speed camera during the plectrum-string interaction. The linear regression associated to its vertical displacement indicates that the jack velocity is constant during plucking. Indeed, the computed coefficients of determination are 0.97 ± 0.02 , and 0.99 ± 0.00 for the normal and the fast *tempi*, respectively. Eventually, the jack velocities are estimated in the two former contexts at 0.10 ± 0.05 m/s and 0.25 ± 0.03 m/s. These estimations are relevant regarding the *tempo* of the played sequence: the higher the *tempo*, the higher the jack velocity. Besides, it confirms that the instrumentalist control is repeatable along a musical sequence, as expected for an expert gesture [18]. Finally, these measurements are consistent with those previously measured or used to feed a harpsichord plucking model [8, 11]. However, note that these values are highly dependent on the plectrum's shape and the material as well as on the string's properties.

Because of the arduousness of the high-speed camera films post-processing, the laser doppler vibrometer is used to get an insight on the instrumentalist's control during the entire sequence. Fig. 6 shows the G-key velocity curve measured on one plucking action in a musical context. This curve has been chosen because of its representativeness among the whole database. It indicates that the player depresses first the key until the plectrum reaches the string $\forall t \in [t_i; t_c]$.

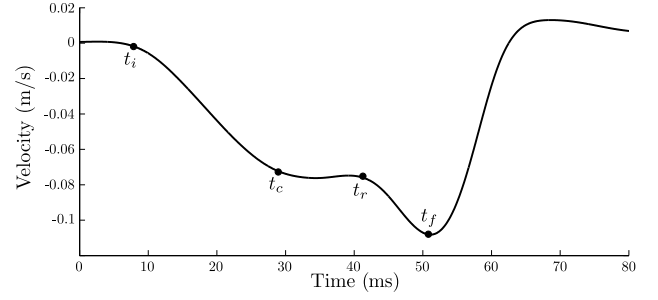


Figure 6. G-key velocity curve during a plucking action and its preparation. t_i corresponds to the instant where the finger starts pushing the key. The plectrum-string interaction begins at t_c , while t_r and t_f are the instants where the string and the key are released, respectively.

Then, during the plucking action ($\forall t \in [t_c; t_r]$), the plectrum velocity presents a second shape. Eventually, from the instant the string is released until t_f , the plectrum velocity presents the same pattern than during the first phase. After this last instant, the finger releases the key toward its rest position.

3.3 Initial condition of the string vibration

Although the free oscillations of the string do obviously not convey all the informations relative to the radiated sound, they reveal variations between plucking actions. Indeed, assuming the string flexible, stretched to a tension T , of uniform linear density ρ_l and fixed at its end, the transverse vibration of the string along the vertical direction can be described by [19, 20]

$$\vec{r}(z, t) = \sum_{n=1}^{\infty} (A_n \cos(2\pi f_n t + \Psi_n) + B_n \sin(2\pi f_n t + \Psi_n)) \vec{\Phi}_n e^{-\alpha_n t}, \quad (1)$$

where $f_n = nc/2L$ are the eigenfrequencies, Φ_n are the modal deflections, α_n are the damping coefficients and $\Psi_n = \arctan \alpha_n / \omega_n$. The modal amplitudes A_n and B_n are written as

$$A_n = \frac{2D_{t_r} \sin(k_n z_0)}{k_n^2 z_0 (L - z_0)}, \quad (2)$$

$$B_n = \frac{2V_{t_r} \sin(k_n z_0)}{k_n^3 z_0 (L - z_0)c}, \quad (3)$$

with k_n the wavenumber, and z_0 the plucking point along the length of the string. $D_{t_r} = \sqrt{x_s(t_r)^2 + y_s(t_r)^2}$ and V_{t_r} are the initial position and velocity of the string regarding its rest position (see Fig. 2), respectively. Further, the magnitude of each polarization of the string is of great importance relatively to the radiated sound. It can be estimated by using the initial angle between the two components $x_s(z_0, t)$ and $y_s(z_0, t)$ of the transverse string motion during the first vibration instants. It is referred to as Γ_{t_r} , as shown in Fig. 2.

Tab. 1 presents the initial conditions of the string's vibrations for the two performed sequences. Let us remind

	Slow	Fast
$x_s(t_r)$ (mm)	0.27 ± 0.02	0.28 ± 0.02
$y_s(t_r)$ (mm)	0.71 ± 0.07	0.90 ± 0.01
V_{t_r} (m/s)	0.06 ± 0.02	0.15 ± 0.01
Γ_{t_r} (deg)	17 ± 2	20 ± 1

Table 1. Initial conditions of the string vibration. The mean is computed on eight samples. The reported uncertainty represents a 95% confidence interval.

that the jack velocity was substantially increased for the faster sequence. As expected, the *tempo* does not influence the string position in the horizontal direction: this value is completely determined by the initial position of the string relatively to the jack. However, results indicate that the more the jack velocity, the higher the string is released. In the present case, we measure an elevation of the string of about 27 % in the case of the faster *tempo* relatively to the slower one. This obviously implies a higher string's velocity magnitude as well as a more important force applied by the string on the plectrum's end. The latter directly conveys to a higher plectrum's deformation and therefore to a more important initial polarization angle Γ_{t_r} .

4. COMPARISON BETWEEN MEASUREMENTS AND MODELING

In this section, a comparison between a mechanical model of the plectrum-string interaction and experimental results is presented.

4.1 Plectrum-string interaction model

The model of the plectrum-string interaction presented in this section is based on that of Perng and colleagues [7, 8, 12]. In this model, based on the assumption that no friction occurs between the string and the plectrum, the string is seen as a punctual force applied perpendicularly to the plectrum. The latter is considered as a cantilever beam with small strain, made in an isotropic material, and without any twisting motion. Another assumption is the uniformity of the plectrum section. However, as shown in Fig. 4, the plectrum is beveled under its tip to slip on the string when the jack falls down. This thickness variation induces a modification of the moment of inertia at the tip of the plectrum. In order to take it into account in the model, the plectrum is considered as a beam composed of two segments: the first one from 0 to L_1 with a moment of inertia of I and the second one from L_1 to L with a moment of inertia of $I/8$. With these assumptions, the deflection angle of the plectrum is

$$\begin{aligned} l \in [0 - L_1], \Gamma(l) &= \frac{F}{EI} \left(Ll - \frac{l^2}{2} \right), \\ l \in]L_1 - L], \Gamma(l) &= \frac{F}{EI} \left(\frac{7}{2}L_1^2 - 7LL_1 + 8Ll - 4l^2 \right), \end{aligned} \quad (4)$$

where E is the Young modulus of the plectrum material and F the force applied by the string at a distance l to the plectrum (see Fig. 2). The second Newton's law is applied

String length L_s	47.5cm
String tension T	85.7N
Plectrum Young Modulus E	1.5GPa
Plectrum length L	3.5mm
Beginning of the bevel L_1	2.9mm
Moment of inertia I	0.011mm^4

Table 2. Parameters of the plectrum and of the string.

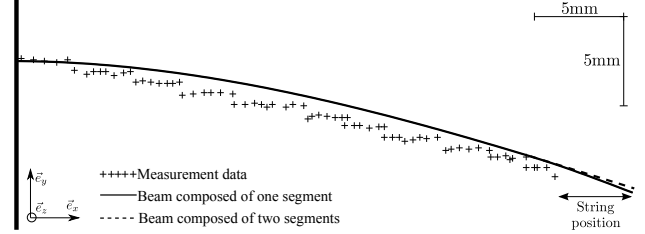


Figure 7. Experimental and theoretical plectrum deflection at the release instant. Referred axes are the same as in Fig. 2.

on the string's element in contact with the plectrum during the interaction. Its projection along x -axis and y -axis are written as

$$\Delta m \frac{\partial^2 x_s}{\partial t^2} = -Kx_s + F_{px}(l), \quad (5)$$

$$\Delta m \frac{\partial^2 y_s}{\partial t^2} = -Ky_s + F_{py}(l). \quad (6)$$

With a geometrical approach, K is computed from the string's tension T , length L_s and the plucking ratio β using $K = T/(L_s\beta(1 - \beta))$. Eqs. 5 and 6 indicate that the plectrum-string interaction depends on the string's position on the plectrum l . This position can be extracted from the jack displacement (x_j, y_j) by computing at each time step the following system of equations (with the small angle approximation):

$$x_s(t) - x_j(t) = \int_0^l 1 - \frac{1}{2}\Gamma(x)^2 dx, \quad (7)$$

$$y_s(t) - y_j(t) = -\int_0^l \Gamma(x) - \frac{1}{6}\Gamma(x)^3 dx. \quad (8)$$

In this way, at each time, the string position (x_s, y_s) , the plectrum deflection Γ and the plectrum force F_p may be known.

4.2 Results

In order to compare experimental data to theoretical ones, the model is computed with parameters given in Tab. 2.

In Figure 7, the plectrum deflection at the release instant is shown. Unfortunately, all the plectrum's shape cannot be extracted because, at this instant, the string hides the plectrum tip. Nevertheless, theoretical results are found to be consistent with the experimental data. In this figure, we also present the model based on a beam composed of

one segment of moment of inertia I with the one with two segments (as explained in Sec. 4.1). The proposed refinement induces a very slight modification of the plectrum shape prediction. However, the comparison of the plectrum deflection angle ($\Gamma(L) = \Gamma_{t_r}$) for the two configurations indicate that for an unique segment the angle is 13.4° whereas for two segments it is 20.9° . The second result is closer to the experimental value (see Fig. 1). The beam composed of two segments is thus valuable to model the harpsichord plectrum.

Based on the plectrum-string model, the initial conditions of the string vibration can be computed. At the release instant, we obtain $x_s(t_r) = 0.31$ mm and $y_s(t_r) = 0.80$ mm. These two values are very close to the experimental ones, especially for the fast case (see Tab. 1). Note that our results are obviously independent of the jack velocity because no friction force is taken into account in the model. Therefore, when the string slips fast on the plectrum, less friction forces occur and the model is found to be more appropriate than for the slow case.

5. CONCLUSION

This paper has presented an investigation of harpsichord plucking action. A well-controlled measurement setup conveying the jack and plectrum motion features as well as the produced sound has been carried out in a musical context.

As expected, results indicate that the plectrum-string interaction is longer when the playing *tempo* is slower. This is directly related to the jack velocity during the interaction. Indeed, a slower *tempo* leads to a lower jack velocity while it is in contact with the string. Besides, it has been shown that the plectrum velocity is constant during the plectrum-string interaction. Although these two parameters can be considered as control parameters, the instrumentalist cannot directly act on the initial conditions of the string vibration, which highly contribute to define the produced sound. However, a relationship has been pointed out between the playing *tempo* and the initial position, velocity and initial angle of the string's free oscillations. Indeed, we observed that the string is released higher relatively to its rest position for a higher playing velocity, which obviously imply a higher initial string's velocity. Moreover, the string applies therefore a more important force on the plectrum, conveying to a higher plectrum bending and then variation in the initial string's oscillations polarization. A deeper investigation on several plectrum's shape and material is required to conclude about the implications of these observations on the produced sound. Further, a plectrum-string interaction model has been implemented and confronted to our measurements. Simulations results indicate that the variation of the plectrum geometry along its length has to be taken into account in order to well-reproduce its deformations throughout the plucking action. This is of great importance to accurately predict the initial conditions of the string's vibration. The model is found particularly effective to give results very close to experimental ones, especially when the player plays rapidly.

Further work will be carried out to investigate harpsichord voicing. Several plectra of various shapes will be

used for the measurements. A parametric study on the plectrum's geometrical parameters will be helpful to point out the relationship between the voicing process and the produced sound as well as the instrumentalist control. Further, refinements of the current plectrum-string interaction model will be proposed. For instance, the friction forces during the slipping phase will be modeled in order to improve the prediction of the initial conditions of the string free oscillations. Besides, as the plectrum-string interaction is obviously dependent on the musical context further comparative analyses will be performed on various playing techniques and music styles.

Acknowledgments

The authors acknowledge the Ministère de la culture et de la communication for its financial support to the project, as well as Jean-Claude Battault and Stéphane Vaiedelich for the fruitful discussions, Jean-Marc Fontaine for taking the pictures of the plectrum, Laurent Quartier for his technical help and Hugo Scurto for the preliminary tests.

6. REFERENCES

- [1] M. Spencer, "Harpsichord physics," *The Galpin Society Journal*, vol. 34, pp. 2–20, 1981.
- [2] N. Fletcher, "Analysis of the design and performance of harpsichords," *Acustica*, vol. 37, pp. 139–147, 1977.
- [3] S. Le Moyne, S. Le Conte, F. Ollivier, J. Frelat, J.-C. Battault, and S. Vaiedelich, "Restoration of a 17th-century harpsichord to playable condition: A numerical and experimental study," *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp. 888–896, 2012.
- [4] S. Le Conte, S. Le Moyne, F. Ollivier, and S. Vaiedelich, "Using mechanical modelling and experimentation for the conservation of musical instruments," *Journal of Cultural Heritage*, vol. 13, no. 3, pp. 161–164, 2012.
- [5] H. Penttinen, "On the dynamics of the harpsichord and its synthesis," *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, pp. 115–120, 2006.
- [6] D. Griffel, "The dynamics of plucking," *Journal of Sound and Vibration*, vol. 175, no. 3, pp. 289–297, 1994.
- [7] C.-Y. Perng, J. Smith, and T. Rossing, "Physical modeling of the harpsichord plectrum-string interaction," *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)*, 2010.
- [8] —, "Harpsichord sound synthesis using a physical plectrum model interfaced with the digital waveguide," *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, pp. 329–335, 2011.
- [9] C. Sloane, "The effects of plectrum parameters on harpsichord sound," *Journal of Sound and Vibration*, vol. 125, no. 1, pp. 185–186, 1988.

- [10] S. Carral and M. Paset, "The influence of plectrum thickness on the radiated sound of the guitar," in *Proc. Acoustics '08*, Paris, France, 2008.
- [11] N. Giordano and J. Winans II, "Plucked strings and the harpsichord," *Journal of Sound and Vibration*, vol. 224, no. 3, pp. 455–73, 1999.
- [12] C.-Y. Perng, "Physical modeling of the harpsichord plectrum-string interaction," Ph.D. dissertation, Stanford University, 2012.
- [13] D. Chadeaux, J.-L. Le Carrou, B. Fabre, and L. Daudet, "Experimentally based description of harp plucking," *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp. 844–855, 2012.
- [14] D. Chadeaux, J.-L. Le Carrou, and B. Fabre, "A model of harp plucking," *J. Acoust. Soc. Am.*, vol. 133, no. 4, p. 24442455, 2013.
- [15] S. El-Azim, "An efficient object tracking technique using block-matching algorithm," in *Radio Science Nineteenth National Conference of the Proceedings of NRSC*, Alexandria, Egypt, 2002, pp. 427–433.
- [16] C. Duxbury, J. Bello, M. Davies, and M. Sandler, "Complex domain onset detection for musical signals," *6th International Conference on Digital Audio Effects (DAFx)*, vol. 1, pp. 90–94, 2003. [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.57.9197&rep=rep1&type=pdf>
- [17] J. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. Sandler, "A tutorial on onset detection in music signals," *IEEE Transactions on Speech and Audio Processing*, vol. 13, pp. 1035–1047, 2005.
- [18] D. Chadeaux, "Interaction musicien/instrument : Le cas de la harpe de concert. (musician/instrument interaction: The case of the concert harp.)," Ph.D. dissertation, Université Pierre et Marie Curie, 2012.
- [19] N. Fletcher and T. Rossing, *The Physics of Musical Instruments*, 2nd ed. New York, US: Springer, 1998.
- [20] A. Chaigne and J. Kergomard, *Acoustique des instruments de musique*. BELIN, 2008.

Bio-inspired robot to study stringed instruments: application to the harp

Delphine Chadeaux

LAM / D'Alembert,
UMR CNRS 7190
UPMC Univ Paris 06

Alexandre Roy

LAM / D'Alembert, UMR CNRS 7190
ISIR, UMR CNRS 7222
UPMC Univ Paris 06

Jean-Loïc Le Carrou

LAM / D'Alembert,
UMR CNRS 7190
UPMC Univ Paris 06

Marie-Aude Vitrani

ISIR, UMR CNRS 7222
UPMC Univ Paris 06

Benoît Fabre

LAM / D'Alembert,
UMR CNRS 7190
UPMC Univ Paris 06

ABSTRACT

This paper is a review of a previous musician/harp interaction study leading to the specifications of a configurable excitatory mechanism. Such robotic tools can be valuable to investigate the musician/instrument interaction and the mechanical behavior of an instrument in a repeatable and realistic musical context. To design a robotised excitatory mechanism for the concert harp, the mechanical descriptors' typical orders of magnitude defining a musical performance have to be highlighted. For this purpose, two experimental setups have been designed. The first one focuses on the sound-producing gesture (i.e. on the plucking action). A high-speed camera has been used to accurately measure the finger and string motions in a realistic musical context. The second measurement set-up consists in capturing the whole harpist's body motion through infra-red cameras during a performance. The set of mechanical parameters extracted from these measurements is of great interest to get an insight on an ideal robotic tool to pluck harp strings. Based on these considerations, a highly-configurable and repeatable robotic finger has been designed.

1. INTRODUCTION

A configurable and repeatable excitatory mechanism may be particularly adapted for the study of musical instruments or of the musician/instrument interaction. For wind instruments, researchers designed artificial mouths to play the trumpet [1], the trombone [2] or the clarinet [3]. These systems may be automated [3,4]. For stringed instruments, classical mechanical excitatory systems used in the literature are wires rolled around the strings pulled up until they break or mechanical plectrums. They can generate repeatable initial conditions applied to the string but cannot reproduce the complexity of the plucking gesture and therefore the typical initial conditions. To do so, the idea would be to use a bio-inspired robot. This kind of robots are

used to develop *virtuous* instrumentalist as Waseda flutist robot [5] but not to exactly reproduce the instrumentalist action. For plucked instruments, this action is performed by a finger or a plectrum and imposes particular initial conditions to the string. Designing bio-inspired robot needs above all an accurate description of the instrumentalist gesture. In the light of recent studies on the harp, a robot called DROPIC (Doigt RObotisé PInceur de Cordes) [6] has been designed to reproduce the musicians' expert gestures. The goal of the present paper is to suggest the development steps of a bio-inspired musical robot from former studies presented in [7], [8] and [9] and to specify its limitations.

The paper is organized as follows: first a summary of the plucking action in the case of the harp will be presented. Then an analysis of this gesture will draw the DROPIC specifications. Eventually the final design and the choices made during the development of the robot will be presented.

2. ANALYSIS OF THE INSTRUMENTALIST GESTURE

Playing a musical instrument is a complex task, relying on a long training by the player. From the instrumentalist's point of view, the performance may be seen as series of gestures, that turn the musical intention into movements and results in sound production. The first investigation in the study of the playing is the segmentation of the gestures and movements. While the writing of the music on a score suggests a note based segmentation, the analysis of everyday practice of players suggests that scales, arpeggios, melodic and harmonic grouping of notes may be more significant.

2.1 Finger trajectory analysis

In the playing of a note or of a chord on a harp, several time periods can be identified. Starting from the postural and gestural preparation, the player brings his/her hand and fingers in contact with the strings, pulls them until the force applied by the finger reaches a threshold. This phase is called the sticking phase. Before the string is released with controlled initial oscillating conditions, it slips on the finger pulp of the musician, this is called the slipping phase.

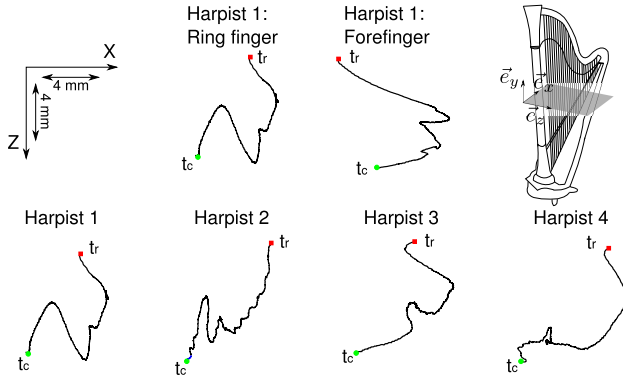


Figure 1. Example of finger trajectories. The instants denoted t_c and t_r correspond to the sticking phase beginning and the string release, respectively. The four bottom figures are trajectories extracted from an arpeggio sequence played by four harpists with the ring finger. The top-left and top-right figures presents finger trajectories extracted from an arpeggio sequence played by the first harpist with the ring finger and the forefinger, respectively.

Works presented in [8], [9] and [10] focus on the movements of the finger and string during the finger-string interaction. In these papers, finger's and string's trajectories are measured in the plane perpendicular to the strings direction (\vec{e}_x, \vec{e}_z) defined in figure 1. The samples given in figure 1 show harpists' finger trajectories projected in this plane. They indicate that non negligible differences occur between harpists for a given musical context. Besides, considering one harpist, the finger pattern depends on the finger and on the technique used to pluck the string [10].

The estimation of mechanical parameters needed to design the robotic finger are determined according to [6, 10]. They are summarized in Table 1. First, the peak force applied by the harpist's finger to the D \flat 3 string has been estimated up to 20 N, regardless of the musical context. Further, figure 2 presents the maximal D \flat 3 string displacement during the ten harpists' ring finger and forefinger plucking actions. Let us note that no significant influence of the musical context on this descriptor has been outlined. Although the trajectories are specific to each harpist, figure 2 tends to indicate that all the maximum displacements D_{max} given to the string are contained in a 24 mm radius quadrant.

Meanwhile the maximum fingertip velocity has been estimated to be less than 1.5 m/s [6].

Criterion	Value
maximum velocity	1.5 m/s
nominal force	20 N
area of use	4 cm ²

Table 1. Criteria the robot should match to correctly reproduce the motion of a musician finger

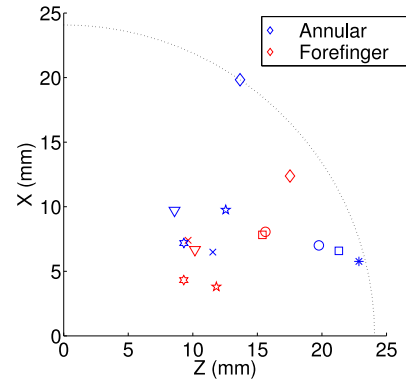


Figure 2. Maximum displacement of a harp string for the ten harpists depending on the finger used (forefinger or ring finger). Each harpist is denoted by a symbol. The blue and the red ones referred to the ring finger and the forefinger displacements, respectively. The dashed line represents the circle which radius is the maximum displacement for all harpists.

2.2 Arm gesture analysis

Although the description of the finger gesture is valuable to understand the plucking action itself, it is not sufficient for a thorough comprehension of a harp performance. Indeed, the “actuators” of the movement, i.e. the muscles and articulations of the arm, have to be analyzed too.

The research performed in [11] with a motion capture system presents the movement analysis of the whole harpist's body during a performance. Three harpists have been asked to perform the beginning of Debussy's *Danse Profane*. Efforts in the upper limbs (hand, arm and forearm) are deduced from kinematic measurements using a well-known human mass distribution model [12]. Figure 3 shows the computational process [11]. First, infra-red cameras accurately provide the 3D-position of reflective markers fixed on the musician in a global reference system. Then anthropometric data lead to the trajectories of each limb's center of mass. Besides, as they are not directly reachable using the experimental set-up, the more distal segment moment and force have to be estimated. The moment is neglected and the force is found to be between 2 N and 8 N (i.e. the minimal and maximal forces measured for a chord performance [10]) according to the sound level of each note. Then, a link-segment model [12] leads to the forces and moments at the first proximal joint. This process is recursively done to determine the forces \vec{F} and moments \vec{M} occurring at each arm joints. Each joint's mechanical power are computed from $P = \vec{M} \cdot \vec{\omega}$ where $\vec{\omega}$ is the angular velocity of the limb's center of mass. Finally, the work done by the group of muscles related to each limb-segment is computed by time-integrating the power curve under the considered movement duration.

Results presented in figure 4 indicate that the shoulder controls the entire arm gesture while performing the octave intervals contained in the beginning of Debussy's *Danse Profane*.

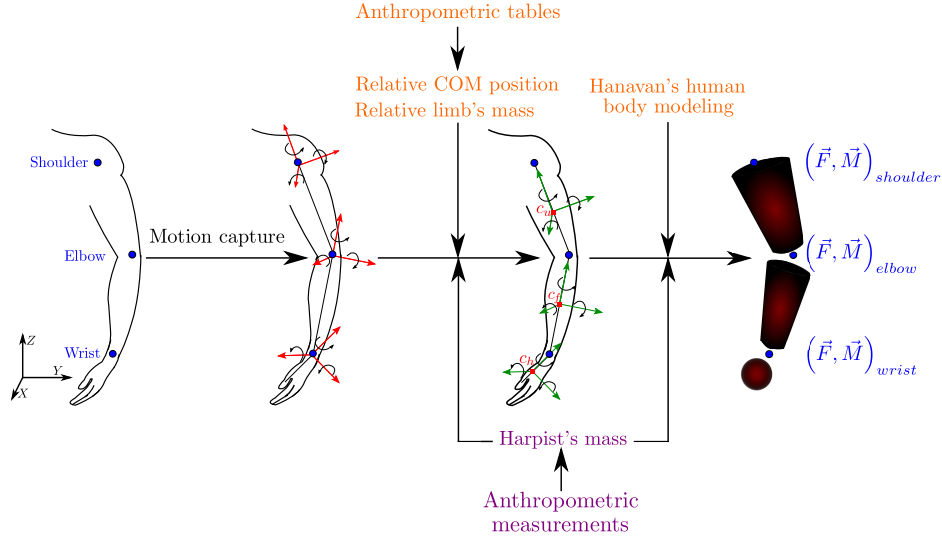


Figure 3. Retrieval process of the forces and moments occurring at a harpist's arm joints while playing octave intervals.

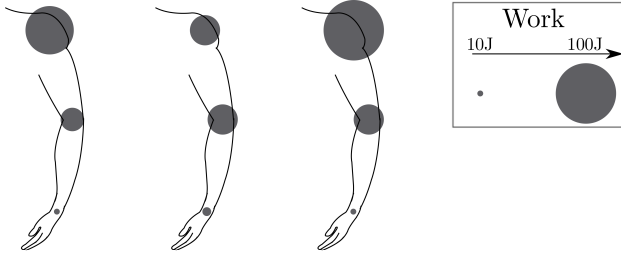


Figure 4. Average work estimated at each arm joint for the three involved harpists.

3. ANALYSIS OF THE ROBOT CAPABILITIES

In the present section, an ideal behavior of the robot is presented, according to the plucking action analysis. Then the design of the robot will be done to meet the specifications. Finally, a fingertip will be designed to reproduce the harpists finger pulp and thus the initial conditions provided to the string.

3.1 Robot specifications

Table 1 points out three parameters the robot must reproduce to provide to the string the same initial conditions as a musician. Besides, information provided in [10] about the finger's trajectories duration is required. This study indicates that the typical sticking and slipping phases durations are about 300 ms and 3 ms, respectively. Obviously, details about the forces involved in harp plucking are also required. For this, the harp strings' tension is needed. It has been computed from formulas defined in appendix A and presented in figure 5. Thus, assuming each string is plucked at the third of its length with an initial displacement corresponding to D_{max} (see Subsection 2.1), the maximal efforts needed to pluck the whole set of the harp's strings are deduced and shown in figure 5. Note that the strings parameters used for this estimation are those of a

concert harp (Camac, *Atlantide Prestige* model). This harp has 47 strings, and its *tessitura* extends from Cb0 (30.8 Hz) to Gb6 (2960 Hz).

Figure 5 indicates that the strings can be split into two groups. The first corresponds to the wrapped, nylon and a few gut strings (i.e. from Cb0 (30.8 Hz) to Gb1 (92.5 Hz) and from Fb4 (698.5 Hz) to Gb6 (2960 Hz)) for which the effort is comprised between 20 N and 40 N. The second group corresponds to the other gut strings (i.e. from Ab1 (104 Hz) to Eb4 (622 Hz)), for which the effort is comprised between 10 N and 20 N. For practical reasons, the first set of strings was discarded from the present study, implying that the maximum effort the robot has to apply on the strings is 20 N.

3.2 Design of the robot

Based on the previous musician/harp interaction analysis (see section 2), the shoulder has to develop important efforts to allow the finger to pull the string with a 20 N force. In the case of a bio-mimetic robot design, all the involved efforts and arm's dimensions should be taken into account. However, in the present case, considering the fingers trajectories are almost contained in a plane, we can reduce the robot to a single finger to limit its size. The basic idea is to design a planar robot with two rotational joints that mimic a human finger [16]. Its dimensions are l_1 and l_2 and it is parametrized by the angles ϑ_1 and ϑ_2 , as shown in figure 6. The numerical values for l_1 and l_2 are arbitrarily inspired from a real finger, because they lead to a small robot that can deliver the desired efforts, and are:

$$\begin{cases} l_1 &= 6.21 \text{ cm}, \\ l_2 &= 4.75 \text{ cm}. \end{cases}$$

Using these previous parameters, a classical static analysis of the robot provides a relationship between the actuators parameters, i.e. the motor torques C_{m1} and C_{m2} for each joint, the geometrical parameters of the robot and the

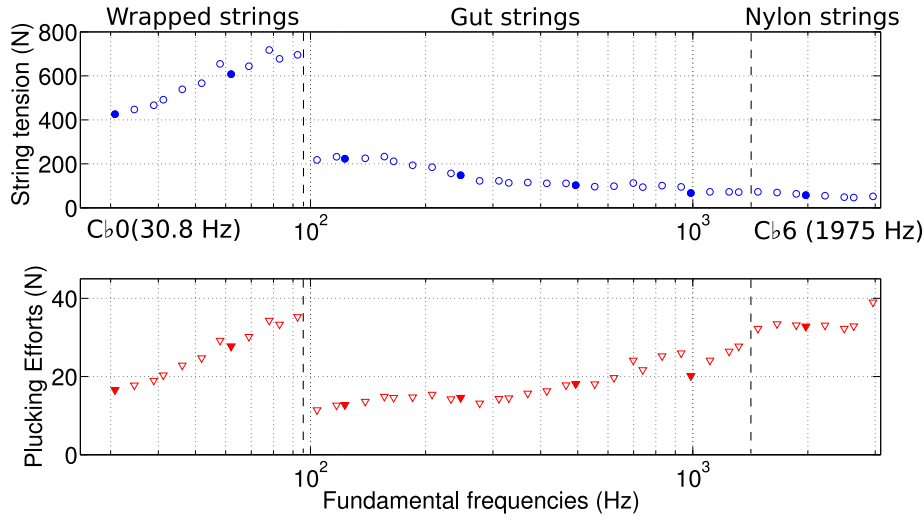


Figure 5. String tension (upper plot) over the concert harp compass. Plucking effort for each string (lower plot) in order to reach a displacement of 24 mm. String tensions are estimated from pitch and mass measurements. The plain markers on the figure show the C_b strings.

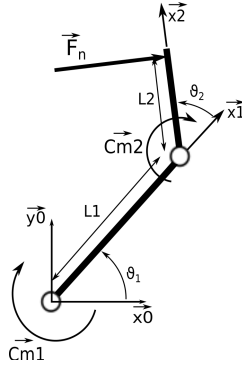


Figure 6. Parametrization of the robotic finger.

maximum force $F_{n \max}$ as follows:

$$\begin{cases} C_{m1} = l_1 F_{n \max}, \\ C_{m2} = l_2 F_{n \max}. \end{cases}$$

As the maximum effort $F_{n \max}$ is chosen to be up to 20 N (see subsection 2.1), the motor torques numerical values are estimated to

$$\begin{cases} C_{m1} = 1.21 \text{ N.m}, \\ C_{m2} = 0.926 \text{ N.m}. \end{cases}$$

Based on these values, the motors chosen to actuate the joints are two Maxon motors, model RE35, equipped with two 12:1 Maxon GP42 reducers, which leads to a practical maximal torque for each actuator:

$$C_{m1} = C_{m2} = 1.26 \text{ N.m}.$$

The choice of the actuators and of the geometry of the robot allows the maximum speed at the end effector at $V_{\max} = 1.16 \text{ m/s}$. Moreover, at this maximum speed, the

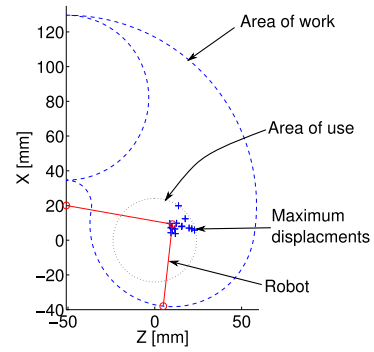


Figure 7. Area of work of the robot and area of work specified. The maximum distances are reminded on the graph with '+' symbol.

duration Δt takes by the robot to do a straight line with a length of D_{\max} is

$$\Delta t = \frac{D_{\max}}{V_{\max}} \approx 20 \text{ ms}.$$

This value is lower than the duration of the trajectory. Therefore, the robot meets all the specifications.

Figure 7 shows the area of work of the robot which contains the area of use specified in subsection 3.1. The origin of the robot can be adjusted to contain its whole area of use. Thus, the robot and its actuators can reproduce the trajectories measured on the musicians. It fulfills the requirements to mimic finely the finger's trajectory during a plucking action.

3.3 Fingertip material

In a previous study [16], the shape and the material of a fingertip was investigated to reproduce human finger pulp

behavior. Silicone was chosen because of its friction coefficient, which is close to the human skin [11].

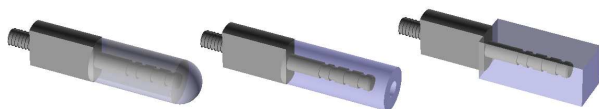


Figure 8. Fingertip shapes enhancing the aluminium end effector of the robot.

Figure 8 presents several fingertips molded with silicone. During the moulding process, a set of parameters can be controlled: the mold's shape, the quantity of filler and oil added to the mixture. Let us note that the filler increases the hardness of the material while the oil decreases it. Figure 9 shows a picture of DROPIC, enhanced by a silicone fingertip, mounted on the concert harp.

An experiment presented in [16] has been carried out to validate the robotic finger and to point out the best fingertips shape and hardness. For this purpose, a harpist has been asked to perform isolated plucking actions. Then, the extracted finger's path has been reproduced by the robotic finger enhanced with the various fingertips presented in figure 8. The resulting string's trajectories as well as the generated soundboard vibration have been compared to those produced by the harpist. Results indicate that the harder the fingertip is, the better the trajectory is followed. Moreover, better results have been obtained for fingertips with a cylindrical shape, a round extremity and a hardness close to 30 shore A. Eventually, let us remark that the silicone's drawback is the loss of its mechanical properties over time. Therefore, new materials are currently investigated to enhance the robotic finger as for instance polyurethane.

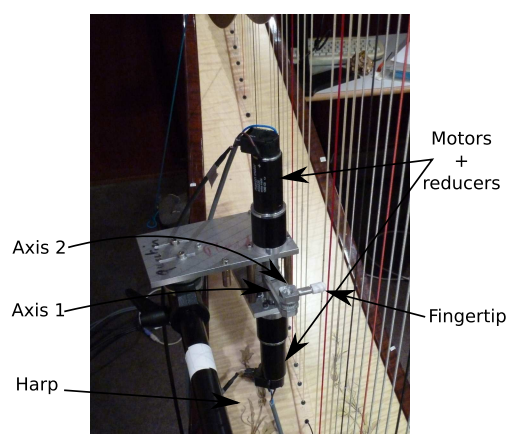


Figure 9. The robot DROPIC mounted on the harp for measurements.

4. CONCLUSION

This paper has presented the developing steps of a bio-inspired musical robot based on a review of a previous

musician/interaction analysis. The proposed methodology is carried out to conceive a robot designed to pluck harp strings as harpists do. A phase of simplification of experimental results is necessary to propose specifications of the robot. Therefore, the analysis presented in [8], [9], [10] and [11] are based on simplified models of the finger-string interaction: the motion of the finger is only analyzed in the plane perpendicular to the string and the torsion of the string during the contact is ignored. Observation of actual plucking action shows that the rotation of the finger around the contact point as well as the return force due to string torsion may be decisive in the triggering of the slipping of the string on the finger. The development of a two axes robot for string plucking is therefore a first step in understanding the influence of the different elements (finger motion, finger elastic behavior or hardness, skin contact...) that govern the initial oscillating conditions on the string. This robot can be used to gather dynamical information on the forces, both normal and tangential to the contact surface, and therefore on the triggering of the slipping of the string on the finger. Future versions may include more complex motions such as those observed in actual human plucking. Moreover, this methodology can now be applied to others string instruments to analyse the musician/instrument interaction.

Acknowledgments

The authors acknowledges Laurent Quartier, for his participation in the setup of the experiments as well as his precious advice, and Sylvère Billout for his investment in developing the first version of the robot.

5. REFERENCES

- [1] J-F. Petiot, F. Teissier, J. Gilbert and M. Campbell. "Comparative analysis of brass wind instruments with an artificial mouth: first results", *Acustica* 89, pp.974-979 (2003).
- [2] J. Gilbert, Sylvie Ponthus and J-F. Petiot. "Artificial buzzing lips and brass instruments: Experimental results", *J. Acoust. Soc. Am.* 104(3), pp.1627-1632 (1998)
- [3] D. Ferrand and C. Vergez. "Blowing machine for wind musical instrument: toward a real-time control of the blowing pressure", 16th IEEE Mediterranean Conference on Control and Automation, Ajaccio, France (2008), pp.1562-1567
- [4] D. Ferrand, C. Vergez, B. Fabre, F. Blanc, "High-precision regulation of a pressure controlled artificial mouth: the case of recorder-like musical instruments", *Acustica united with Acta Acustica* 96, pp.700-711 (2010)
- [5] J. Solis, K. Chida, K. Taniguchi, S.M. Hashimoto, K. Suefuji and A. Takanishi, "The Waseda Flutist Robot WF-4RII in Comparison with a Professional Flutist", *Computer Music Journal* 30(4), pp.12-27 (2006)

- [6] J-L. Le Carrou, D. Chadeaux, M-A. Vitrani, S. Billout and L. Quartier. “DROPIC: A tool for the study of string instruments in playing conditions”, *Proceedings of Acoustics 12*, Nantes, France (2012), pp.451-456.
- [7] J-L. Le Carrou, F. Gautier, F. Kerjan and J. Gilbert. “The string-finger interaction in the concert harp”, in *Proceedings of ISMA*, Barcelone, Spain (2007)
- [8] J-L. Le Carrou, E. Walhem, E. Brasseur and J. Gilbert. “Two dimensional finger-string interaction in the concert harp”, in *Proceedings of Acoustics 08*, Paris, France (2008), pp.1495-1500
- [9] D. Chadeaux, J-L. Le Carrou and B. Fabre. “A model of harp plucking”, *J. Acoust. Soc. Am.* 133(4), pp.2444-2455 (2013)
- [10] D. Chadeaux, J-L. Le Carrou, B. Fabre and L. Daudet. “Experimentally based description of harp plucking”, *J. Acoust. Soc. Am.* 131(1), pp.844-855 (2012)
- [11] D. Chadeaux, “Interaction musicien / instrument: le cas de la harpe de concert. (Musician/Instrument Interaction: The case of the concert harp.)”, PhD Thesis, Université Pierre et Marie Curie, (2012)
- [12] D.A. Winter, “Biomechanics and Motor Control of Human Movement”, 4th ed., John Wiley & Sons, Inc. Hoboken, New Jersey, 2009.
- [13] C. Valette and C. Cuesta. “Mécanique de la corde vibrante”, *Traité des Nouvelles Technologies série Mécanique*, Hermes, Paris, France (1993), pp.27-45.
- [14] I. M. Firth, C. Sykes. “Harp Strings of the Eighteenth and Twentieth Centuries: An Acoustical Comparison”, *The Galpin Society Journal*, Vol. 43 (1990), pp.46-59.
- [15] J. Chabassier. “Modélisation et simulation numérique d’un piano par modèles physiques”, PhD Thesis (2012), pp.51-56.
- [16] D. Chadeaux, J-L. Le Carrou, M-A. Vitrani, S. Billout and L. Quartier. “Harp plucking robotic finger”. in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Portugal (2012), pp.4886-4891.

A. FORMULAS FOR THE STRING TENSION

For a i^{th} string fixed at both ends its tension is [13]:

$$T_i = 4\rho_{li}L_i^2f_i^2,$$

where ρ_{li} its linear density, L_i its length and f_i its fundamental frequency. For the case of the harp, strings are in gut, in nylon or wrapped. For the latter, an equivalent linear density ρ_l can be deduced as follows [14, 15]:

$$\rho_l = \rho_{vw}\frac{\pi}{4}D^2 + \left[\rho_{vc}\frac{\pi}{4} - \rho_{vw}\left(\frac{\pi}{4}\right)^2\right]d^2,$$

where ρ_{vw} is the density of the wrapping element, ρ_{vc} is the density of core material, D and d are the whole string’s and its core’s diameters, respectively.

In quasi-static approximation, the relationship between the string’s tension and the magnitude of the plucking effort is a straightforward geometrical relation [13]:

$$F_{ni} = h_i T_i \frac{L_i}{y_{0,i}(L_i - y_{0,i})}$$

where h_i is the displacement of the string relatively to its rest position, $y_{0,i}$ the plucking position and L_i the string’s length.

ACOUSTIC RADIATION OF A SPLIT CARBON COMPOSITE SOUNDBOARD GUITAR COMPARED TO WOODEN CLASSICAL GUITARS

Sylvie LE MOYNE

François OLLIVIER

Joël FRELAT

Charles BESNAINOU

UPMC, Univ. Paris 6, UMR 7190, Institut Jean Le Rond d'Alembert, Paris, France.

CNRS UMR 7190, Institut Jean Le Rond d'Alembert, Paris, France.

charles.besnainou@upmc.fr, sylvie.le_moyne@upmc.fr, francois.ollivier@upmc.fr, joel.frelat@upmc.fr

ABSTRACT

A new design carbon composite soundboard guitar (2012) is compared to traditional wooden guitars: a Antonio Torrès (1887), its copy by Thomas Norwood (2013) and a Ignatio Fléta (1981). Comparison leads to the absolute power radiation and the spectrum shapes. Globally, the split guitar is the most powerful and its spectrum is better balanced.

INTRODUCTION

The soundboard of a string instrument is an essential element of its sound radiation. Many studies ([1]-[5]) have already led to a better understanding of the radiating behavior of guitar soundboards. Regarding the results of these studies, a new design of soundboard has been proposed [6, 7]. The new soundboard is made of a composite material (*i.e.* a sandwich made of a foam core between two unidirectional carbon fibers layers reproducing mechanical properties of best wood [8]). The major innovation (subject of a patent [6]), is a split table that allows on one hand to enrich the modal density of the soundboard and on the other hand to systematize the asymmetry of even modes. The overall objective is to increase the acoustic radiation.

The objective of our work is to estimate the effects of the novel design concepts on the soundboard behavior. The final objective is to investigate different composite guitar design options, and to compare these new concepts to traditional wooden guitars. The present paper proposes a first experimental step of this process. A numerical study (FEM and BEM) is also led simultaneously ([8]).

A first split guitar was made. The bracing of this guitar is asymmetric (unlike those of Torres and Fleta). The two half soundboards of this first prototype are symmetric in shape, thickness and static deformation (which is important). Some traditional wooden guitars were selected for comparison. The objective of this first step is to present a measurement process able to provide relevant comparison criteria of the vibro-acoustic and acoustic radiation of different guitar soundboards.

The technique uses a 120 microphones array to measure the acoustic impulse response of the soundboard in different situations. The radiated power and the radiated pressure spectrum are evaluated. The modal shapes

of the soundboard are estimated using the NAH method adapted by the authors to impulse response field.

EXPERIMENTAL METHOD

A non-intrusive method was used: the Impact Planar Nearfield Acoustical Holography (INAH). This technique, implements the well-known inverse method NAH on the basis of the acoustic impulse response field. It has been implemented previously for the study of a 17th century harpsichord [10], of a pianoforte [11] and of a monochord model [12].

This technique is used to achieve a structural modal analysis. The NAH process of planar harmonic pressure fields is exhaustively described in [13], its adaptation for impulse source excitation (INAH) was initially presented in [14].

The impulse response of the vibrating source is measured in terms of radiating acoustic field with a microphone array. The impulse response is obtained by a point shock excitation of the structure. The vibration behavior of the source is deduced in terms of normal flexural vibration velocity, with the help of an inverse calculation method based on spatial 2D Fourier transforms.

EXPERIMENTAL SETUP

The acoustic impulse responses of the guitars soundboard are all measured in similar conditions.

The instrument is installed in a baffle in order to minimize the radiated sound field from the vibrating parts of the instrument not concerned by the study (Figure 1).

The strings are chosen not to be removed, in order to measure the acoustic radiation of the soundboard in its actual stress state [12, 15-16]. In order to avoid their sound production, the strings are muffled.

A point impulse excitation is provided by an automated hammer driven by an electromagnet that produces a reproducible shock.

A 12 by 10 electret microphones array, with a 50 mm step, is used to collect the pressure field (Figure 1). So as to fit the measurement grid, the array is moved according to 4 or 16 positions. The 120 impulse pressure responses for each position of the array are collected using a home-made 128 channels synchronous digital recorder. Each measurement associated to one shock on the soundboard

has to be phase referenced. Therefore the constant impulse force is systematically recorded along with the acoustic signals.

Two series of measurements are conducted, in two different objectives:

- The first objective is to compare the sound radiation of the different guitars
- The second objective is to measure the modal behavior of the soundboards.



Figure 1. Experimental setup. The guitar is located in a baffling box, the impact hammer excites the soundboard on the bridge, the strings are muffled. The microphone array is located on a parallel plane to the soundboard.

Acoustic radiation

For the acoustic radiation measurement, the excitation points are chosen on the bridge (Figure 2) as it is usually done for modal analysis of soundboards [17-18]. Two excitation locations are chosen: between strings 1 and 2 (S12); between strings 5 and 6 (S56).

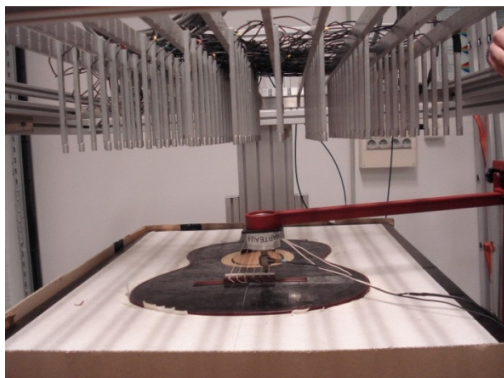


Figure 2. Experimental setup for acoustic radiation measurement

The microphone array records the acoustic impulse response field over a plane parallel to the soundboard at a distance $z_h = 213$ mm. The array is moved according to 4 overlapping positions covering a 900×1100 mm rectangle. The field is finally sampled with a 50 mm step. The final set of measurements counts 437 point acoustic impulse responses.

Vibration field identification

So as to identify the modal behavior of the soundboards, the array is located at a close distance. The excitation impact is applied on the edge of the back plate.

The resulting acoustic impulse response field is measured over a parallel plane at a distance $z_h = 25$ mm. The array is moved according to 16 positions. The field is finally sampled according to a thin grid with a 12.5 mm step and limited to a 500×600 mm rectangle. The final set of measurements counts 1920 point acoustic impulse responses. It undergoes the NAH process at each frequency of interest.



Figure 3. Experimental setup for holographic modal analysis.

RESULTS

Acoustic radiation

On figure 4, the overall acoustic radiated field is presented for the four guitars under measurement.

On figure 5 one can observe the spectra of the space averaged acoustic pressure for the four guitars tested.

In average the split composite guitar has a more balanced spectrum between 700 Hz and 2000 Hz.

It can be then interesting to compare the radiated power of the four guitar in different frequency bands:

- the low frequency band [50-700]Hz corresponding to the musical range of fundamental frequencies of the guitar.
- The high frequency band [700-2000]Hz corresponding of the harmonics range of the instrument.

The radiated power of the four guitars is not so different in the low frequency band, but in the high frequency band, the split guitar is more powerful than the classical guitars.

Then, because the response of ears is more sensitive in the range of [700-2000]Hz, perception of intensity is stronger compare to wooden guitars: and thus giving an enhancement of almost 3 dB [7].

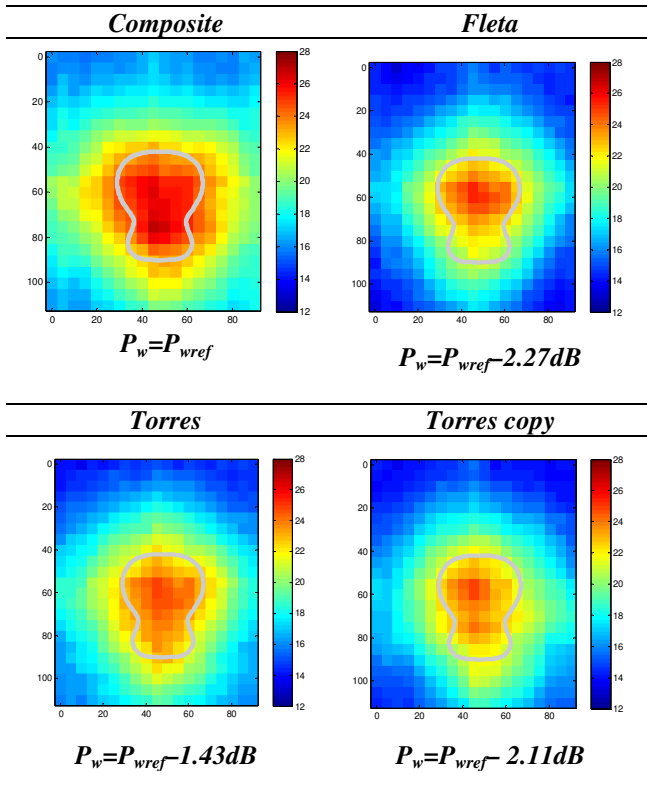


Figure 4. Relative RMS map and radiated power in the [50 2000]Hz band, of the radiated acoustic field of the four guitars. P_{wref} represents the acoustic power radiated by the split guitar. P_w is the radiated power.

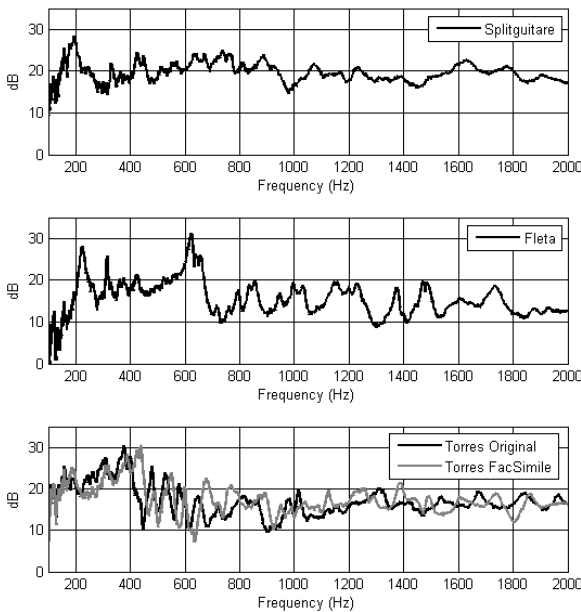


Figure 5. Space averaged acoustic pressure radiated in the nearfield (25mm) for the tested guitars

	Split guitare	Fleta	Torres	Torres Copy
50-2000 Hz	P_{wref}	P_{wref} -2,27dB	P_{wref} -1,43dB	P_{wref} -2,11dB
50-700 Hz	P_{wref}	P_{wref}	P_{wref} +0,79dB	P_{wref} -1,92dB
700-2000Hz	P_{wref}	P_{wref} -4,77dB	P_{wref} -3,69dB	P_{wref} -2,23dB

Table 1. radiated power comparison for different frequency bands. The acoustic power radiated by the split guitar (P_{wref}) is taken as reference.

Operational mode shapes

Operational deflection shapes are reconstructed by means of NAH process using sparse regularization [19] for the four guitars in the [50 – 800] Hz frequency band.

The first modes of the four guitars are presented in table 2-4. Asymmetry and density of split guitar first mode shapes does not seem as obvious as one could expect ([7]). It is probably due to the fact that the prototype tested is split but not so asymmetric: the thickness and the initial soundboard deformation are identical for both half soundboards. The bracing is the only asymmetric conception part of this guitar. The numerical part of the project (in progress) should give us in the future some responses and optimization indications for forthcoming prototype.

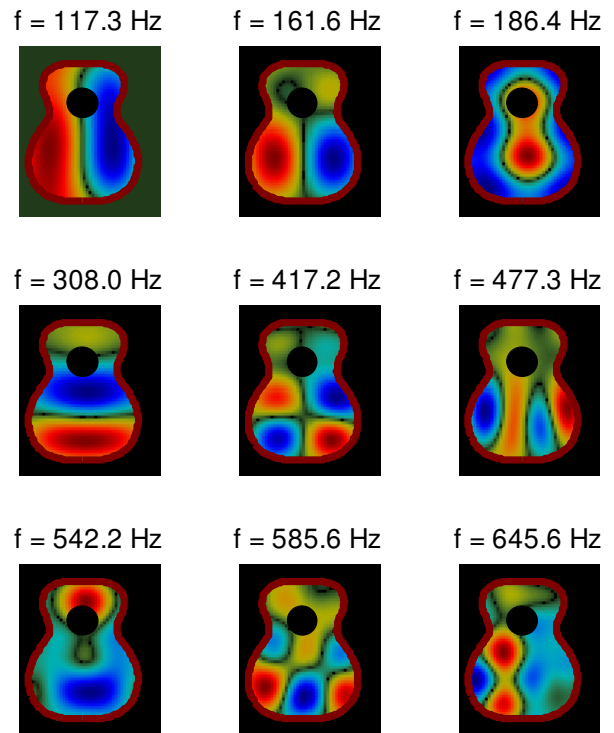


Table 2. Operational mode shapes of the Torres guitar

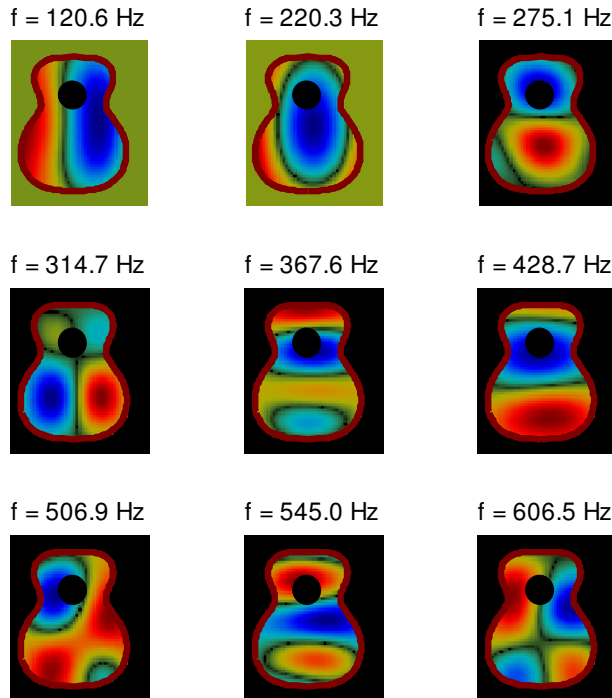


Table 3. Operational mode shapes of the Fleta guitar

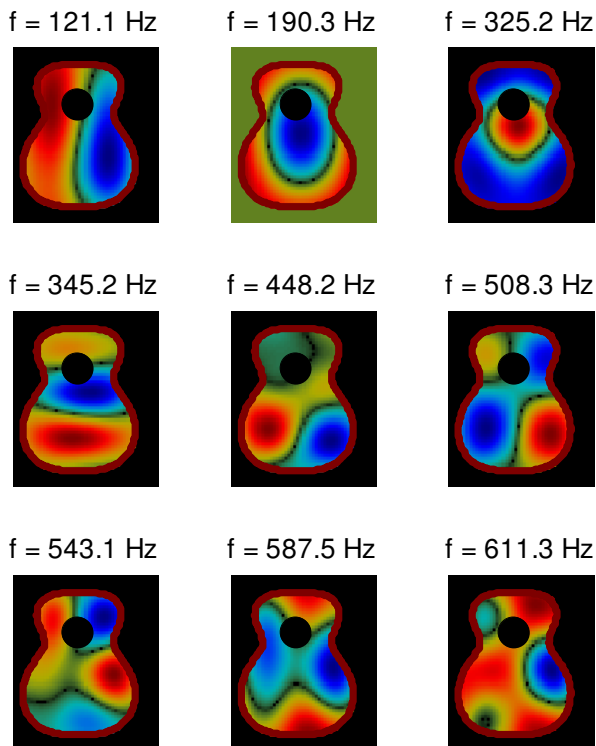


Table 4. Operational mode shapes of the split guitar

CONCLUSION

We have used a device that can give access to both the power radiated by musical instruments and the shapes and frequencies of its eigenmodes. This allows the evaluation of the radiation efficiency in the near and far field.

It will permit to connect objective physical measurements with perceptual judgments of musicians. It also opens a field to optimizing the sound quality of musical instrument. In a further step we will extend measurements at higher frequencies (5000 Hz) and complete the study with numerical models (finite elements and boundary elements, currently in progress). Some other new design composite guitars will also be measured and modeled in order to evaluate the effect of various design concepts.

Acknowledgments

The authors want to thank Dominique Busquet (Université Pierre et Marie Curie, France) for her contribution to the experiments. We also acknowledge Denis Mairat and Hervé César who allowed us to make measurements on so exceptional guitars. The authors want to thank the UPMC students who participated to experimental campaign, Emile Blanc and Marina Vannier.

REFERENCES

- [1] E. Bécache, A. Chaigne, G. Derveaux, P. Joly, "Numerical simulation of a guitar", *Comput. Struct.* 83, pp107–126, 2005.
- [2] S. Griffin, H. Luo, S. Hanagud, "Acoustic guitar function model including symmetric and asymmetric plate modes", *Acta Acust. Acust.* 84, 144–152, 1998.
- [3] I. Curtu, M. D. Stanciu, N. C. Cretu, C. I. Rosca, "Modal Analysis of Different Types of Classical Guitar Bodies", *Proc. of the 10th WSEAS International Conference on Acoustics Music. Theory Appl.*, pp. 30-35, 2009.
- [4] J. A. Torres, R. R. Boullosa, "Radiation efficiency of a guitar top plate linked with edge or corner modes and intercell cancellation", *J. Acoust. Soc. Am.* 130 (1), pp. 546–556, 2011.
- [5] E. Gorrostieta-Hurtado, J-C. Pedraza-Ortega, J-M. Ramos-Arreguin, A. Sotomayor-Olmedo, J. Perez-Meneses, "Vibration analysis in the design and construction of an acoustic guitar", *International Journal of Physical Sciences* Vol. 7(13), pp. 1986 - 1997, 2012.
- [6] UPMC, C. Besnainou, J. Frelat, A. Mamou-Mani, *Instrument de musique à cordes et table d'harmonie*, patent N° 9F-1949CAS19, 18/01/2010, 2010.
- [7] C. Besnainou, J. Frelat, Kuriijn Buys, "A new concept for string-instrument soundboard: the splitting board", in *Proc. ISMA 2010*, Sydney, 2010.
- [8] C. Besnainou, "From Wood Mechanical Measurements to Composite Materials for Musical Instruments : New Technology for Instruments Makers", *Materials Research Society – MRS Bulletin*, march 1995.
- [9] S. Le Moyne, J. Frelat, C. Besnainou, "Un nouveau concept de table d'harmonie de guitare. Etude numérique du comportement vibratoire", in *Proc. CSMA2013*, Gien, France, 2013.

- [10] S. Le Moyne, S. Le Conte, F. Ollivier, J. Frelat, J-C. Battault, S. Vaiedelich, "Restoration of a 17th century harpsichord to playable condition: a numerical and experimental study", *J. Acoust. Soc. Am.* 131(1) pt.2, pp. 888-896, 2012.
- [11] F. Ollivier, S. Le Moyne, S. Le Conte, "Acoustics radiation and modal analysis of a piano forte and its fac-simile", in *Proc. Acoustics2012*, Nantes, 2012.
- [12] A. Mamou-Mani, S. Le Moyne, F. Ollivier, C. Besnainou, J. Frelat, "Prestresses effects on soundboard vibration " *J. Acoust. Soc. Am.* 131(1) pt.2, 872-877, 2012.
- [13] E.G. Williams, *Fourier Acoustics. Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, 1999.
- [14] F. Ollivier, P. Alais, A. Karkaletsis "Fast modal analysis by means of impulse acoustical holography", in *Proc. 11th ICSV*, 2004, pp 2395-2402.
- [15] A. Mamou-Mani, J. Frelat, and C. Besnainou, "Numerical simulation of a piano soundboard under downbearing" *J. Acoust. Soc. Am.* 123, pp.2401–2406, 2008.
- [16] A. Mamou-Mani, J. Frelat, and C. Besnainou, "Pre-stressed soundboards: analytical approach using simple system including geometric nonlinearity", *ActaAcust. Unit. Acust.* 95, pp.915–926, 2009.
- [17] N. Giordano, "Sound production by a vibrating piano soundboard: Experiment", *J. Acoust. Soc. Am.* 104(3), pp.1648-1653, 2001.
- [18] H. Suzuki, "Vibration and sound radiation of a piano soundboard", *J. Acoust. Soc. Am.* 80(6), pp. 1573-1582, 1986.
- [19] G. Chardon, L. Daudet, A. Peillot, A. F. Ollivier; N. Bertin, R. Gribonval, "Near-field acoustic holography using sparse regularization and compressive sampling principles", *J. Acoust. Soc. Am.*, 132(3), pp.1521-1534, 2012.

VIBROACOUSTIC BEHAVIOR OF A VIHUELA

Sylvie LE MOYNE

UPMC, Univ Paris 6.
Institut Jean Le Rond d'Alembert,
CNRS UMR 7190, Paris, France
sylvie.le_moyne@upmc.fr

François OLLIVIER

UPMC, Univ Paris 6.
Institut Jean Le Rond d'Alembert,
CNRS UMR 7190, Paris, France
francois.ollivier@upmc.fr

Sandie LE CONTE

Laboratoire de Recherche et
Restauration du Musée de la
musique, Paris, France
sleconte@cite-musique.fr

ABSTRACT

The music museum of Paris possesses one of the three existing instruments certified to be vihuelas. None of these instruments is playable, and therefore no one can tell with certainty what the vihuela sounded like. The objective of the museum is to improve our knowledge of this unknown instrument and allow the general public to hear it. In this objective a *facsimile* was ordered and realized. An Impact Near Field acoustic holography (INAH) measurement of the *facsimile* top plate was performed. The sound radiated by the back of the *facsimile* instrument was measured with the same technique. For both the top plate and the back, the spectrum of the radiated sound is investigated, and a modal analysis is proposed. The influence of the string tension on the top plate acoustic radiation is also studied.

Measurement on the copy, as it is complete and playable have the principal objective of improving our knowledge of this instrument.

INTRODUCTION

The "Musée de la musique" in Paris preserves more than 7000 objects related to music and among them more than 5000 are musical instruments.

The mission of the "Musée de la Musique" is to conserve, display and make available its collection, especially musical instruments. Conservation primarily involves the management of collections inventories, the preventative conservation of instruments, the acquisitions policy, and research and documentation on cultural heritage objects and their restoration.

Preserving cultural heritage objects consist in evaluating all the cultural values borne by the object. These cultural values are the historical one, the aesthetic one, the rarity of the object, and for musical instruments one could add the acoustical value. As a result of this evaluation, among the 5000 musical instruments less than 5% are considered in a playable state. This means that for less than 5 % the acoustical value is the most important. Indeed maintaining a musical instrument in playable conditions means doing some adjustment or restoration interventions, not reversible, which could not agree with the deontological principles of the conservation. Nevertheless to have a better objective evaluation of the cultural values, the "Musée de la Musique" has set up a laboratory working in a range of areas and developing activities of increasing scientific importance. Many projects are carried out in collaboration with different institutions, like University

Pierre et Marie Curie for vibroacoustic researches on several instruments.

It is of particular importance for the museum to enable the public to hear the sound of the instruments presented. When the cultural values evaluation leads to favor the acoustical function of the cultural heritage object it can be chosen to restore the instrument in a playable state (according the reversible principles) [1] or to build a *facsimile* [2]. The making process of a *facsimile* is a pluri-disciplinary project since it involves a maker of course but also historians, organologists, scientists and most of the time a musician. It consists in reproducing an historical instrument according to ancient know-hows, techniques and using the same materials because it is possible to copy only what it is known.

Dated back to the 16th century, the vihuela is a 12 strings instrument widely considered to be the Iberian counterpart to the lute. As of today, only three certified instruments exist:

- The first one, in the Iglesia de la Compañía de Jesús de Quito, in Quito, Ecuador, is treated as a relic,
- The vihuela Guadalupe is exposed in the Jacquemart-Andrée museum in Paris, and is considered to be a true lute-making masterpiece. It is a controversial instrument due to its stringing and to its shape [3,4],
- The vihuela Chambure, rediscovered only in 1998, is exposed at the Cité de la Musique in Paris. This original instrument is supposed to be from the 16th century. It is the object of our study.

None of the three specimens is playable, so that no one can tell with certainty what the vihuela sounded like. The restoration of the Vihuela Chambure is not possible because the soundboard structure is too weak (many cracks, and some parts are missing). Restoring this instrument would lead to the loss of historical information because too many interventions are needed. The objective of the museum is to improve the knowledge of this unknown instrument and to allow the general public to hear it. In this objective a *facsimile* was ordered. It was realized by Sandi Harris & Stephen Barber (London) in 2010. The original instrument and its copy are presented on Figure 1 and Figure 2.

This paper presents a preliminary study of the vibroacoustic properties of the *facsimile* using the impact near-field acoustic holography technique (INAH). The INAH method and the experimental setup are presented in the first section. The radiated sound of both the top plate and the back were measured. In the case of the soundboard response, it was also decided to examine the influence of

the string tension. A first analysis of results is proposed in terms of spectral response and modal analysis.



Figure 1. Original vihuela Chambure.



Figure 2. Making of the *facsimile* of the vihuela Chambure.

MEASUREMENT TECHNIQUE AND EXPERIMENTAL SETUP

The Impact Nearfield Acoustical Holography (IN-AH) used in this study is a fast technique allowing collecting the vibration characteristics of a plane structure. This technique implements the well known NAH inverse method on the basis of the acoustic impulse response field. It has been used previously for the monitoring of a 17th century harpsichord [1] and on a monochord model [5].

This technique is used here in the aim of achieving a structural modal analysis. The NAH process of planar harmonic pressure fields is exhaustively described in [6], its adaptation for impulse source excitation (IPNAH) was primary presented in [7].

The plane structure is excited using an automated hammer applying a controlled and repeatable point force. A plane regular microphone array located in a close parallel plane is used to record the spatial acoustic impulse response (AIR) of the structure, including the evanescent waves signature. Iterating the process for several positions of the array allows sampling very sharply the acoustic near field. NAH can be implemented afterwards at each frequency of interest. It consists in solving an inverse propagation problem to reconstruct the bending vibration field of the structure.

The impulse response of the vihuela is measured in the semi-anechoic room of the museum. Ambient tempera-

ture and hygrometry were controlled during all along the experiment.

It was chosen not to remove the strings in order to measure the acoustic radiation of the top plate in its actual stress state [5, 8-9]. In order to avoid the strings to sound and perturb the measurement, these were muffled (Figure 3).



Figure 3. Experimental setup for top plate impulse response measurement. The instrument is installed in its baffle, the strings are muffled. The excitation hammer and the reference accelerometer are in place.

The instrument is installed in a baffle in order to minimize the radiated sound field from the vibrating surface not concerned by the study (Figure 3).



Figure 4. Experimental setup for the back impulse response measurement. The instrument is installed in its baffle. The excitation hammer and the reference accelerometer are in place.

A point impulse excitation is provided by an automated hammer driven by an electromagnet that produces a reproducible shock. The excitation point is chosen on the bridge (Figure 3) as it is usually done for modal analysis of top plate [10-11]. In the case of the measurement of the back radiation, an excitation point on the back was chosen (Figure 4). A 12 by 10 electret microphones array, with a 50 mm step, has been used to collect the pressure field (Figure 5). So as to fit the measurement grid, the array is moved according to 16 positions. The 120 impulse pressure responses for each position of the array are collected using a custom-made 128 channels synchronous digital recorder. Each measurement associated to one

shock on the top plate has to be phase referenced. Therefore an accelerometer has been positioned on the soundboard near the excitation point (Figures 3 and 4), and its constant impulse response is systematically recorded along with the acoustic signals.

The resulting acoustic impulse response field is measured over a parallel plane at a distance $z_h = 103$ mm. The field is finally sampled according to a thin grid with a 12.5 mm step and limited to a 500 x 600 mm rectangle. The final set of measurements counts 1920 point acoustic impulse responses. After time frequency analysis, this so called hologram can undergo the NAH process at each frequency of interest in order to identify the bending vibration field of the structure.



Figure 5. Experimental setup for the top plate impulse response measurement. The microphone array is located at a distance $z_h = 103$ mm from the soundboard.

RESULTS

Acoustic radiation

Strings influence

It can be expected that the string tension modifies the stress state of the instrument. In order to evaluate this effect on the vihuela, the impulse response of the top plate was measured for three different conditions:

- With strings at full tension
- In the few minutes following the strings removal
- 2h30 after the strings were removed

Figure 6 presents the space averaged acoustic spectra of the top plate of the *facsimile* for the three conditions.

No difference can be noticed between the two cases without strings. This implies that no relaxation effect affects the acoustic radiation of the top plate when the strings are removed.

The radiated spectrum is weakly affected, resonance frequencies are not significantly modified.

This small influence can be explained by the low height of the bridge. The normal force induced by the strings tension on the bridge is consequently very small.

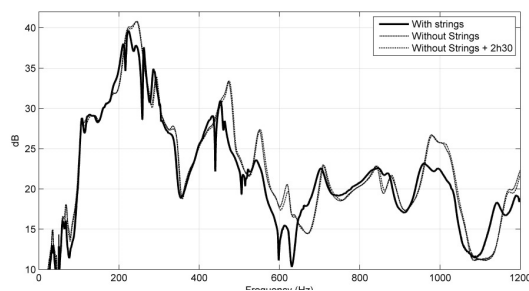


Figure 6. Average spectra on the microphone array. Influence of string tension.

Soundboard and back

Figure 7 presents a comparison of the space averaged acoustic spectra for the top plate and the back. In both cases the strings were tensed on the instrument.

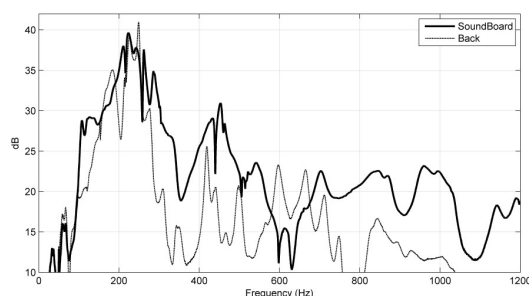


Figure 7. Space averaged acoustic impulse response spectra of the soundboard and back.

These curves show the important coupling between both sides of the instruments. Exploration of modal patterns for both sides shows also modal participation of both sides in the 100-1000Hz frequency band. This non-negligible participation of the back in terms of vibrations is in concordance with musicians verbatim.

Operational mode shapes

Operational mode shapes shown in figures 8 and 9 are the result of the holographic reconstruction of a virtual vibrating source on the top plate plane for six radiating frequencies. Figure 9 shows the vibration field reconstructed on a virtual plane on the back plate position. The reconstruction for the back plate is less precise and only gives some indications. Indeed the real back plate is curved with a arch height of 2.8cm.

These results have some similarities with classical guitars results ([12]-[13]).

Mode shapes and resonances of the top plate and back are strongly coupled at low frequencies (modes (1,1), (1,2)). The sequence of modal shapes of the top plate is also quite similar to the behavior of classical guitar. The modal behavior of the back is more different as it is curved and quite rigid. Resonance frequencies of modes (1,1), (1,2) and (2,1) are significantly lower than those of classical guitars [13], this is not surprising assuming that the bracing in the classical guitar is more complex than in the vihuela (only two struts). On the response spectra, an important resonance area can be observed in the [200-

275] Hz frequency band, where the top plate and the back plate vibrate in piston mode. It is probably the Helmholtz frequency domain.

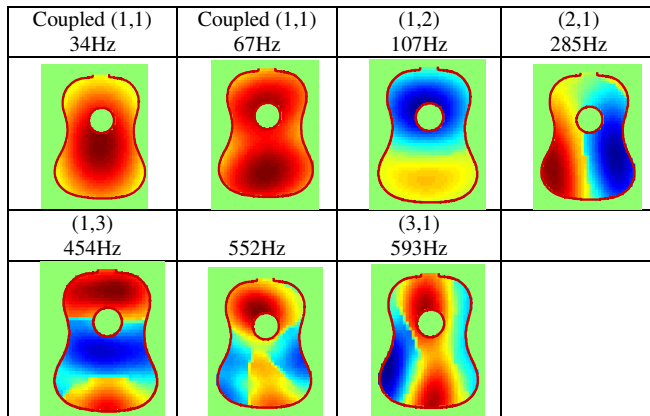


Figure 8.Operational mode shapes of the top plate

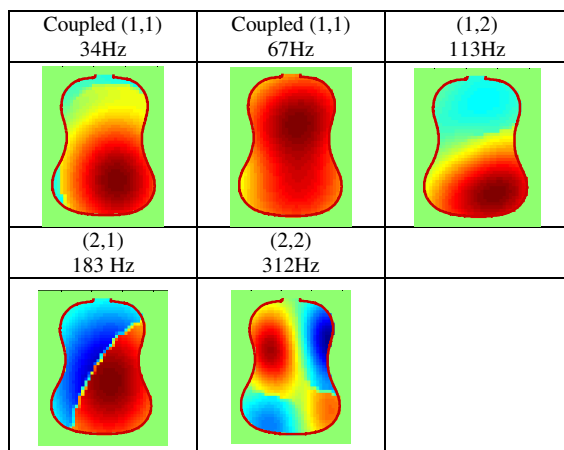


Figure 9.Operational mode shapes of the back plate

CONCLUSION

A first experimental study was conducted on a recent *facsimile* of the vihuela Chambure owned by the music museum of Paris. The impulse acoustic response of the top plate and back plate was measured with a 120 microphone array. Using a near field acoustical holography process, the corresponding vibrating behavior was identified for some resonance frequencies.

Acknowledgments

The authors want to thank Dominique Busquet (Université Pierre et Marie Curie, France) for her contribution to the experiments and the Musée de la musique curator J Dugot, who allowed us this study.

REFERENCES

[1] S. Le Moyne, S. Le Conte, F. Ollivier, J. Frelat, J-C. Battault, S. Vaiedelich, "Restoration of a 17th century harpsichord to playable condition: a numerical and experimental study", *J. Acoust. Soc. Am.* 131(1) pt.2, pp. 888-896, 2012.

[2] F. Ollivier, S. Le Moyne, S. Le Conte, "Acoustics radiation and modal analysis of a piano forte and its fac-simile", in *Proc. Acoustics2012*, Nantes, 2012.

[3] A. Corona-Alcade. "L'organologie de la vihuela". In "Aux origines de la vihuela de mano", *les cahiers du musée*, n°5, pp. 16-28, ed. Cité de la musique.96 p., 2004.

[4] R. Bader."Radiation characteristics of multiple and single sound hole vihuelas and a classical guitar", *J. Acoust. Soc. Am.* 131 (1), pp. 819–828, 2012.

[5] A. Mamou-Mani, S. Le Moyne, F. Ollivier, C. Besnainou, J. Frelat, "Prestresses effects on soundboard vibration " *J. Acoust. Soc. Am.*131(1) pt.2, 872-877, 2012.

[6] E.G. Williams, *Fourier Acoustics. Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, 1999.

[7] F. Ollivier, P. Alais, A. Karkaletsis "Fast modal analysis by means of impulse acoustical holography",in *Proc.11th ICSV*, 2004, pp 2395-2402.

[8] A. Mamou-Mani, J. Frelat, and C. Besnainou, "Numerical simulation of a piano soundboard under downbearing" *J. Acoust. Soc. Am.*123, pp.2401–2406, 2008.

[9] A. Mamou-Mani, J. Frelat, and C. Besnainou, "Prestressed soundboards: analytical approach using simple system including geometric nonlinearity", *Acta Acust. Unit. Acust.* 95, pp.915–926, 2009.

[10] N. Giordano, "Sound production by a vibrating piano soundboard: Experiment",*J. Acoust. Soc. Am.*104(3), pp.1648-1653, 2001.

[11] H. Suzuki, "Vibration and sound radiation of a piano soundboard", *J. Acoust. Soc. Am.*80(6), pp. 1573-1582, 1986.

[12] N. Fletcher, T. Rossing, *The physics of musical instruments*, Springer-Verlag, 1991.

[13] G. Caldersmith, "Designing a guitar family", *Applied Acoustics*, 46, pp.3-17, 1995.

Eigenspectrum of the sarode membrane

Manaswi Mishra

Department of Physics,
IIT Madras

manaswimishra17@gmail.com

Gowtham Suresh

Department of Physics,
IIT Madras

gowthams001@gmail.com

R. Adhikari

Institute of
Mathematical Sciences

rjoy@imsc.res.in

ABSTRACT

The sarode is a multi-stringed musical instrument widely used in performances in the classical music of northern India. Its main parts consist of a solid body of carved wood, a fretless metal fingerboard, and a goatskin membrane radiator. While the instrument is second only in popularity to the sitar in north Indian music, to the best of our knowledge, no mathematical modeling of the physical instrument has been attempted to characterize its sound. Here, as a first step towards a complete mathematical model of the sarode, we compute the eigenvalues and eigenfunctions of the sarode membrane, the main radiating element of the instrument. We obtain the shape of the membrane from laser scans of a real instrument and then project it onto a computational mesh for finite element eigenanalysis. While the eigenfunctions can be characterised by the number of open and closed nodal contours, as is customary for highly symmetric circular or elliptic membranes, they are totally non-degenerate and have no symmetry about the nodal lines. The enhanced number of distinct eigenmodes that appear due to the peculiarity of the membrane shape adds to the timbre of the instrument. We also compute the eigenspectrum including radiation loss in a phenomenological manner to aid possible numerical modal sound synthesis.

1. INTRODUCTION

Almost every culture around the world has developed their musical traditions rooted in local practices, customs and available materials. Central to these musical traditions are a variety of musical instruments. Indian music is no exception to this and overtime a large number of musical instruments have evolved, each with a specific musical form in mind. The complexity of the instruments also varies substantially from simple drums to complicated loaded drums, from single stringed instruments to multistringed instruments which are either plucked or bowed. Surprisingly, for a musical tradition that have evolved such complex array of musical instruments, the scientific analysis and mathematical

modeling of such instruments is still in its infancy though the initial foray was made almost a century ago by C V Raman. This is unlike the musical instruments used in western musical traditions where the scientific analysis of acoustics of instruments encompasses a wide variety. Raman[1, 2] first analysed the Indian drums, the tabla and the mridangam, and the long-necked lutes the tanpura and the veena[3]. Raman noticed that the Indian drums have harmonic overtones due to the central density loading and that the lutes had a powerful series of overtones, absent in other string instruments, due to their curved "jawari" bridge. Though not studied by Raman, this feature is common to the sitar and has been the subject of recent acoustic studies[4, 5].

A popular and widely played musical instruments that has not been acoustically modeled, to the best of our knowledge, is the sarode.¹ Sarode is extensively used in performing the north Indian or hindustani classical music, one of the two major traditions in the highly evolved Indian music system. Sarode seems to have evolved around two centuries ago with its origins in the Afghan rubab. The body of the instrument consists of (see Fig.1) a solid carved teak wood frame covered on top by a fretless metal, usually made of steel, finger board and a goat skin membrane radiator. The total length is around a metre depending on the construction. Strings are tied at the end of the goat skin top through a bridge and held at the other end tied to the pegs. The main playing strings are either four or five with a number of sympathetic strings and drone strings. The excitation of strings is by plucking. Unlike other stringed Indian instruments where the radiator is usually made of wood, sarode has a goat skin membrane which gives the instrument its rich tonal quality.

A complete acoustical analysis must include the dynamics of the excitation of the string by plucking, the transfer of the string energy to the bridge and from there, to the membrane, which is coupled to the enclosed air within the body of the instrument.[6] The multiple sympathetic strings are coupled at the bridge to the main playing strings and can, thus, absorb energy from them, producing a long aftersound similar to the piano. Such an acoustical analysis is a formidable problem involving the coupled vibrations of canonical acoustical elements like strings, membranes, and enclosed air, where the membrane and the enclosed air volume have non-trivial

Copyright: ©2013 Manaswi Mishra et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ For a brief description of the instrument, see <http://en.wikipedia.org/wiki/Sarod>

geometry. Our goal in this paper is to take a first step towards the complete analysis by studying the eigenmodes of sarode membrane, which is the main radiating element of the instrument.

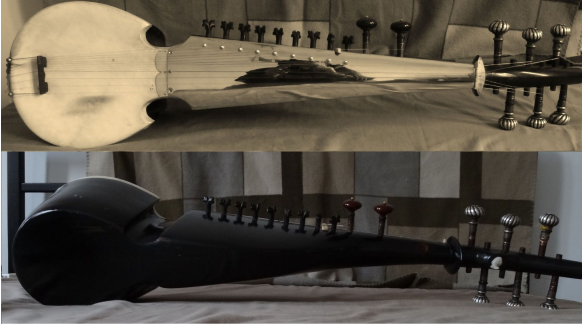


Figure 1. Front and rear views of the late 19th century sarode used in this study

The paper is organised as follows: In Sec.II we discuss the accurate experimental determination of the shape of the sarode membrane through laser scanning and the extraction of relevant parts of the experimental data for mathematical analysis. In Sec.III we describe in detail the mathematical model applied to the two dimensional sarode membrane including damping terms and the appropriate boundary conditions. In Sec.IV, we present numerical results for the eigenspectrum of the membrane and compare them with regular shaped membrane such as circular or elliptic membrane. Finally in Sec.V we present a summary of our observations and conclusions.

2. SARODE MEMBRANE

As our instrument for analysis, we have chosen an approximately 140 year old instrument manufactured by artisans in Darbhanga in present day Bihar. It was the concert instrument of Mohammed Amir Khan and is now played by Arnab Chakrabarty. The instrument has more sustain but is less loud compared to sarodes that have been manufactured more recently. We plan to compare the evolution of membrane shapes and their impact on the sarode sound in a future study.

In order to obtain the sarode membrane shape as an input to our finite element analysis algorithm, a laser scan of the sarode skin was performed. ScanWorks V5 3D laser scanner from Perceptron² was used for this purpose, with a mean point to point resolution of 0.013mm and a measurement accuracy of 0.024mm.

This gives us a point cloud data from which just the boundary must be extracted. The extraction procedure was three fold and is described sequentially as follows.

2.1 Cleaning of Scanned object

The Laser scan of the Sarode membrane produces a three dimensional *STL* (stereo lithography) object. As the laser scanning works on a principle of creating a point cloud

² <http://www.perceptron.com/index.php/en/-industrial/3d-scanning-solutions-g/manual-3d-scanning.html>

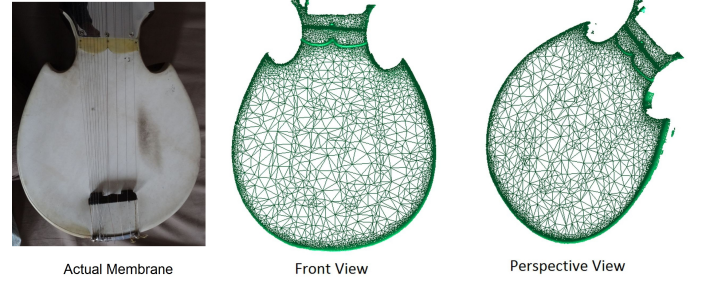


Figure 2. The pre-processed point cloud data of the sarode skin.

data whose aberrations depend on various parameters like maximum scan angle, scan rates, scan density, exposure noise reduction filters etc, the *STL* file exhibits various undesired features like spilling over of meshing at corner points and unconnected components. The software gmsh[7] is used for all the pre-processing steps. We clean the raw scanned object by deleting the above mentioned unconnected components, not part of the main sarode body.

2.2 Stencil Extraction

The connected spill overs at the boundaries of the sarode membrane are avoided and the boundary is isolated by performing a gradient operation on the point cloud data. This gradient operation performs a simple edge extraction. The two dimensional stencil boundary is then obtained from this edge by projecting the extracted edge onto its perpendicular plane.

2.3 Boundary Meshing

After our boundary stencil has been extracted it must be meshed for performing finite element analysis. The density of meshing is one of the parameters and convergence of solutions can be obtained from further refining this mesh.

This meshed boundary is passed as an input data set to the finite element algorithm module which applies the Dirichlet boundary conditions and solves for the wave equation thereby generating its eigenvalues and eigenmodes as described in detail in the following section.

3. MODEL AND METHODS

In order to solve the eigenspectrum of the sarode membrane, we use the equation of motion governing the damped membrane with uniformly loaded tension T given by

$$\rho \ddot{u} = T \nabla^2 u + \eta_1 \dot{u} + \eta_2 \nabla^2 \dot{u} \quad (1)$$

Here ρ denotes the uniform density of the membrane, T is the uniform unloaded tension of the membrane and $u = u(\mathbf{r}, t) = u(r, \theta, t)$ is the transverse displacement of a point on the membrane at time t . The eigenspectrum is

obtained by solving the above differential equation satisfying the Dirichlet boundary condition:

$$u(R_i, \theta_i, t) = 0,$$

where R_i, θ_i denote the points on the boundary.

The eigenvalue equation is obtained by assuming a solution of the form $u(r, \theta, t) = \Psi_{mn}(r, \theta)e^{\omega t}$. Substituting this in Eq.1, we have the eigenvalue equation for Ψ_{mn} given by

$$[\omega_{mn}^2 \rho - \omega_{mn}(\eta_1 + \eta_2 \nabla^2) + T \nabla^2] \Psi_{mn} = 0 \quad (2)$$

Here Ψ_{mn} refers to the eigenmode of the eigenvalue equation. The eigenvalues and eigenfunctions in general depend on the damping terms. The coefficients η_1 and η_2 capture the effect of damping, while η_1 represents an overall frequency independent damping of all modes, η_2 can be thought of as a frequency dependent damping factor.

We first set up the eigenvalue problem using the separation of variables as described below. For this purpose, it is most convenient to use dimensionless variables and rewrite the Eq.1 in terms of these variables. We introduce a length scale a such that

$$u(\mathbf{r}, t) = aw(\mathbf{r}, t); \quad \mathbf{r} \rightarrow a\mathbf{r}$$

where w is now dimensionless and all length scales are measured in units of a . Furthermore the time t is rescaled by the factor

$$t \rightarrow \sqrt{\frac{T}{\rho a^2}} t$$

so that the equation of motion for w is given by

$$\ddot{w} = [\nabla^2 w + \gamma_1 \dot{w} + \gamma_2 \nabla^2 \dot{w}], \quad (3)$$

where

$$\gamma_1 = \frac{\eta_1 a}{\sqrt{\rho T}}; \quad \gamma_2 = \frac{\eta_2}{a \sqrt{\rho T}}.$$

The differential equation 3 may now be solved by separation of variables. Let

$$w(\mathbf{r}, t) = \psi(\mathbf{r})\phi(t). \quad (4)$$

Inserting this into the dimensionless equation of motion gives

$$(\ddot{\phi} - \gamma_1 \dot{\phi})\psi(\mathbf{r}) = (\phi + \gamma_2 \dot{\phi})\nabla^2 \psi(\mathbf{r}) \quad (5)$$

The condition under which the separation of variables is valid is then given by

$$\frac{(\ddot{\phi} - \gamma_1 \dot{\phi})}{(\phi + \gamma_2 \dot{\phi})} = \frac{\nabla^2 \psi(\mathbf{r})}{\psi(\mathbf{r})} = \lambda. \quad (6)$$

This gives the eigenvalue equation for the spatial eigenmodes

$$\nabla^2 \psi(\mathbf{r}) = \lambda \psi(\mathbf{r}) \quad (7)$$

and the ordinary differential equation for the temporal evolution

$$\ddot{\phi} - (\gamma_1 + \lambda \gamma_2) \dot{\phi} - \lambda \phi = 0. \quad (8)$$

The spectrum of the Laplacian with Dirichlet boundary conditions in a bounded domain is real and negative, corresponding to localised states with finite energy. Thus, in what follows, we take $\lambda < 0$. The temporal evolution is solved by assuming a solution of exponential form

$$\phi(t) = \phi(0)e^{\mu t}. \quad (9)$$

Substituting this in Eq.8 we get the quadratic equation for μ

$$\mu^2 - (\gamma_1 + \lambda \gamma_2) \mu - \lambda = 0 \quad (10)$$

with a pair of complex conjugate solutions

$$\mu = \frac{1}{2}(\gamma_1 + \lambda \gamma_2) \pm \frac{1}{2}\sqrt{(\gamma_1 + \lambda \gamma_2)^2 + 4\lambda} \quad (11)$$

for $(\gamma_1 + \lambda \gamma_2)^2 + 4\lambda < 0$, which is ensured for small damping. The real parts of the solutions

$$Re(\mu) = \frac{1}{2}(\gamma_1 + \lambda \gamma_2) \quad (12)$$

are the dampings, while the imaginary parts

$$Im(\mu) = \pm \frac{1}{2}\sqrt{(\gamma_1 + \lambda \gamma_2)^2 + 4\lambda} \quad (13)$$

So we can get the damping part, both wave number dependent and wave number independent parts analytically from the eigenvalues of the spatial laplacian, obtained by solving Eq.8.

Since the shape of the Sarode membrane does not have rectangular symmetry we have to solve the equation piece wise through elemental meshing of the shape. This finite element analysis is done using the GetFem++ libraries [8]. In our case since the membrane boundary is fixed in temporal confinement we apply the dirichlet boundary conditions for the eigenvalue equation.

This requires us to add a constraint on all the corresponding boundary points. Our eigenvalue problem is modeled as

$$Au = \lambda Mu$$

A represents the stiffness matrix and M represents the mass matrix. Since we have an absorbing boundary, the following changes have to be made in the mass and stiffness matrix to capture the same -

$$|A_{i,j}| = \begin{cases} \approx LARGE & \text{if } i = j \text{ and } (i, j) \in \text{boundary} \\ A_{i,j} & \text{otherwise} \end{cases}$$

$$|M_{i,j}| = \begin{cases} 1 & \text{if } i = j \text{ and } (i, j) \in \text{boundary} \\ M_{i,j} & \text{otherwise} \end{cases}$$

The only input required to our model is the meshed boundary that has been scanned and pre-processed, and the results obtained are discussed in the following section.

4. RESULTS

The eigenmodes and eigenvalues were calculated from the finite element analysis of the sarode membrane and the disparities with circular and elliptical membrane were observed.

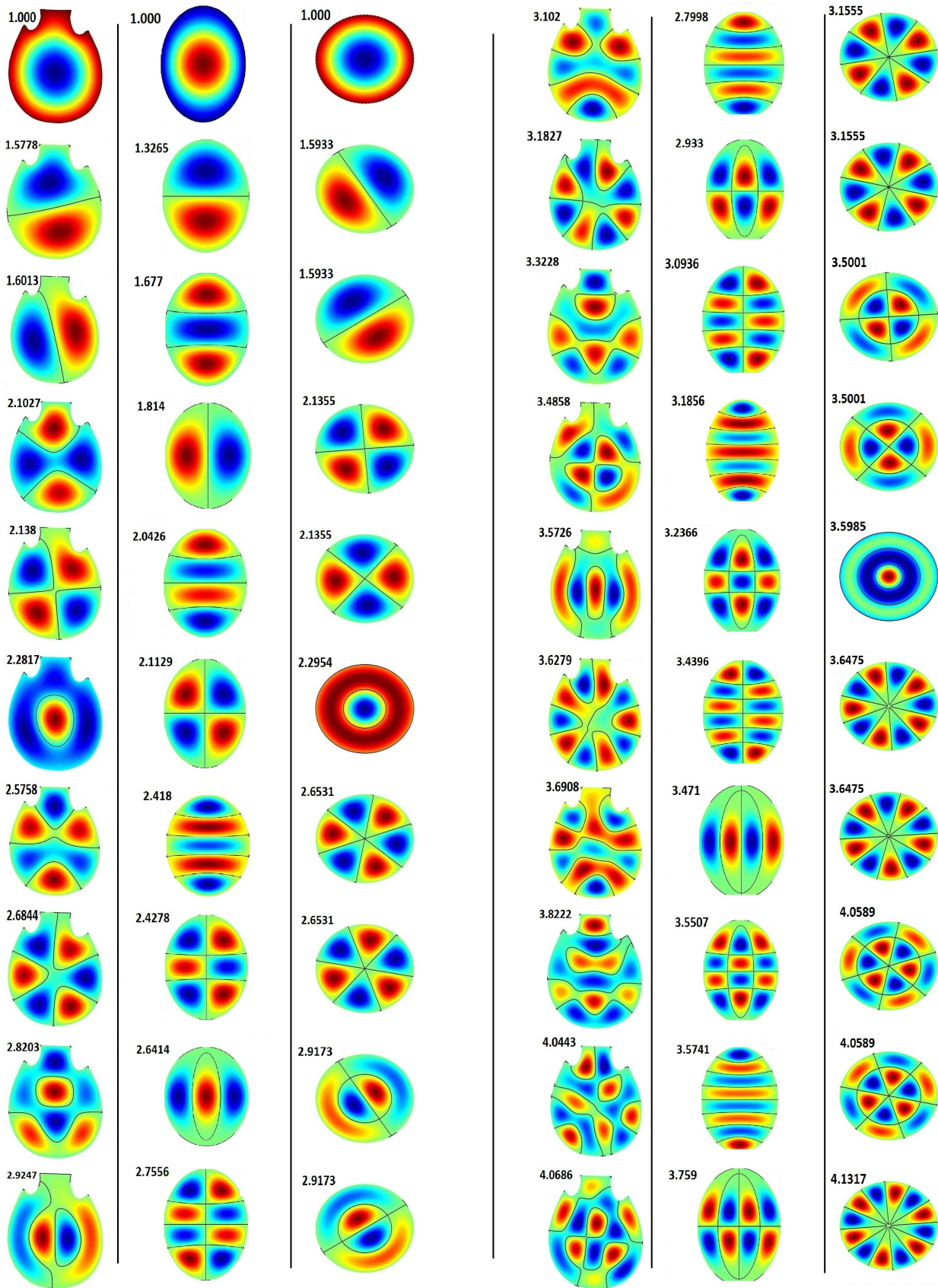


Figure 3. The first 20 eigenmodes of the sarode skin with their eigenvalue ratios compared with the ellipse and the circle eigenmodes.

4.1 Predicted Eigenmodes

The eigenvalues of the sarode were computed and compared with the circular membrane. One of the main observations was the non-degenerate eigenvalues of the sarode as against the degenerate eigenvalues of the circle. In other words, a circle has multiple eigenmodes contributing the same energy unlike the sarode eigenmodes. This was expected due to the lack of radial symmetry as in the circular membrane. We also observe the order in which the eigenmodes occur indicating the energy contributions from each axis of symmetry.

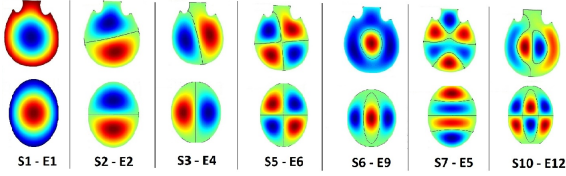


Figure 4. Comparison between similar elliptical and sarode eigenmodes based on number of modal diameters and circles

The fig. 4 indicates a comparison between the eigenmodes of elliptical and sarode shaped membrane based on their corresponding number of modal diameters and modal circles. The clear variation from circular symmetry is noticed in the heat maps and the lack of degeneracy in the eigenratios. Since the sarode membrane is a uniform unloaded membrane its only variation is its unsymmetrical shape. Hence we compare this disparity by solving for the same equation for an area normalized circular membrane. The following plot shows the eigenratios obtained for both the circular membrane and the sarode membrane.

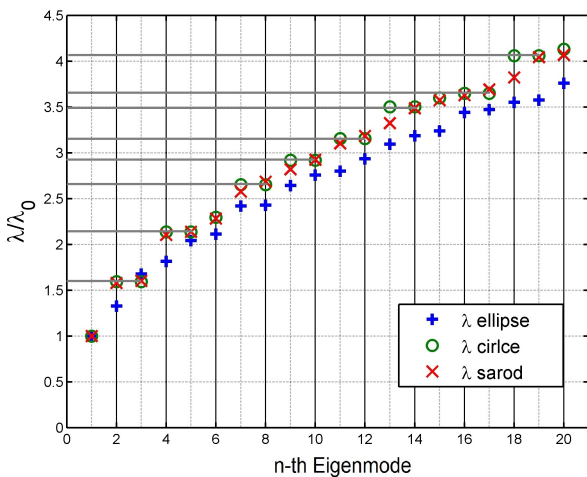


Figure 5. Comparison of first 20 eigenvalues of sarode, circle and elliptical shaped membrane boundaries

The solid horizontal lines in fig.5 indicate the consecutive degenerate eigenvalues of the circular

membrane. In comparison to this, it is observed that the sarode eigenvalues do not exhibit this degeneracy.

4.2 Damped membrane

We also depict a plot of the damped eigenvalues indicating the real part and the imaginary part separately on a phase plane. These are the eigenvalue solutions for non zero damping in 10. The plot indicates that varying the wavenumber independent damping coefficient, γ_1 , shifts all frequencies by an equal amount while the wavenumber dependent coefficient, γ_2 , shifts the frequencies by an amount proportional to the undamped values.

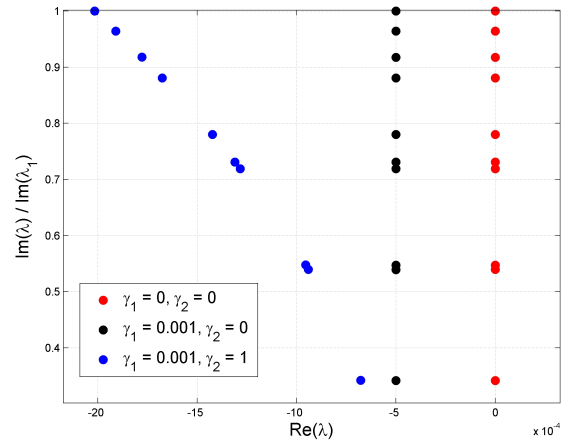


Figure 6. Phase plot indicating frequencies with different values of γ_1 and γ_2

5. CONCLUSION

We have presented a model for the vibrations of the sarode membrane and have performed an experiment predicting the eigenmodes and eigenvalues of the sarode membrane shape. A comparison with a similar analysis on the regular circular and elliptical shaped membranes gives us an insight into the richness of the sarode tone added by its unique shape.

There are several ways in which the present work needs to be elaborated for a more realistic physical modeling of the vast class of unique Indian musical instruments. First, we have neglected the effect of the bridge that is fixed onto the membrane while playing that would contribute to a source term in the wave equation leading to non linearity as shown by S Siddiq [5]. Also the other contributors to the sarode tone, the air column, unique bridge and wooden body have not been considered. The unique effects of the bridge and the smooth membrane folding edges (like in the tabla, Indian drum) that add a multitude of otherwise non existent overtones also needs to be studied deeper.

Second, with further refinement, our numerical method should find application in numerical sound synthesis.

Acknowledgments

The authors thank Arnab Chakrabarty for generously allowing experiments on the rare sarode used in the study. The work developed from initial suggestions by Arnab Chakrabarty to study the sarode body eigenmodes by combining laser scanning data and finite element analysis. We thank M. V. N. Murthy for many useful discussions and a critical reading of the manuscript. We also thank Professor R. Balasubramanian, Director, The Institute of Mathematical Sciences and the CompMusic project ERC grant agreement 267583 for grants which facilitated this study.

6. REFERENCES

- [1] C. V. Raman and S. Kumar, “Musical drums with harmonic overtones,” in *Nature(London)*, pp. **104**, 500, 1920.
- [2] G. Sathej and R. Adhikari, “Eigenspectra of indian musical drums,” in *Journ. Acoust. Soc. Am*, pp. **125** 831–838, 2009.
- [3] C. V. Raman, “On some indian stringed instruments,” in *Proc. Indian Association of Cultivation of Sciences*, pp. **7**, 29–33, 1921.
- [4] S. B. C Vyasarayani and J. McPhee, “Modeling the dynamics of a vibrating string with a finite distributed unilateral constraint: Application to the sitar,” in *Journ. Acoust. Soc. Am*, pp. **125**, 3673–3682, 1921.
- [5] S. Siddiq, “A physical model of the nonlinear sitar string,” in *Archives of Acoustics 37.1*, pp. 73–79, 2012.
- [6] N. H. Fletcher and T. D. Rossing in *The physics of musical instruments*, Springer Verlag, 1998.
- [7] C. Geuzaine and J. F. Remacle, “Gmsh reference manual: the documentation for gmsh, a finite element mesh generator with built-in pre-and post-processing facilities.,” in URL <http://www.geuz.org/gmsh>, 2008.
- [8] J. Pommier and Y. Renard in *Getfem++, an open source generic C++ library for finite element methods.*, 2005.

Ebony vs. rosewood: experimental investigation about the influence of the fingerboard on the sound of a solid body electric guitar

Arthur Paté, Jean-Loïc Le Carrou, Benoît Fabre

LAM / Institut Jean Le Rond d'Alembert

UPMC Univ Paris 06, CNRS UMR 7190

pate@lam.jussieu.fr

ABSTRACT

Beyond electronics, lutherie also has something to do with the sound of the solid body electric guitar. The basis of its sound is indeed the conversion of the string vibration to an electrical signal. The string vibration is altered by coupling with the guitar at the neck. Electric guitar lutherie being a huge topic, this paper focuses on the influence of the fingerboard on the string vibration. An experimental study is carried out on two guitars whose only intentional difference is the fingerboard wood: ebony or rosewood. The well-known "dead spot" phenomenon is observed, where a frequency coincidence of string and structure at the coupling point leads to an abnormal damping of the note. Striking is the different behaviour of each fingerboard wood about dead spots: affected notes, as well as how much they are affected, differ with the wood.

1. INTRODUCTION

Physical studies about the solid body (without soundbox) electric guitar have been mainly focused on electronics, whether it is on the string transduction by the pickup [1,2], the effects and processing chain [3] or the amplifier [4], often with the purpose of doing numerical synthesis. The characteristics of the pickup (transducing the velocity of the string into an electrical signal), effect pedals (transforming this signal with endless possibilities), amplifier (far away from high-fidelity), loudspeakers (reproducing and distorting the final sound) are of course of significant importance. But lutherie is at least partially responsible for the sound. The vibration of the string is altered by the coupling to a moving structure (the guitar) at its ends. The structure may vibrate and exchange energy with the string, like it is the case for e.g. the classical guitar [5] or the harpsichord [6].

The coupling of a string to a structure is described in [7]. The admittance of the structure at the coupling point causes the frequencies and dampings of the coupled-string partials to differ from those in the uncoupled case (string with two rigid ends). This admittance at the coupling point is known as the "driving-point admittance". It is defined by the ratio in the frequency domain between the velocity $V(\omega)$ of

the structure at the coupling-point and the force $F(\omega)$ applied on the same point. Driving-point admittance can be obtained by classical measurements on relevant coupling points between the string and the structure, typically on the neck [8,9]. The real part of this driving-point admittance is called the driving-point conductance. It provides additional damping to the string [7]. Measurements in [9] qualitatively link a measured high conductance value at a specific frequency with the fast decay of the note at the same frequency, when fretting point and measurement point are the same. Notes affected by an abnormally big damping are known as "dead spots". Damping inhomogeneity among notes is known to be disturbing for the players.

The vibrational behaviour of the structure, seen by the string as end conditions, is influenced by many parameters. Electric guitars can differ in many things [10]: shape of the body and headstock, wood used for body, fingerboard or neck, bridge type, nut material, size and material of the frets, neck profile... Each of these lutherie parts changes the vibrational behaviour of the structure and then may alter the sound.

Fleischer and Zwicker [8] studied a *Gibson Les Paul* and a *Fender Stratocaster*, which have been the two reference models in the electric guitar industry [10]. Differences in modal behaviour are found and are attributed to the symmetry of the headstock. However, these two guitars differ not only in the headstock shape, but also in the wood species, the body shape, ...

In order to draw conclusions about the influence of a lutherie parameter, this parameter should be the only varying one. This paper is part of a broader project [11,12] aiming at studying the influence of each lutherie parameter taken separately. Here the spotlight is on the study of the fingerboard wood on the sound. An experimental investigation of ebony and rosewood fingerboards is presented. These are two out of three (the other one being maple) typical woods used for solid body fingerboards.

Section 2 gives a simple model of string-structure coupling and its consequences on string frequency and damping. Section 3 describes the experimental protocol, and quantitatively checks the model of section 2: the string damping value can be predicted from the conductance value. Section 4 discusses the change in sound induced by the change in fingerboard wood.

2. MODEL

A simple model of a stiff lossy string connected at one end to a moving body is proposed. The moving-end string model is derived as small perturbations of the stiff string model simply-supported at its two ends. The theory has already been detailed by [7] and it is briefly described here.

Let x be the axis of the string at rest position and y its motion normal to the fingerboard plane. The string is simply-supported at $x = 0$ and $x = L$. It is stretched with tension T . The string is also characterised by its mass per unit length ρ_L , second moment of area I and Young's Modulus E . The dispersion relation is:

$$\omega_n^0 = ck_n^0 \left(1 + (k_n^0)^2 \frac{EI}{2T} \right) \quad (1)$$

where $c = \sqrt{\frac{T}{\rho_L}}$ is the wave velocity, $\kappa = \sqrt{\frac{T}{EI}}$ is the stiffness term and $k_n^0 = \frac{n\pi}{L}$ is the quantized (n is a positive integer) wavenumber for simply-supported end conditions.

The string loses energy through three damping mechanisms: visco-elasticity, thermo-elasticity and air damping. Let ξ_n^0 be a damping coefficient taking into account those three damping mechanisms. It is frequency-dependent because it depends on the partial number n . The damping ξ_n^0 is generally added as the imaginary part of the complex angular frequency, so that the dispersion relation becomes:

$$\omega_n^0 = ck_n^0 \left(1 + (k_n^0)^2 \frac{EI}{2T} - 2j\xi_n^0 \right) \quad (2)$$

Yet the string is not simply-supported at its ends. One end is connected to the bridge and the other end is connected at the neck to a fret or to the nut. What [8,9] showed was checked: most of the time the motion of the end connected to the bridge is small compared to the motion of the end connected to the neck. The bridge end (at $x = 0$) is still assumed to be rigid whereas the other end (at $x = L$) is connected at the neck to the admittance of the moving guitar. The moving end at $x = L$ only causes small perturbation $\delta k_n \ll 1$ to the wavenumber k_n^0 , so that the corrected wavenumbers $k_n = k_n^0 + \delta k_n$ are used.

At $x = L$, the string's admittance is defined as the ratio between its velocity and the force being applied on it:

$$Y_{string}(L, \omega_n) = \frac{\frac{\partial y}{\partial t}(L, t)}{-T \frac{\partial y}{\partial x}(L, t)} = j \frac{\tan(k_n L)}{Z_c} \quad (3)$$

where $Z_c = \sqrt{\rho_L T}$ is the characteristic impedance of the string. At $x = L$, the string and the structure are connected and must have the same admittance. Letting $Y(L, \omega)$ be the admittance of the structure at the connection point, one has :

$$Y(L, \omega) = Y_{string}(L, \omega) \quad (4)$$

Remembering that $\tan(k_n^0 L) = 0$ ¹ and assuming that $Z_c Y(L, \omega_n) \ll 1$ ², equation 3 leads to the expression of

¹ k_n^0 is the simply-supported end solution for the wavenumber

² The impedance of the structure is much greater than the characteristic impedance of the string, resulting in a reflection of travelling waves in the string at the connection point

k_n :

$$k_n = k_n^0 + \delta k_n = \frac{n\pi}{L} - j \frac{Y(L, \omega_n) Z_c}{L} \quad (5)$$

with which equation 1 becomes:

$$\omega_n = \frac{n\pi c}{L} \left[1 + \frac{n^2 \pi^2 EI}{2TL^2} - 2j\xi_n^0 - j \frac{Y(L, \omega_n) Z_c}{n\pi} \right] \quad (6)$$

ω_n are the complex angular frequencies of a stiff lossy string having a moving end. Modal frequency is defined as :

$$f_n = \frac{\text{Re}(\omega_n)}{2\pi} = \frac{nc}{2L} \left[1 + \frac{n^2 \pi^2 EI}{L^2} \frac{1}{2T} + \frac{Z_c}{n\pi} \text{Im}(Y(L, \omega_n)) \right] \quad (7)$$

and modal damping as:

$$\xi_n = \frac{-\text{Im}(\omega_n)}{2k_n c} = \xi_n^0 + \frac{Z_c}{2n\pi} \text{Re}(Y(L, \omega_n)) \quad (8)$$

The imaginary part of the body admittance implies a shift in the simply-supported string frequencies, affecting the inharmonicity [13]. Nevertheless, measured admittance imaginary parts on the tested guitars never lead to a frequency shift larger than 1Hz. For this reason this paper only discusses the influence of the real part of the admittance, the conductance.

3. EXPERIMENTAL STUDY

The main effect of string-structure coupling is the damping due to the conductance. The experimental study identifies the conductance terms $C(\omega) = \text{Re}(Y(\omega))$ at the points where the strings couple to the structure³, that is on the fingerboard.

3.1 The two guitars of the study

This experimental study is intended to determine what differs in the sound when changing the fingerboard wood. The fingerboard should therefore be the only varying lutherie parameter. In order to fulfil this recommendation, a collaboration with instrument-makers was developed. Two guitars were made by luthiers from Itemm⁴, a french leading lutherie training-center. The two guitars follow the specifications of the *Gibson Les Paul Junior DC*, a version of one of the two most important solid body electric guitars in history [10]: original shape, quartersawn mahogany for body, neck and head, set-in neck, same equipment (bridge, bone nut, P-90 pickup). The only intentional difference between the two guitars is the fingerboard wood. One guitar has an ebony fingerboard (E) and the other one a rosewood fingerboard (R). It should be kept in mind that other parameters may differ between the two guitars, mainly because of the wood variability and the hand-made process. For schedule reasons, the guitars could not be measured prior to the gluing of the fingerboard. However samples of the woods used for the fingerboards were

³ Measurements of this section are made at the connected end of the string, so the dependence in L of $Y(L, \omega)$ is no longer specified.

⁴ <http://www.itemm.fr>

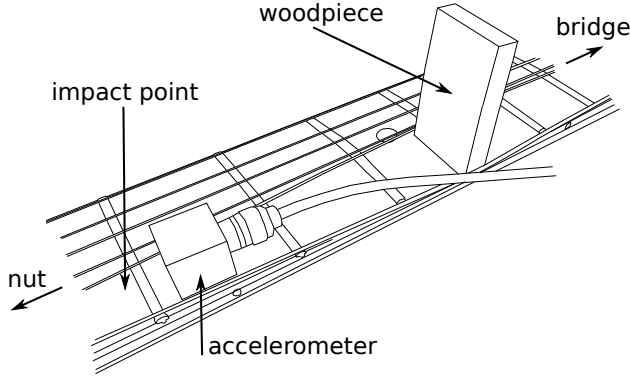


Figure 1. Setup for driving-point conductance measurement along the 5th string's axis, at a particular fret. An accelerometer is put on the one side of the fret. The hammer strikes at the other side of the fret. A very light piece of wood moves aside the strings and allows the accelerometer to stay between the two strings.

provided. Ebony density $\rho_E = 1180 \text{ kg.m}^{-3}$ and rosewood density $\rho_R = 751 \text{ kg.m}^{-3}$ are simply measured. Longitudinal Young's moduli $E_E = 3.02 \cdot 10^{10} \text{ Pa}$ and $E_R = 2.30 \cdot 10^{10} \text{ Pa}$ are identified with simple bending test. The two fingerboard woods have different characteristics: fitting the neck with one or another fingerboard wood may then change the vibrational behaviour of the instrument.

3.2 Experimental setup

The experimental setup is sketched in figure 1. The conductance is measured at every potential coupling point between string and structure, that is at every fret-string crossing on the neck. As in [9], only the conductance normal to the plane of the fingerboard is studied. Only the coupling of the string polarisation in this direction is studied in this paper. As usual, force $F(\omega)$ is applied with an impact hammer equipped with a force sensor, and velocity $V(\omega)$ is measured with an accelerometer. Impact and measurement points must be as close as possible in order to obtain actual driving-point conductance. The modal domain (where peaks and modes are well identified) is from 20Hz to 700Hz. The useful impact bandwidth is from 20Hz to 2000Hz. It is decided to consider the coupling of only the fundamental frequency with the structure, so that $n = 1$ in all equations of the section 2. The guitar is laid on elastic straps supported by a frame. Resonant frequencies of the system {frame-straps} is below the resonant frequencies of the guitars, so that this setup provides a good approximation for free boundary conditions. Modeling clay is put on the pegs and on the screw of the truss rod to prevent them from vibrating. Paper is used to avoid string vibrations, which are unwanted here for the study of the structure only.

Section 3.3 experimentally checks the model of section 2.

3.3 Validation of the model

In order to validate the model of section 2, a simple check is done. Figure 2 is an example (further discussed in sec-

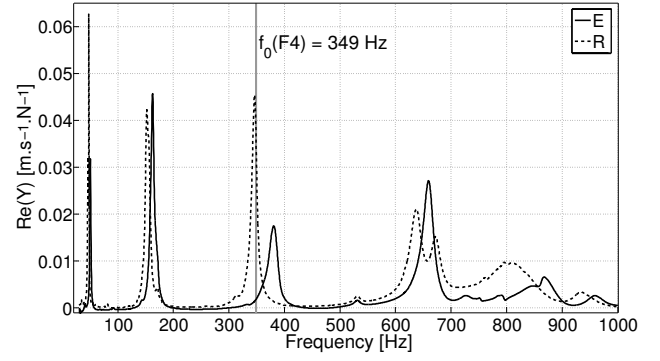


Figure 2. Driving-point conductance at the 6th fret along the 2nd string's axis. Solid line is used for the guitar E and dashed line for the guitar R. Gray line highlights the fundamental frequency of the F4 played at this place.

tion 3.4) of measured driving-point conductance: here at the 6th fret along the 2nd string for both guitars. The corresponding note is F4 with fundamental frequency $f_{F4} = 349 \text{ Hz}$. This note is also recorded by picking the string with a guitar pick, fretting the 6th fret with a capo and taking the output signal of the guitar pickup. Figure 3 shows the temporal evolution of the fundamental of this note. This temporal evolution is extracted from the recorded signal. It is obtained by computing a short-time Fourier transform of the signal and determining the envelope of the bin centered on the fundamental frequency. The time constant τ , defined as the time needed for the amplitude to get divided by Euler's number e , can be estimated from the fundamental envelope signals for guitars E and R. The estimation of the global damping terms $\xi_{F4}^E = (2\pi f_{F4} \tau_{F4}^E)^{-1} = 2.8 \cdot 10^{-4}$ and $\xi_{F4}^R = (2\pi f_{F4} \tau_{F4}^R)^{-1} = 1.4 \cdot 10^{-3}$ is straightforward. Superscripts "R" and "E" refer to guitar R and E respectively. Subscript refers to the note.

Identical string sets provided the string for both guitars. String damping ξ_{F4}^0 is then assumed to be the same for both guitars. The magnetic pickup could be a cause of string damping as well. Since the guitars are equipped with pickups of the same model series, the magnetic damping is assumed to be the same for both guitars. This magnetic damping can be included in ξ_{F4}^0 . According to equation 8 with $C(\omega) = \text{Re}(Y(\omega))$, one expects :

$$\xi_{F4}^R - \xi_{F4}^E = \frac{Z_c}{2\pi} [C^R(2\pi f_{F4}) - C^E(2\pi f_{F4})] \quad (9)$$

Estimating the conductance values from measurements presented in figure 2 ($C^E(2\pi f_{F4}) = 1.9 \cdot 10^{-3} \text{ m.s}^{-1}.\text{N}^{-1}$ and $C^R(2\pi f_{F4}) = 3.2 \cdot 10^{-2} \text{ m.s}^{-1}.\text{N}^{-1}$) leads to:

$$\xi_{F4}^R - \xi_{F4}^E = 1.1 \cdot 10^{-3} \quad (10)$$

$$\frac{Z_c}{2\pi} [C^R(2\pi f_{F4}) - C^E(2\pi f_{F4})] = 1.0 \cdot 10^{-3}$$

ξ_{F4}^0 can be identified by subtracting the term $\frac{Z_c}{2\pi} C(2\pi f_{F4})$ from the experimental ξ_{F4} . The line with crosses in figure 3 is the decay curve with identified $\xi_{F4}^0 = 2.3 \cdot 10^{-4}$.

The small difference between the two lines of equation 10 can be explained by our estimation of τ , and by the ac-

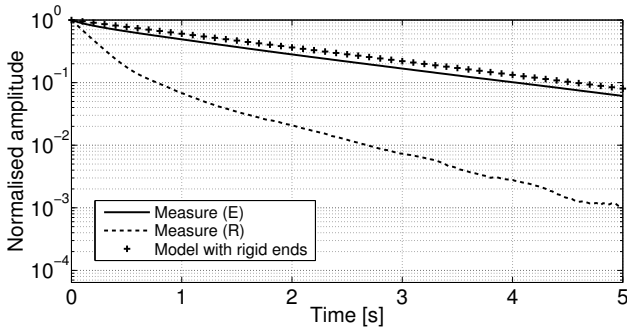


Figure 3. Temporal evolution of the fundamental frequency of the F4 (6th fret and 2nd string) played on both guitars. Solid line is for guitar E, dashed line is for guitar R, crosses show the computed decay of the same string with rigid ends.

curacy of our measurements. Nevertheless, the two values are quite close and the model of section 2 is validated.

3.4 Observation of dead spots

Figure 2 shows that at the fundamental frequency of the note, the conductance takes a low value for the guitar E. Thus the factor $\text{Re}(Y)$ in equation 8 is small and so should be the string damping due to coupling with support. This is checked in figure 3 and in the calculations of section 3.3: the experimental computed decay is close to the intrinsic decay (i.e. the rigid-ends case).

On the other hand, figure 2 shows a high conductance value for the guitar R. Figure 3 and calculations of section 3.3 confirm the "abnormal" damping of the fundamental. This damping is indeed higher for the guitar R, and the decay curve exhibits two slopes instead of the single slope decay for "normal" cases.

These two phenomena are consistently checked on the two guitars: a low conductance value leaves the note's decay unperturbed (live spot), a high conductance value makes the decay of the note shorter (dead spot).

When looking at other notes, it is found that both fingerboards exhibit dead spots. However, the note studied in this section showed a difference between the two fingerboard woods: for the same note at the same place on the neck, a guitar exhibited a live spot whereas the other exhibited a dead spot. Section 4 deals with the differences in sound that may appear between the two guitars.

4. SOUND DIFFERENCES BETWEEN THE TWO FINGERBOARDS

The two fingerboard woods lead to the same dead/live spot phenomenon. However, it does not break out the same way depending on the instrument.

4.1 Dead spot location

Section 3.4 indicates that a difference between the two guitars is the places where dead spots occur. Since the representation of figure 2 is hard to handle if one wants to have

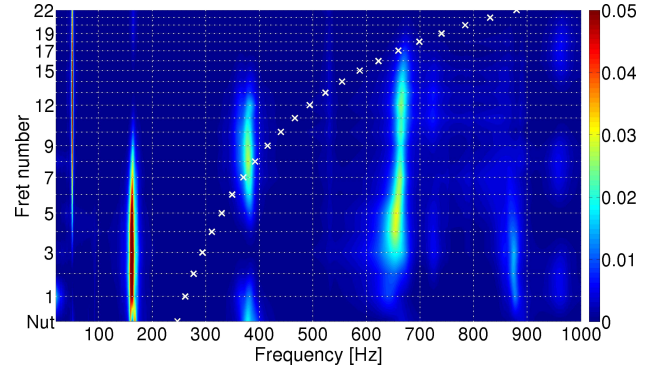


Figure 4. Guitar E: driving-point conductance values in the frequency range [20Hz–1000Hz] for all frets along the 2nd string. White crosses spot the fundamental frequencies of the notes played at each fret of the 2nd string. Unit of conductance is $m.s^{-1}.N^{-1}$.

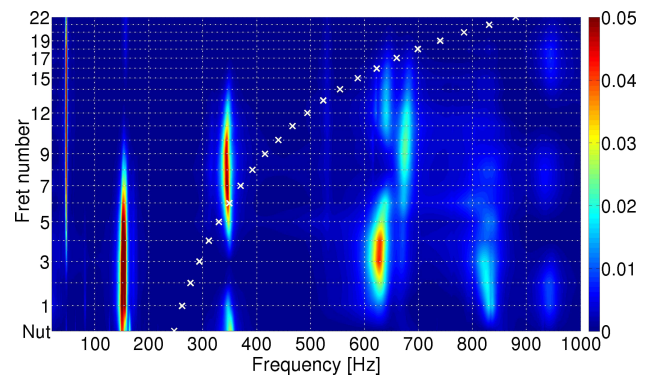


Figure 5. Rosewood-fingerboard guitar: driving-point conductance values in the frequency range [20Hz–1000Hz] for all frets along the 2nd string. White crosses spot the fundamental frequencies of the notes played at each fret of the 2nd string. Unit of conductance is $m.s^{-1}.N^{-1}$.

an overview of every fret of one string, figures 4 and 5 propose a synthetic view of the string-structure frequency coincidences, a "dead spot map". The frequency-dependant driving-point conductance values at every fret along one string are represented on the same plot. Conductance value is transcribed as a continuous color coding from blue (very low conductance) to red (conductance peak). For each fret, the fundamental frequency of the note is plotted with a white cross. Hence, whenever a white cross gets close enough to a red spot, a dead spot is reached.

Figures 4 and 5 can be used to quantify the number of dead spots of a guitar. Here on the 2nd string, the rosewood-fingerboard guitar has one dead spot (6th fret) and none at any other place. "Corresponding" dead spot for the guitar E is moved to the 7th and 8th frets. A first remark can be done, when looking at the "dead spot maps" for all six strings (five are not showed in this paper): the number of dead spots is roughly the same between the two guitars, but their locations often slightly (one or two frets) differ.

A second remark is that the conductance peaks (for example figures 4 and 5 around 400Hz and between 600Hz

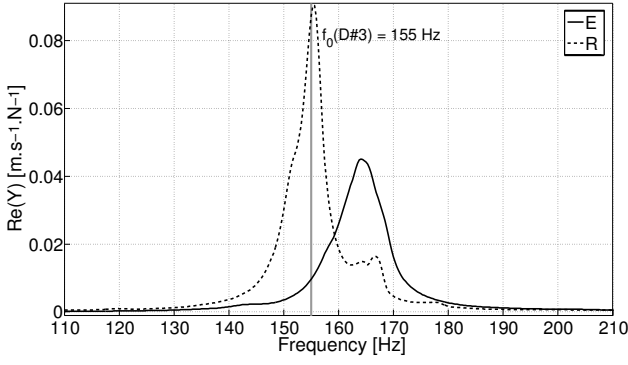


Figure 6. Driving-point conductance at the 1st fret along the 4th string's axis. Solid line is used for the guitar E and dashed line for guitar R. Gray line highlights the fundamental frequency of the D#3 played at this place.

and 700Hz) seem to be higher for the guitar R than for the guitar E. This is the purpose of section 4.2.

4.2 Dead spot dangerousness

Another kind of difference between the two fingerboard woods is the amplitude of conductance peaks, that is the potentially high damping of the note. That is what we call the "dangerousness" of a dead spot. Most of the measurements show that higher values of driving-point conductance are reached for the rosewood-fingerboard guitar. Figure 6 illustrates this tendency. It shows the measurement at 1st fret along the 4th string for both guitars. The conductance at the frequency of the note (D#3, 155Hz) is high for both guitars. Estimation of experimental damping coefficient ξ as in section 3.3 leads to $\xi_{D\#3}^E = 1.0 \cdot 10^{-3}$ and $\xi_{D\#3}^R = 2.4 \cdot 10^{-3}$ in this case. $\xi_{D\#3}^0$ is estimated as in section 3.3: $\xi_{D\#3}^0 = 3.0 \cdot 10^{-4}$. For both guitars, $\xi_{D\#3}$ is much higher than $\xi_{D\#3}^0$: this clearly reveals a common dead spot. However, the dead spot is more pronounced for the guitar R than for the guitar E. Guitar R's conductance peak is closer to the frequency of the note than guitar E's one, and guitar R also has a higher peak guitar E's one.

Whatever the tuning (determining the frequencies of the notes) is, the string-structure coupling still occurs because the neck conductance still takes non-zero values. In order to characterise this conductance amplitude difference tendency between guitars E and R in a more tuning-independent way, a mean conductance value is computed. For each measurement (each fret) along a string, the mean of the conductance is computed in the frequency range [20Hz–2000Hz]. Figure 7 presents these computed mean conductance values as a function of the place on the neck along the 2nd string. For every fret the guitar R clearly stands out from the guitar E with systematically higher mean conductance. This would mean that whatever the tuning is, the rosewood-fingerboard guitar is likely to grasp more vibrating energy from the string.

In figure 7 it can be seen that above 15th fret, the two curves become closer and the mean conductance tends to become smaller. This is because 15th fret and upper frets are close to the neck-body junction, an area where the neck

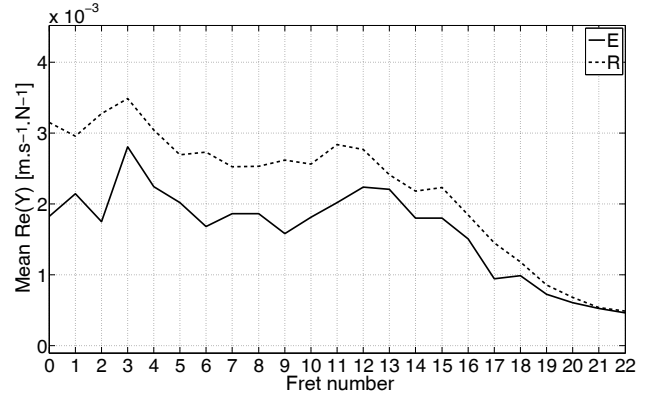


Figure 7. Mean value of conductance in the frequency range [20Hz–2000Hz] as a function of fret number/ measurement place along the 2nd string. Solid line is used for the guitar E and dashed line for the guitar R.

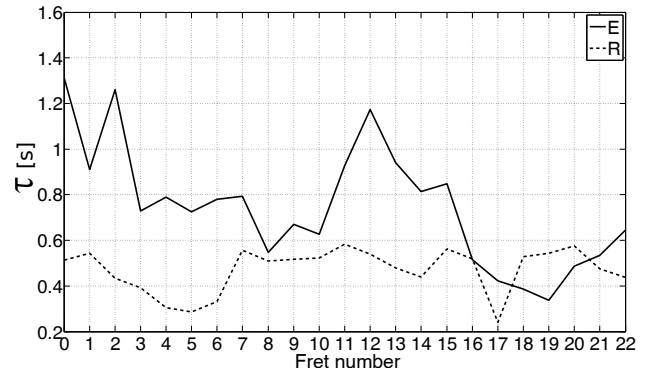


Figure 8. Computed time constants τ for each note of the 2nd string. A bandpass filter ([20Hz–2000Hz]) was applied to each note. Solid line is used for the guitar E and dashed line for the guitar R.

motion is smaller.

These mean conductances can be linked to the computed time constants τ for each note along the 2nd string showed in figure 8. This computation is slightly different from section 3.3: a bandpass-filter (20Hz to 2000Hz) is applied to the pickup signal and τ is computed from this filtered signal. Guitar E almost always has a higher time constant. The smaller damping (for every partial in the frequency range [20Hz–2000Hz]) due to the smaller mean conductance for guitar E results in a higher time constant τ .

The mean conductances and the time constants are computed along the five other strings. The tendency is confirmed: the guitar E almost always has a lower mean conductance value and a higher time constant. Rosewood might then perturb the string more than ebony.

5. CONCLUSION

Previous results on the influence of the structure on the vibration of the string have been confirmed. Because the vibrational behaviour of the electric guitar is highly dependent on the lutherie parts, which are numerous, it was decided to focus on the influence of a single lutherie parameter: the most prominent difference between the two guitars

of the study was the wood of the fingerboard (ebony or rosewood). Comparative study of sound and driving-point conductance on these two guitars indicate that the wood of the fingerboard may have an influence upon the:

- **dead spot location:** the spatial and frequency coincidence of string and guitar resonances happens at different places depending on the fingerboard wood
- **dead spot dangerousness:** when this coincidence happens, the string damping may be bigger for rosewood-fingerboard guitar

Experimental investigation about dead spots and the related discussion are naturally not only valid for the fundamental of the string but also for partials. As equation 8 shows, each string partial may couple with higher structure modes. Hence the timbre is affected by the fingerboard.

The sound differences that may be induced by the change of fingerboard wood can then have consequences in:

- **instrument-making:** the guitar maker could attempt to change the resonance coincidences: for example fingerboard thickness, shape (the so-called "slim" and "slapboard" fingerboards by Fender) or sawing angle are parameters changing the modes of the structure. Hence the instrument-maker can reduce the differences between the woods or on the contrary increase them.
- **playing:** the same note can be played at different places on the neck. Depending on the location and dangerousness of the dead spots, the player may be forced to avoid certain places on the neck and to conform his playing to the guitars' sound.
- **tuning:** actually, the frequency coincidence between the string and the structure depends on the tuning of the string. In order to avoid a too strong coupling, the guitar player can slightly change the tuning of the strings. This could be an explanation to the fact that some guitar players say that a guitar sounds better with a special tuning (e.g. all the strings a whole-tone lower) than with the standard tuning A-440Hz.

A perceptual study involving the two guitars of this paper has been carried out. The analysis is in progress and is expected to tell us to what extent the differences found here are perceptible for the guitar player.

Acknowledgments

The authors pay a special tribute to the luthiers Vincent Charrier, Loïc Keranfor, Lisa Marchand, Bela Pari, Alexandre Paul and Julien Simon and warmly thank as well Yann-David Esmans, Fred Pons and Pierre Terrien from Itemm. The efficient mediation of Vincent Doutaut made the collaboration with instrument-makers possible.

The authors also thank Rémi Blandin for his helpful contribution to the measurements on the guitars, and Benoît Navarret for fruitful discussion.

6. REFERENCES

- [1] D. Queen, "From boom boxes to Beatles, Baez and Boccherini, the electric guitar at the crossroads," in *Proceedings of the 31st Convention of the AES*, 1966.
- [2] N. G. Horton and T. R. Moore, "Modeling the magnetic pickup of an electric guitar," *American Journal of Physics*, vol. 77 (2), pp. 144–150, 2009.
- [3] O. Kröning, K. Dempwolf, and U. Zölzer, "Analysis and simulation of an analog guitar compressor," in *Proceedings of DAFx11*, 2011.
- [4] I. Cohen and T. Hélie, "Real-time simulation of a guitar power amplifier," in *Proceedings of DAFx10*, 2010.
- [5] C. Lambourg and A. Chaigne, "Measurements and modeling of the admittance matrix at the bridge in guitars," in *Proceedings of SMAC 93*, 1993, pp. 448–452.
- [6] N. H. Fletcher, "Analysis of the design and performance of harpsichords," *Acustica*, vol. 37, pp. 139–147, 1977.
- [7] A. Hirschberg, J. Kergomard, and G. Weinreich, *Mechanics of musical instruments*. Wien-New York: Springer Verlag, 1995.
- [8] H. Fleischer and T. Zwicker, "Mechanical vibrations of electric guitars," *Acta Acustica united with Acustica*, vol. 84, 1998.
- [9] H. Fleischer and T. Zwicker, "Investigating dead spots of electric guitars," *Acta Acustica united with Acustica*, vol. 85, pp. 128–135, 1999.
- [10] A. Paté, B. Navarret, R. Dumoulin, J.-L. Le Carrou, B. Fabre, and V. Doutaut, "About the electric guitar: a cross-disciplinary context for an acoustical study," in *Proceedings of Acoustics 2012, Nantes, France*, 2012, pp. 2717–2722.
- [11] J.-L. Le Carrou, J. Frelat, A. Mancel, and B. Navarret, "Guitare électrique : quel rôle pour les éléments de lutherie ?" in *10ème Congrès Français d'Acoustique*, 2010.
- [12] A. Paté, J.-L. Le Carrou, B. Navarret, D. Dubois, and B. Fabre, "A vibro-acoustical and perceptive study of the neck-to-body junction of a solid-body electric guitar," in *Proceedings of Acoustics 2012, Nantes, France*, 2012, pp. 2711–2716.
- [13] C. E. Gough, "Acoustical studies of stringed instruments using string resonances," in *Proceedings of SMAC 83*, 1983, pp. 19–45.

HIGH-SPEED CAMERA DISPLACEMENT MEASUREMENT (HCDM) TECHNIQUE OF STRING VIBRATIONS

Niko Plath

Institute of Musicology at the University of Hamburg

Neue Rabenstr. 13

20354 Hamburg, Germany

niko.plath@systematicmusicology.de

ABSTRACT

In this paper we describe a method for string vibration measurement through analysing high-speed camera recordings (HCDM). A short introduction into most common string motion detection methods is given and discussed. The experimental design for HCDM is described, covering the necessary settings needed to produce quantifiable recordings. The applied method of data analysis can be summarized as: String motions are converted to time series by tracking of string positions in every discrete time step through searching for high contrasts of RGB values in a row of recorded pixels. Examples of measurements of string motion are presented: The formation of the Helmholtz motion in a bowed violin string, the transversal versus the torsional motion of a plucked vihuela string, the transversal versus the longitudinal motion of a struck piano string and the effect of the una corda pedal for the transversal motion of three unison piano strings. Benefits and drawbacks of the presented method are discussed followed by possible improvements for future research.

1. INTRODUCTION

Among the earliest published measurement methods for string vibrations Hermann von Helmholtz describes the characteristic motion of bowed strings [1]. He uses a method invented by Jules Antoine Lissajous he calls the "Vibrationsmikroskop" [1, p.138], a modified Lissajous apparatus. The superposition of the string vibration and a perpendicularly orientated vibration provided by a tuning fork is observed. If one vibration has the same frequency or is an integral multiple of the other vibration the superposition gives a stationary fixed curve.

Today electrodynamic methods are common for visualizing string motion. Magnets attached close to the string supply a magnetic field. The motion of the string induces a voltage proportional to its vertical velocity. By integrating the velocity signal the displacement of the string can be obtained. A discussion of the benefits and limits of electrodynamic string detection is given in [2]. [3] uses a

microphone preamplifier placed near a string as a capacitive transducer. The grounded string works as the second capacitor plate. For electrodynamic methods a few principal drawbacks have to be considered: They work well with metal strings but do not work with nylon strings. Additionally both described measurement methods influence the behaviour of the measured system: A current induced in a string introduces additional damping of its vertical motion. The width of the electrodynamic detector introduces an averaging of the velocity over the measured string segment. This implies an upper frequency limit for the detection. Attaching an acceleration transducer adds weight to the measured object resulting in a shift of the fundamental frequency of the object.

As a non-invasive tool a high-speed camera is a powerful device to visualize mechanical processes in an instrument. The possibility to watch parts of the instrument move as it produces sound helps to gain insight into its complex vibrational motions - the behaviour of the system is directly visible. With the presented method it is possible to produce evaluable data obtained by high-speed recordings of moving strings. It is non-invasive, meaning the behaviour of the system is not changed due to measurement. Current works on high-speed string detection include [4] measuring the movements of a piano hammer striking a string by tracking applied markers on hammer and string. [5, 6] describe a high-speed detection method for tracking a finger plucking a harp string.

2. METHOD

2.1 Equipment

For the recordings a *Vision Research Phantom V711* high speed camera is used. The monochrome sensor has 1280 x 800 pixels, maximum frame rate at full resolution is 7350 fps. At minimum resolution (128 x 8) maximum frame rate is 1,400,000 fps. Table 1. shows the relationship between possible image resolution and time resolution dependent of 7 Gpx/s maximum throughput. The recordable fps rate is limited by the possible provided illumination intensity. A fps rate of 200 kHz implies a maximum exposure time of 5 μ s. With the two used GSVitec MultiLed LT V6 LED-based lamps it is possible to record frame rates of up to 200 kHz. At maximum throughput the 16 GB internal RAM of the camera allows recordings of 1.5 s length. The maximum amplitude resolution of a captured motion is lim-

Resolution	fps
1280 x 800	7,530
1280 x 112	48,000
1280 x 8	650,000

Table 1. Possible maximum frame rates for some resolutions resulting in approximately 7 Gpx/s throughput.

ited by the number of pixels in a row or column - in this case 1280, which is equivalent to a resolution of a little more than 10 bit. The resolution can be enhanced subsequently by a factor of 10-40 with sub-pixel discretization algorithms. Implementing sub-pixel tracking in the data analysis process is subject of present work but not presented here. With an amplitude of 1280 pixels and a field of view of 15 mm x 24 mm a resolution of approx. 20 μm is possible. The use of a Tokina AT-X M100 Pro D (100mm F2.8) macro objective lens allows a distance to the measured object of down to ≈ 12 cm. Based on optical laws some trade-off decisions must be made: The bigger the objective diaphragm the less illumination is needed (so higher fps rates are possible) but the smaller is the depth of field (DOF). The longer the focal length the bigger is the magnification but the smaller is the DOF. The smaller the distance to the object the bigger is the captured motion but the smaller is the DOF. Because the subsequent analysis needs defined edges in the recorded images the movement has to be in the DOF range. That often leads to multiple readjustments of the experimental design.

2.2 Procedure

This section describes the procedure of filming with its special requirements to experimental design. Because the captured amplitude resolution is limited by the resolution the camera sensor supports it is important to capture the amplitude of the movement with the largest possible image section. Vibrating parts of musical instruments move with high frequency but small amplitude. Instrument strings often have amplitudes of a fraction of a centimeter, depending on the used material (steel, plastic), diameter, length and the applied tension demanded for efficient sound production. Therefore the camera is placed as close to the measured object as possible in consideration of the focal length of the used objective lens. Due to the fact that the implemented motion detection algorithm searches along a pixel row or column the camera has to be placed in an angle that the captured motion is parallel to the rows or columns of the camera sensor.

A one dimensional oscillation with amplitude A^+ should be shot orthogonal to the direction of motion. This angle is defined $\alpha = 0^\circ$. For $\alpha \neq 0^\circ$ the captured projection is $A^\angle = \cos \alpha \cdot A^+$.

A two dimensional oscillation with amplitude A^{v+} orthogonal to the angle of view ($\alpha = 0^\circ$) and amplitude $A^{h\parallel}$ parallel to the angle of view and $A^{v+} \perp A^{h\parallel}$ is defined (See Fig.1). The captured projection is

$$A^{v\angle} = \cos \alpha \cdot A^{v+} \quad (1)$$

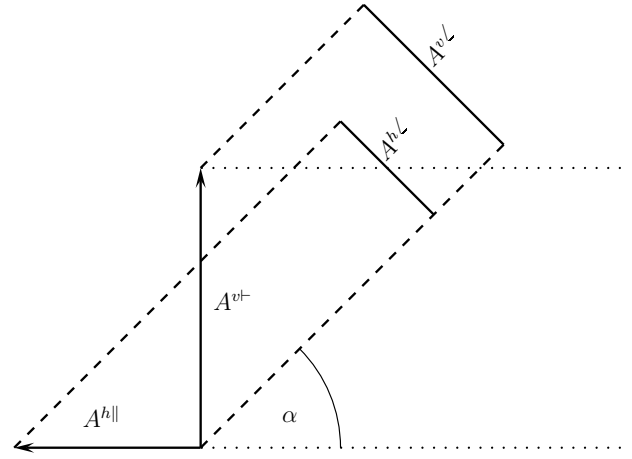


Figure 1. Captured amplitudes of the two transverse polarizations when shot with an angle $\alpha \neq 0^\circ$.

and

$$A^{h\angle} = \sin \alpha \cdot A^{h\parallel} \quad (2)$$

For $\alpha = 0^\circ$ the failure for A^{v+} is min ($A^{v\angle} = A^{v+}$), the failure for $A^{h\parallel}$ is max ($A^{h\angle} = 0^\circ$). For $\alpha = 90^\circ$ the failure for A^{v+} is max ($A^{v\angle} = 0$), the failure for $A^{h\parallel}$ is min ($A^{h\angle} = A^{h\parallel}$). For $\alpha \neq 0^\circ, 90^\circ$ the captured projection is an unknown combination of both polarizations. All measured instruments show transversal string oscillations in two polarizations. For the piano the interaction between vertical and horizontal polarization of the transversal motion is essential for producing its characteristic decay [7]. With the equipment used here it is not yet possible to capture both polarizations simultaneously. Experiments with mirrors do not work satisfactory because of a too small DOF. Capturing the vertically polarized transverse motion of several neighbouring strings is possible by shooting with an $\alpha > 0^\circ$. α should be large enough to project space between the strings (overlapping strings cant be detected uniquely) but as small as possible to maximize $A^{v\angle}$ and minimize the failure (influence of $A^{h\angle}$).

2.3 Data Analysis

After a sequence is recorded the result can be inspected via a control software. Important questions regarding the quality (calculability) of the recording are the following: Does the string leave the frame? In these frames the string position can not be tracked properly. Does the camera or the whole instrument move? That would add a low-frequency offset to the recorded time series. Is the lighting adequate for the used exposure time?

If the recording quality is satisfactory the strings position can be tracked for every frame. Each recorded frame is stored as single files (8 bit RGB monochrome)¹. An algorithm written in *Mathematica* imports all frames and searches in a selected row (or column, depending on the direction of the movement) for the strings by checking the RGB values in that row (column). Fig.2 depicts one recorded

¹ In a monochromatic picture all three RGB channels have the same value.

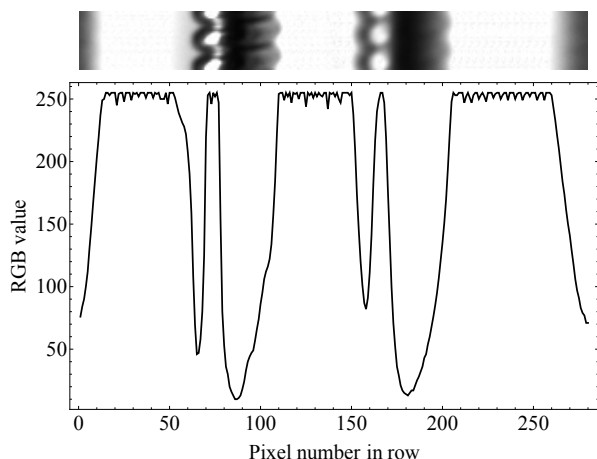


Figure 2. RGB values (range 1-255) of the first pixel row in the captured frame above. The strings positions are clearly visible and can be tracked over a threshold.

frame showing two unison strings of a grand piano. The diagram below the picture shows the RGB value for each pixel in one row. A manually set threshold value is used to find the edge between the background and the string. This edge is tracked for every given frame. Each measured position is stored in a list, giving a time series for the measured point on the string.

2.4 Results

In this section some examples of measurements for different string instruments are presented and compared to current publications.

2.4.1 Bowed string - Violin

Fig.3 shows the transverse motion of a bowed violin string captured with a frame rate of 35 kHz. The captured motion is in good accordance to classical as well as recent violin research publications. Raman [8] is the first to formulate an idealized model for the formation of the characteristic violin string motion Helmholtz [1] observed and described². Cremer [9] and Fletcher and Rossing [10] give an extensive overview on recent violin research.

2.4.2 Plucked string - Mexican Vihuela

Torsional motion of strings can be found in many instruments like for instance violins [9–11], harps [5, 11] or guitars [12]. The mexican vihuela also shows such modes of vibration that were measured with the proposed method. By marking one string with a pen and tracking the movement of this point it is possible to measure its torsional motion. To measure the torsion two transverse movements are tracked:

1. The contrast between the string and the background
2. The contrast between the background and the point on the string

² In honour of Hermann von Helmholtz the characteristic motion of the violin string is called *Helmholtz-motion*.

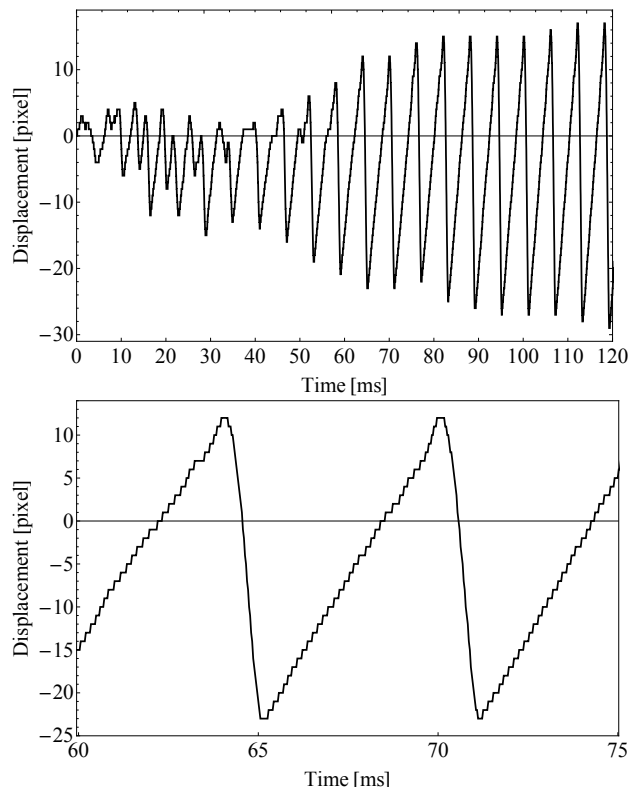


Figure 3. Transient formation of Helmholtz motion in a bowed violin string.

Both resulting time series are subtracted from each other yielding the torsional motion.

Fig.4 shows the transversal versus the torsional motion of a mexican vihuela string captured with 40 kHz. The torsional motion is scaled by a factor of 1.5 for a better visibility. The amplitude of the torsion is too small to determine spectral information but shows a frequency of 2.86 times the transversal frequency (transversal: 280 Hz, torsional: 800 Hz). The appearance of the torsional vibration strongly depends on the way the vihuela string is plucked. The string documented in Fig.4 is plucked with the finger tip. In this case the skin twists the string torsionally before releasing it. After release the torsional motion vanishes within 100 ms. When plucked with a finger nail the torsion is not measurable. The question if and how this impacts the process of transient sound production can not be answered here and is of interest for further research.

2.4.3 Struck string - Piano

Besides transversal and torsional motions, longitudinal string vibrations can be found in most string instruments. Particularly for the lower notes of the piano they contribute significantly to the tonal character [13, 14]. Conklin measured the pitch relation of the transverse and longitudinal modes and showed its influence on the quality of the tone. Listening tests suggest that longitudinal components are audible up to C_5 (523 Hz) [15]. Fig.5 shows the transverse and parts of the longitudinal motion of a forte played A_1 tone of a grand piano. The transverse motion is detected as described above. For time intervals when the transverse im-

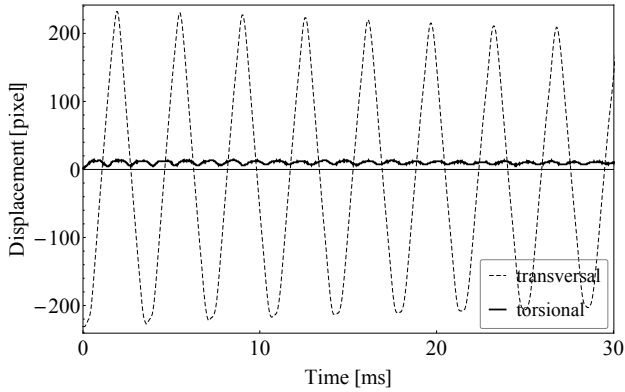


Figure 4. Transversal vs. torsional motion of a plucked vihuela string.

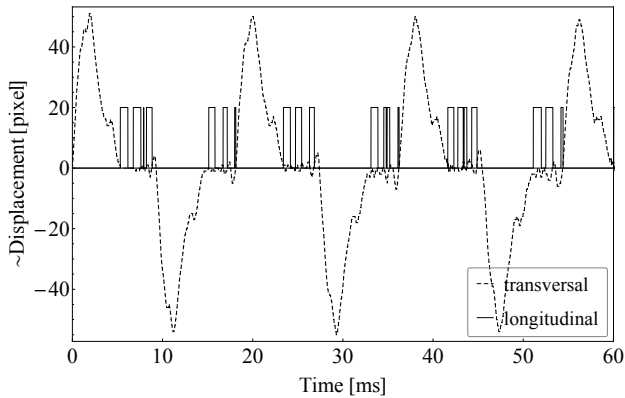


Figure 5. Transversal vs. longitudinal motion of a struck piano string.

pulse, travelling up and down the string, is not crossing the observation point the string is nearly motionless transversally. In these time intervals the longitudinal motion can be detected. With the given set-up the longitudinal amplitude resolution is only 2 bits and does not allow spectral analysis. By counting the peaks of the longitudinal motion as shown in Fig.5 one can approximate the pitch relation of transversal and longitudinal motion. The longitudinal period is 12,5 times shorter than the transversal which is consistent with Conklins measurements.

For Fig.6 a forte G_3 on a grand piano is recorded. The use of the una corda pedal shifts the action to the right, so that the hammer only strikes two of three unison strings. This results in a softer note and a changed timbre. In terms of the two characteristic piano decay components described by Weinreich [7] the una corda pedal weakens the mainly vertical polarized attack component and enhances the mainly horizontal polarized long tail component. Because the action is shifted the contact point between the hammer and the string is at a softer part of the hammer felt tip. This contributes to a softer tone quality due to damping of higher modes resulting from a longer lasting hammer string contact time [16]. The strings in Fig.6 are shot with an $\alpha \approx 15^\circ$.

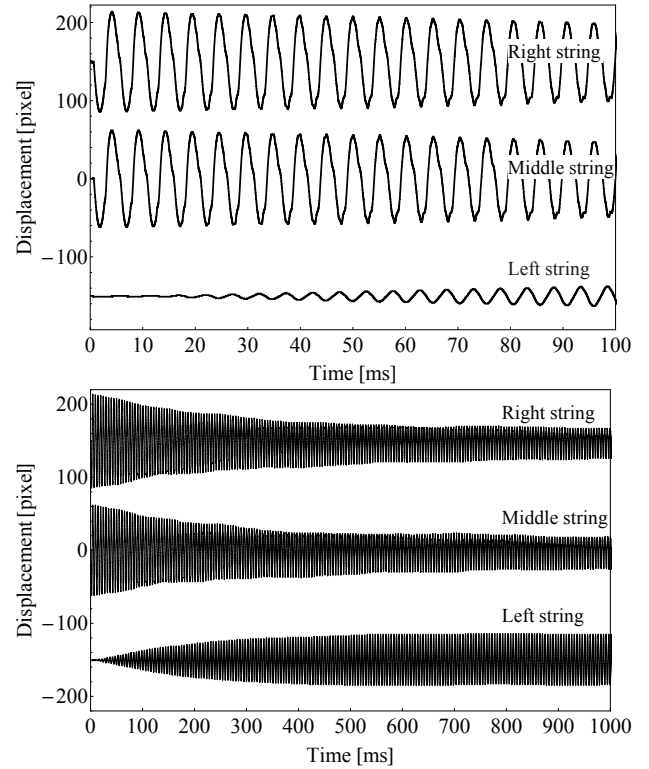


Figure 6. Effect of the una corda pedal. Two of three unison strings are struck, the left string obtains energy only from the bridge resulting in a slow attack.

3. DISCUSSION

A high-speed camera can be used in instrument acoustics as a valuable tool for visualization of acoustical vibrations. In many cases it gives a qualitative insight into the behaviour of vibrating instrument parts. The presented HCDM method is able to produce evaluable data on the basis of recorded string motion at sample rates as high as 200 kHz, which is more than sufficient for most audio purposes. It is non-invasive, easy (intuitive) to use and works with nylon strings. Aspired enhancements of the method for future work are the following:

- Implementation of sub-pixel discretization to increase the amplitude resolution.
- Comparison of different tracking algorithms.
- Increasing the depth of field to allow the application of mirrors to enable three dimensional tracking.
- Synchronization with other measurement methods such as microphone array measurements or laser speckle interferometry.

4. REFERENCES

- [1] H. von Helmholtz, *Die Lehre von den Tonempfindungen als psychologische Grundlage für die Theorie der Musik*. Vieweg, 1870.
- [2] A. Askenfelt and E. V. Jansson, "From touch to string vibrations. iii: String motion and spectra," *The*

- Journal of the Acoustical Society of America*, vol. 93, no. 4, pp. 2181–2196, 1993. [Online]. Available: <http://link.aip.org/link/?JAS/93/2181/1>
- [3] X. Boutillon, “Model for piano hammers: Experimental determination and digital simulation,” *The Journal of the Acoustical Society of America*, vol. 83, no. 2, pp. 746–754, 1988. [Online]. Available: <http://link.aip.org/link/?JAS/83/746/1>
- [4] S. Birkett, “Experimental investigation of the piano hammer-string interaction,” *The Journal of the Acoustical Society of America*, vol. 133, no. 4, pp. 2467–2478, 2013. [Online]. Available: <http://link.aip.org/link/?JAS/133/2467/1>
- [5] D. Chadeaux, J.-L. L. Carrou, and B. Fabre, “A model of harp plucking,” *The Journal of the Acoustical Society of America*, vol. 133, no. 4, pp. 2444–2455, 2013. [Online]. Available: <http://link.aip.org/link/?JAS/133/2444/1>
- [6] D. Chadeaux, J.-L. L. Carrou, B. Fabre, and L. Daudet, “Experimentally based description of harp plucking,” *The Journal of the Acoustical Society of America*, vol. 131, no. 1, pp. 844–855, 2012. [Online]. Available: <http://link.aip.org/link/?JAS/131/844/1>
- [7] G. Weinreich, “Coupled piano strings,” *The Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1474–1484, 1977. [Online]. Available: <http://link.aip.org/link/?JAS/62/1474/1>
- [8] C. V. Raman, “On the mechanical theory of vibrations of bowed strings,” *Indian Assoc. Cult. Sci. Bull.*, vol. 15, pp. 243–276, 1918.
- [9] L. Cremer, *Physik der Geige*. S. Hirzel Verlag, 1981.
- [10] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*. Springer Verlag, 1998.
- [11] T. D. Rossing, *The Science of String Instruments*, T. D. Rossing, Ed. Springer Science+Business Media, 2010.
- [12] M. Pavlidou and B. E. Richardson, “The string-finger interaction in the classical guitar,” in *Proc. ISMA*, Dourdan, 1995.
- [13] H. A. Conklin, “Design and tone in the mechanoadoustic piano. part iii. piano strings and scale design,” *The Journal of the Acoustical Society of America*, vol. 100, no. 3, pp. 1286–1298, 1996. [Online]. Available: <http://link.aip.org/link/?JAS/100/1286/1>
- [14] A. Askenfelt, *Five Lectures on the Acoustics of the Piano*, ser. Publications issued by the Royal Swedish Academy of Music. Kungl. Musikaliska Akademien, 1990.
- [15] B. Bank and H.-M. Lehtonen, “Perception of longitudinal components in piano string vibrations,” *The Journal of the Acoustical Society of America*, vol. 128, no. 3, pp. EL117–EL123, 2010. [Online]. Available: <http://link.aip.org/link/?JAS/128/EL117/1>
- [16] A. Askenfelt and E. Jansson, “From touch to string vibrations—the initial course of the piano tone,” *The Journal of the Acoustical Society of America*, vol. 81, no. S1, pp. S61–S61, 1987. [Online]. Available: <http://link.aip.org/link/?JAS/81/S61/1>

FINITE ELEMENT MODEL OF A KANTELE WITH IMPROVED SOUND RADIATION

Henna Tahvanainen, Henri Penttinen, Vesa Välimäki

Department of Signal Processing and Acoustics

School of Electrical Engineering

Aalto University

henna.tahvanainen@aalto.fi

Jyrki Pölkki

Soitinrakentajat AMF

Tuikkalantie 2

Leppävirta

jyrki.polkki@estelle.fi

ABSTRACT

In this paper, a plucked string instrument called the kantele is modelled with the finite element method. The aim is to compare two traditional body structures and a modified body structure in terms of vibrational modes and radiation efficiency. The two traditional body structures are the closed box kantele and the top-plate kantele. In the modified structure, the top plate and the back plate are separated with an air gap. The modified structure has more vibrational modes than the traditional body structures, because it incorporates a freely vibrating top plate coupled with enclosed air. The simulations show that when the air gap is between 1-3 mm, the radiation efficiency of the modified kantele is higher than that of the traditional kanteles. This result supports previous research that concluded the modified kantele to be louder than the traditional top-plate kantele.

1. INTRODUCTION

The kantele is a zither-like plucked string instrument. There are 16 peoples in North East Europe and West Siberia that play this kind of a zither. Consequently, instruments with various string arrangements and terminations, as well as shapes and sizes of the sound box belong to the kantele family. The various kanteles are currently being documented in *The Kindred of Kantele-* project ¹.

The common parts for all the kanteles include a wooden sound box, metal or nylon strings, and tuning pins made of metal or wood. The amount of strings varies from 5 to 40. In addition, there may be a sound hole. The common acoustical feature for the kanteles is the special fluctuating timbre caused by the knotting of string around the non-rigid tuning pins [1, 2]. Initially, the kantele was carved from a single piece of wood with no back plate. Later, the carved kantele was turned into a box by adding a soundboard to the carving. These two structures are henceforth referred to as the top-plate kantele and the box kantele, respectively. Nowadays, the kantele is assembled by glueing wooden

¹ <http://kanteleenkielin.maanite.fi/2013/03/31/kindred-of-kantele-overview-in-english/>

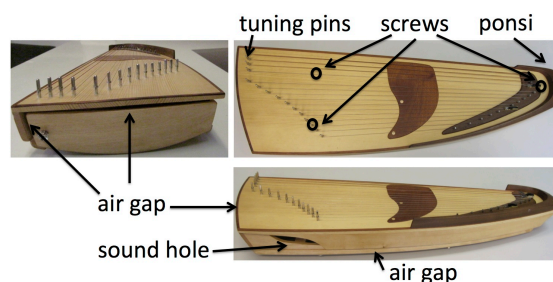


Figure 1. The modified kantele with the air gaps indicated.

plates together. In addition to the two traditional kanteles, a kantele with a freely-vibrating top plate isolated from the back plate with an air gap has been developed [3, 4]. It is called the modified kantele and it is shown in Fig. 1.

The kantele with its complex shape, enclosing and surrounded by air, represents a coupled structure-acoustic problem. To tackle such problems, wave-based modelling methods, such as finite element, boundary element, and finite difference methods, have been used extensively in string instrument acoustics [5, 6]. Other numerical methods include the fictitious domain method applied to the guitar [7].

The finite element method (FEM) is particularly suitable for complicated shapes, especially when the eigenfrequencies of the solid domain and enclosed air are comparable in frequency. Of the family of string instruments, FEM has been applied successfully at least to the guitar [8–13], violin [14–16], and piano [17–20]. The early models considered mainly the instrument soundboards in the low frequency region [6], or instrument-shaped air-cavities [16]. In addition, different stages of construction were simulated [9, 15]. More recently, FEM has been applied to structure-acoustic coupling in string instruments [10–12, 14]. Both enclosed and surrounding air add to the mass of the instrument and thus lower the eigenfrequencies of the string instrument [10, 14]. The mode shapes, on the other hand, are much less affected.

This paper aims at extending the understanding of the contribution of the isolated top plate to the increased loudness of the modified kantele by using FEM. The vibrational modes and sound radiation of three kantele structures are studied: closed box kantele, top-plate kantele, and modified kantele. Additionally, the radiation efficiency for each kantele structure is computed. Moreover, the effect of size of the air gap in the modified kantele is investigated.

This paper is organised as follows. In Sec. 2, the background of the modified kantele is described. In Sec. 3, the

finite element method is briefly reviewed. Section 4 describes the FE-model of the kantele and the results of the simulations are presented in Sec. 5. Section 6 concludes the paper.

2. BACKGROUND OF THE MODIFIED KANTELE

Originally, the kantele was used as a rune-accompaniment in small farm houses. Nowadays, it is played in various ensembles in large spaces. Consequently, it is beneficial to develop an acoustically louder version. A kantele with improved sound radiation was presented at SMAC 2003 by Pölkki et al [3]. The idea for this modified kantele came from old museum copies with no back plate. Interestingly, these seemed to sound louder in a soft-padded case rather than on lap where they are usually played. Based on these observations, in 1999 Pölkki began experimenting with different lightweight wooden structures, which would do the same. The first prototypes with 11, 15, and 19 strings were made. Adjusting the air gap between the plates was a delicate process in order to obtain a favourable timbre. In addition, the larger kanteles benefited from a small sound hole in the gap. Based on the prototypes, it was clear that changes in the top plate were needed if the strong attack and the long decay of the characteristic kantele sound were to be preserved. Consequently, the top plate was made three times thicker and strengthened around the tuning pins.

The collaboration between Pölkki and Penttinen lead to verified design rules for the kantele with improved sound loudness [4]. The rules are 1) increasing string tension, 2) increasing radiating surface area, and 3) isolating the top plate from the back plate with an air gap. In addition, it was confirmed by listening tests that this modified design was on average 3 dB louder than the top-plate kantele. In the meanwhile, the modified design has been fine-tuned and sold commercially for 12 years. The commercial version, *Kirjokansi*, is shown in Fig. 1. The top plate and the back plate are attached together with three screws. The kantele has 11+4 strings spanning the fundamental frequency range of 73-588 Hz. The four lowest strings are used as drones. Finally, the isolating air gap has lately been applied to the 40-string kantele, and based on informal listening tests after fine-tuning, it now sounds warm and rich in bass.

3. FINITE ELEMENT METHOD

In FEM, the continuous domain is discretised into a finite set of solid elements, which each are approximated with simple shape functions such as polynomials [21]. The continuous field, such as the pressure, is then evaluated at the vertices of each element called nodes. At least six elements per bending wavelength are required in order to ensure sufficient accuracy of the simulation [22]. For structure-acoustic vibrations, the equations of motion translate into the following matrix

$$\begin{bmatrix} M_s & 0 \\ -H_{sf}^T & M_f \end{bmatrix} \begin{bmatrix} \ddot{d}_s \\ \ddot{p}_f \end{bmatrix} + \begin{bmatrix} D_s & 0 \\ 0 & D_f \end{bmatrix} \begin{bmatrix} \dot{d}_s \\ \dot{p}_f \end{bmatrix} + \begin{bmatrix} K_s & H_{sf} \\ 0 & K_f \end{bmatrix} \begin{bmatrix} d_s \\ p_f \end{bmatrix} = \begin{bmatrix} F_b \\ F_q \end{bmatrix} \quad (1)$$

where M , D , K , and F , are the global matrices of mass, damping, stiffness, force, and p and d are the pressure and

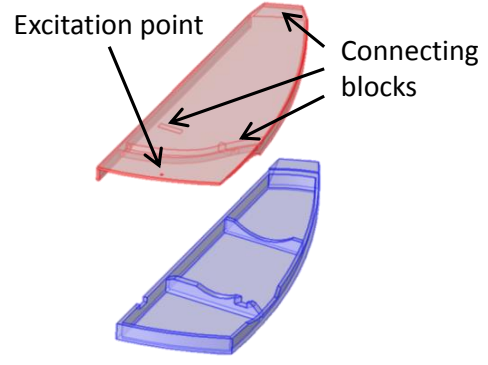


Figure 2. The top plate and the back plate of the kantele for the FE-model. The top-plate kantele considers the top plate only, the box kantele is obtained by attaching the plates together, so that the side edges are fixed. In the modified kantele, the plates are attached together with the three connecting blocks allowing an air gap between the plates.

the displacement at nodal points of the respective domain (subscript s for structure, f for fluid). H_{sf} is the coupling matrix between the two domains. At the structure-acoustic boundary, two conditions for the coupling apply. Firstly, the particle velocity perpendicular to the boundary should be continuous. Secondly, the air exerts a load on the structure, which is equal to the inverse of the air pressure. Finally, when modelling the sound radiation of an instrument in the free field, the calculations are often limited over a finite volume. In order to avoid artificial reflections, the boundaries are modelled as so called radiating boundaries, or using a technique called Perfectly Matched Layer (PML). The numerically more accurate boundary condition, PML, is an additional medium surrounding the computational domain that matches the domain and absorbs all the incoming radiation in all angles with no reflections [23].

4. SIMULATIONS

In this paper, three kantele structures were considered: the top-plate kantele, the box kantele, and the modified kantele with air gaps of several sizes (1 mm, 3 mm, 5 mm, 7 mm). Both eigenfrequency and forced vibration analysis were conducted using an existing finite element software (Comsol Multiphysics) in order to study the vibrational modes and the radiation of the kantele, respectively. The modelling steps taken are described in the following.

1. Importing/drawing the geometry

In these simulations, strings, tuning pins, and bars parallel to the top plate were omitted. The back plate was modelled flat. The three plastic screws were substituted with wood blocks. The top plate and the back plate of the modelled kantele can be seen in Fig. 2. In addition, the kantele was surrounded by air ($r = 0.7$ m).

2. Defining material properties

The kantele was assumed to be made out of one orthotropic material, spruce. The material properties

Parameter	Symbol	Value	Unit
Elastic moduli	E_x	122	$[10^8 \text{ Pa}]$
	E_y	15.6	
	E_z	7.20	
Shear moduli	G_{xy}	15.1	$[10^8 \text{ Pa}]$
	G_{yz}	1.22	
	G_{xz}	14.6	
Poisson ratios	ν_{xy}	0.42	[1]
	ν_{yz}	0.53	
	ν_{xz}	0.46	
Density	ρ_s	440	$[\text{kgm}^{-3}]$
Isotropic loss factor	α	0.01	[1]
Driving force amplitude	A	1.5	[N]
Air density	ρ_0	1.2	$[\text{kgm}^{-3}]$
Speed of sound in air	c_0	343	$[\text{ms}^{-1}]$

Table 1. Parameter values inserted in the FE-model, the material values are obtained from [24].

obtained from Kretschmann [24] and other parameters are listed in Tab. 1.

3. Defining boundary conditions

The air was surrounded by a PML ($r = 0.8 \text{ m}$). The two structure-acoustic boundary conditions according to Sec. 2 were applied.

4. Meshing

The chosen element size was limited by computational capacity. Tetrahedral elements were used, with the size range of 1.5-10 cm. Such meshing resulted to 60 000-200 000 elements depending on the kantele structure. Thus, the simulated frequency range was limited to 570 Hz, which covers the tuning range of the kantele strings.

These four steps are enough to perform an eigenfrequency analysis. In order to study radiation, the following additional steps for the forced oscillation analysis are required:

5. Damping coefficient

The internal damping of the wood material was considered isotropic and frequency-independent for simplicity.

6. Type and location of the force

The kantele was driven sinusoidally by approximating the area and the location of the fifth tuning pin indicated in Fig. 2.

7. Frequency steps at which the analysis is performed

The frequency step was 5 Hz due to limited computational time.

5. RESULTS

5.1 Comparison of the kantele structures

The modelled frequency responses of the three kantele structures at 0.7 m above the top plate are shown in Fig. 3. Based on the figure, the structures differ at least in overall sound pressure level, mode density, and mode frequencies. The box kantele has a lower sound pressure level than the top-plate and the modified kantele, except at regions around 295 Hz and 555 Hz. The modal density between 70-570 Hz is 8, 9 and 11 for the box, top-plate, and modified kantele, respectively. It is desirable to have many modes in the tuning range of strings for better sound quality and sound pressure level [25]. This is one of the factors that explain

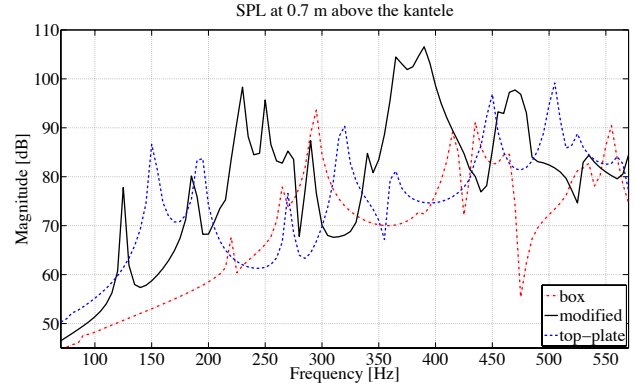


Figure 3. The sound pressure level at 0.7 m above the kantele for three different kantele structures: box kantele (dash-dotted red line), modified kantele (solid black line) and top-plate kantele (dashed blue line).

the increased loudness of the modified kantele compared to the traditional top-plate kantele.

The mode density for the modified kantele is higher than for the two traditional kanteles because the modified kantele has a freely vibrating top plate and the air enclosed by the two plates which can couple to the plate vibrations. An example of this case is illustrated in Fig. 4. Figure 4a shows the first two normal mode shapes of the top-plate kantele (at 150 Hz and 195 Hz), Fig. 4b the first significant normal mode of the box kantele (at 295 Hz), and Fig. 4c the first three normal modes of the modified kantele (at 125 Hz, 185 Hz, and 230 Hz). The color legend indicates the total displacement in millimeters. The values above and below the legend are the maximum and minimum values for the displacement. It can be seen that shapes of the first two top plate modes correspond to the shapes of the first two normal modes of the modified kantele. The frequencies of these modes have lowered; from 150 Hz to 125 Hz, and from 195 Hz to 185 Hz. In addition, the vibration amplitudes (total displacement) are smaller for the modified kantele than for the top-plate kantele. In particular, the ponsi vibrates less. The back plate has a solid ponsi (see Fig. 1), which is fixed to the top plate.

Similarly, the shape of the first significant normal mode of the box kantele corresponds to the shape of the third normal mode of the modified kantele. The mode frequency has changed from 295 Hz to 230 Hz. The only difference between these two shapes is that in the modified kantele the free edge also has a vibration maximum. In addition, the Helmholtz resonance (found to be at 230 Hz) and the back plate participate in the vibration. For this mode, the vibration amplitude is higher for the modified kantele than for the box kantele.

Figure 5 shows the corresponding sound pressure levels for the mode shapes in Fig. 4; the first two normal modes of the top-plate kantele are shown in Fig. 5a, the first significant normal mode of the box kantele in Fig. 5b, and the first three normal modes of the modified kantele in Fig. 5c. The colour legend indicates the sound pressure level in dB units. The values above and below the legend are the maximum and minimum values for the sound pressure level, respectively. It can be seen that the radiation of the modi-

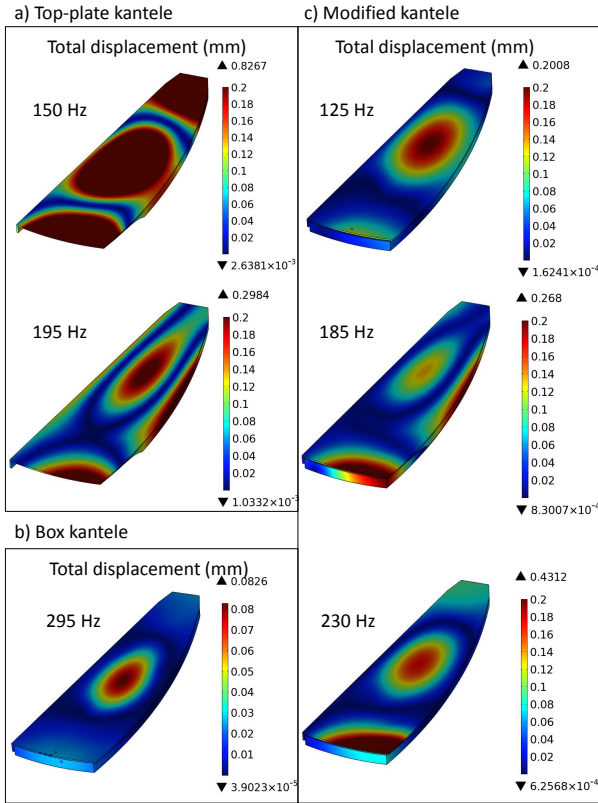


Figure 4. An example of how the modified kantele includes normal modes of both the top-plate kantele and the box kantele in terms of total displacement: a) the first two normal modes of the top-plate kantele, in b) the first normal mode of the box kantele, and c) the first three normal modes of the modified kantele. The mode shapes that are on the same row correspond to one another.

fied kantele at these three modes is more omnidirectional than for the traditional kanteles. The top-plate kantele has a higher maximum sound pressure level, but at some directions on the sides, the sound pressure level is small. This is a demonstration of the acoustic short-circuit. Namely, the air displaced by the top plate is in opposite phase on the top and at the back on the plate. These anti-phasic regions partially cancel each other, and thus reduce sound radiation.

5.2 Effect of the air gap size

The size of the air gap in the modified kantele can be changed. It is of interest to study how the size affects the vibrational and radiation properties of the kantele. Indeed, already a small air gap between the top and back plate changes the frequency response. Figure 6 shows the modelled frequency responses for the kantele with different air gaps: 1 mm, 3 mm, and 7 mm.

Based on Fig. 6, two different frequency ranges for the changes can be distinguished. First of all, below 300 Hz the mode density does not change, but the mode frequencies increase when the size of the air gap increases. The mode shapes do not change, but their vibration amplitudes do. Secondly, above 300 Hz, the mode density decreases when the size of the air gaps increases. In addition, the modes shapes do not appear in the same order in frequency.

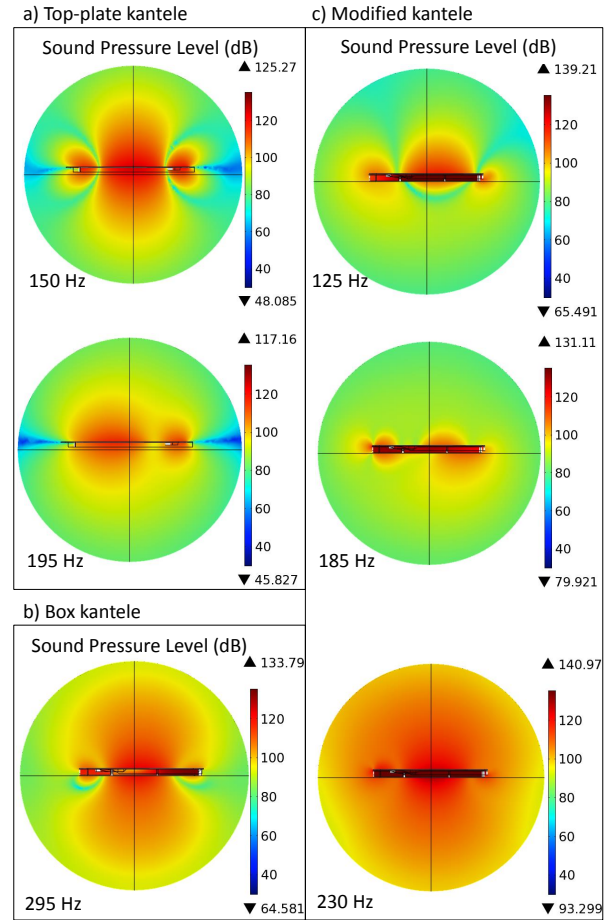


Figure 5. An example of how the modified kantele includes normal modes of both the top-plate kantele and the box kantele in terms of sound pressure levels: a) the first two normal modes of the top-plate kantele, in b) the first normal mode of the box kantele, and c) the first three normal modes of the modified kantele. The mode shapes that are on the same row correspond to one another. The kantele soundhole is facing the reader.

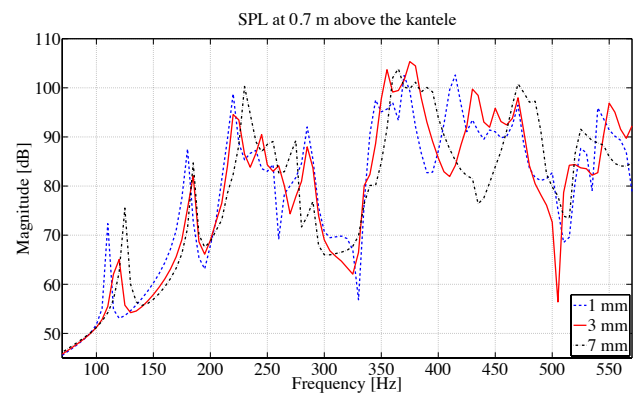


Figure 6. The frequency responses at 0.7 metres above the kantele with different air gaps: 1 mm (dashed blue line), 3 mm (solid red line), and 7 mm (dash-dotted black line).

As the size of the air gap increases, the free edge of the top plate is able to vibrate more freely. What is more, the vibration maxima move closer to the edges and the back plate is less coupled with the top plate. An example of a

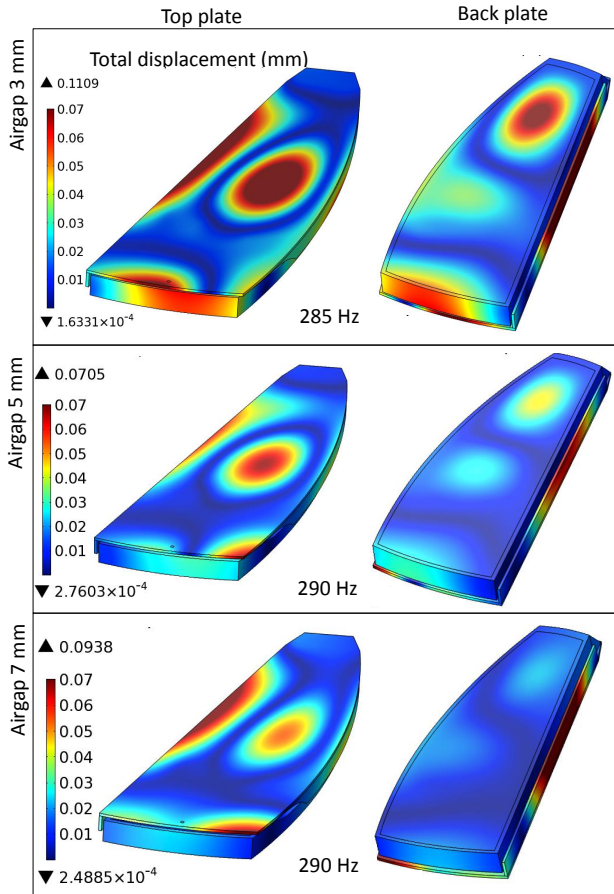


Figure 7. Vibration amplitudes of the sixth normal mode of the modified kantele with an air gap of 3 mm, 5 mm, and 7 mm. In the left-hand column the kantele is seen from the top and in the right-hand column from the bottom.

mode that undergoes such a change is the sixth normal mode of the modified kantele, occurring at 285 Hz. It experiences a drop of 15 dB in sound pressure level when the size of the air gaps increases for 1 mm to 7 mm. In addition, its frequency shifts from 285 Hz to 290 Hz. This mode shape is shown in Fig. 7 for different sizes of air gaps: 3 mm, 5 mm, and 7 mm. The case with the 1-mm air gap is almost identical to the 3-mm air gap. In the left column, the kantele is seen from above and in the right column from the bottom. The corresponding sound pressure level and pressure fields are shown in Fig. 8.

With the increasing air gap size, several phenomena occur. Firstly, the coupling of the top and the back plate decreases. In Fig. 8, the back plate vibrates with the shape (2,0) and its vibration amplitude decreases with the increase of the air gap. Furthermore, the widths of the vibration maxima decrease. At the free edge, the vibration maximum moves towards the corner that is the least fixed. In addition, the radiation patterns become more directive since the widths of vibration maxima decrease. Finally, different air modes couple with the body. At 285 Hz, the air vibrates with the mode shape (2,0) for 3-mm and 5-mm air gaps, and with the mode shape (1,0) for the 7-mm air gap.

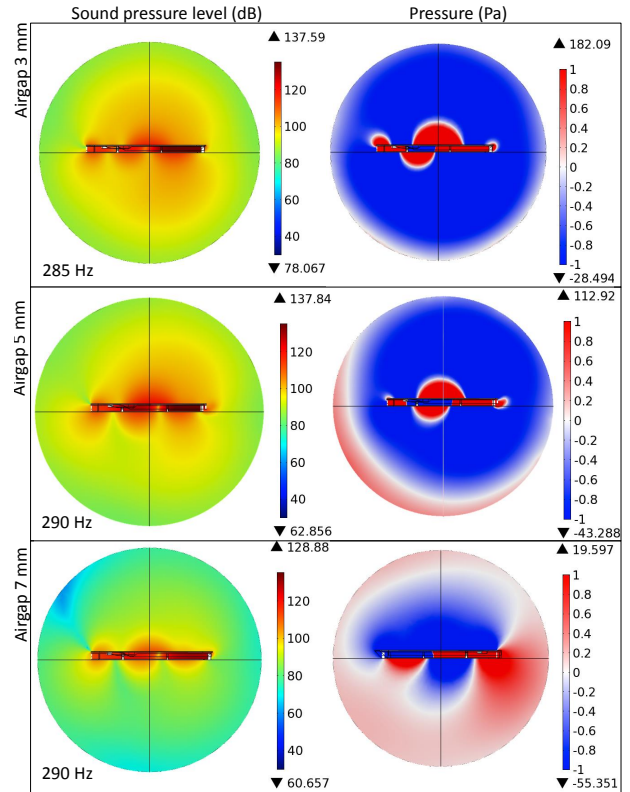


Figure 8. The sound pressure level and pressure of the sixth normal modes on the modified kantele with an air gap of 3 mm, 5 mm, and 7 mm. The sound pressure level is shown on the left-hand column and the pressure on the right-hand column.

5.3 Radiation efficiency

The modified kantele has a higher mode density than the traditional kanteles. The simulations show that the vibration maxima and sound pressure level decrease when the size of the air gap increases, at least for some modes. For the entire frequency range, the average radiation efficiency describes how effective the different structures actually are in transmitting input energy to radiation [26]. The averaged radiation efficiencies obtained via simulations for the modified kantele with different air gaps and the traditional kantele structures are shown in Tab. 2. The radiation efficiency is at its maximum with the 3-mm air gap. Beyond that, the radiation efficiency becomes comparable to that of the box kantele. The top-plate kantele has the lowest radiation efficiency of all the structures studied. It also exhibits more directional radiation patterns.

Air gap	box	1 mm	3 mm	5 mm	7 mm	top-plate
σ	0.0159	0.0242	0.0281	0.0128	0.0083	0.0008

Table 2. Radiation efficiency of the kantele with different air gaps, and the traditional box and top-plate kanteles.

6. CONCLUSIONS

In this paper, the body of a string instrument called the kantele has been simulated with the finite element method. Both the eigenfrequency analysis and forced vibration anal-

ysis with sinusoidal excitation were used to simulate the kantele with the frequency range of 0-570 Hz. Three kantele structures were included; two traditional structures, the top-plate kantele and the box kantele, were compared with a modified kantele which has an air gap separating the top and the back plate. In the modified kantele, the top plate acts as a free-edge vibrator. In addition, together the top and the back plate create an enclosed air mass that has its own vibrational modes.

The simulations confirm that combining both the freely vibrating top plate and the enclosed air leads to an increased density of normal modes and eigenmodes in the modified kantele in comparison to the traditional kanteles. A high mode density of the body in the tuning range of the strings is beneficial for the sound quality and sound pressure level of the instrument, as the string vibrations couple better with the body. This is one of the reasons for the improved loudness of the modified kantele.

In addition, the radiation efficiency of the modified kantele is improved compared to the traditional kanteles, provided that the air gap is smaller than 5 mm. Of the sizes modelled, an air gap of 3 mm had the highest radiation efficiency. Furthermore, the enclosed air allows air modes, in particular the Helmholtz resonance, to couple with the plate modes. Thus, the radiation patterns of the modified kantele are more omni-directional than those of the traditional kanteles. This phenomenon also contributes to the increased radiation efficiency of the modified kantele.

7. REFERENCES

- [1] M. Karjalainen, J. Backman, and J. Pölkki, "Analysis, modeling, and real-time sound synthesis of the kantele, a traditional Finnish string instrument," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, Minnesota, USA, 1993, pp. 229–232.
- [2] C. Erkut, M. Karjalainen, P. Huang, and V. Välimäki, "Acoustical analysis and model-based sound synthesis of the kantele," *Journal of the Acoustical Society of America*, vol. 112, no. 4, pp. 1681–1691, 2002.
- [3] J. Pölkki, C. Erkut, H. Penttinen, M. Karjalainen, and V. Välimäki, "New designs for the kantele with improved sound radiation," in *Proceedings of the Stockholm Music Acoustic Conference*, Sweden, 2003.
- [4] H. Penttinen, C. Erkut, J. Pölkki, V. Välimäki, and M. Karjalainen, "Design and analysis of a modified kantele with increased loudness," *Acta Acustica united with Acustica*, vol. 91, pp. 261–268, 2005.
- [5] S. Bilbao, *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*. John Wiley and Sons, 2009.
- [6] A. Chaigne, "Modeling and simulation of musical instruments," in *Proceedings of Forum Acusticum*, Aalborg, Denmark, 2011.
- [7] G. Derveaux, A. Chaigne, P. Joly, and E. Becache, "Time-domain simulation of a guitar: Model and method," *Journal of the Acoustical Society of America*, vol. 114, no. 6, pp. 3368–3383, 2003.
- [8] B. Richardson and G. Roberts, "The adjustment of mode frequencies in guitars: a study by means of holographic interferometry and finite element analysis," in *Proceedings of the Stockholm Music Acoustics Conference*, Sweden, 1983.
- [9] M. J. Elejabarrieta, A. Ezcurra, and C. Santamara, "Vibrational behaviour of the guitar soundboard analysed by the finite element method," *Acta Acustica united with Acustica*, vol. 87, pp. 128–136, 2001.
- [10] —, "Coupled modes of the resonance box of the guitar," *Journal of the Acoustical Society of America*, vol. 111, no. 5, pp. 2283–2292, 2002.
- [11] A. Ezcurra, M. Elejabarrieta, and C. Santamaria, "Fluid-structure coupling in the guitar box: numerical and experimental comparative study," *Applied Acoustics*, vol. 66, pp. 411–425, 2005.
- [12] J. A. Torres and R. Boullosa, "Radiation efficiency of a guitar top plate linked with edge or corner modes and intercell cancellation," *Journal of the Acoustical Society of America*, vol. 130, no. 1, pp. 546–556, 2011.
- [13] —, "Influence of the bridge on the vibrations of the top plate of a classical guitar," *Applied Acoustics*, vol. 70, pp. 1371–1377, 2009.
- [14] A. Runnemalm, "Sound and vibrations in structures with air closures," Ph.D. dissertation, Lule University of Technology, Sweden, 1999.
- [15] J. Bretos, C. Santamaria, and J. A. Moral, "Vibrational patterns and frequency responses of the free plates and box of a violin obtained by finite element analysis," *Journal of the Acoustical Society of America*, vol. 105, no. 3, pp. 1942–1950, 1999.
- [16] —, "Vibrational patterns of a violin-shaped air cavity obtained by finite element modeling," *Acta Acustica united with Acustica*, vol. 85, pp. 584–586, 1999.
- [17] N. Giordano, "Simple model of a piano soundboard," *Journal of the Acoustical Society of America*, vol. 102, no. 2, pp. 1159–1168, 1996.
- [18] J. Berthaut, M. Ichchou, and L. Jezequel, "Piano soundboard: structural behaviour, numerical and experimental study in the modal range," *Applied Acoustics*, vol. 63, pp. 1113–1136, 2003.
- [19] A. Mamou-Mani, J. Frelat, and C. Besnainou, "Numerical simulation of a piano soundboard under downbearing," *Journal of the Acoustical Society of America*, vol. 123, no. 4, pp. 2401–2406, 2008.
- [20] L. Ortiz-Berenguer, F. Casajus-Quiros, E. Blanco-Martin, and D. Ibanez-Cuenca, "Modeling of piano sounds using FEM simulation of soundboard vibration," in *Proceedings Forum Acusticum*, Paris, France, 2008.
- [21] F. Fahy and P. Gardonio, *Sound and Structural Vibration – Radiation, Transmission and Response*, 2nd ed. Academic Press, 2007.
- [22] F. Ihlenburg, *Finite Element Analysis of Acoustic Scattering*. Springer, 1998.
- [23] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Computational Physics*, vol. 114, pp. 185–200, 1994.
- [24] D. Kretschmann, "Chapter 5. Mechanical properties of wood," in *Wood Handbook*, ser. Technical Report. U.S. Department of Agriculture, Forest Service, Forest Laboratory, 2010, no. FPL-GTR-190.
- [25] A. Peekna and T. D. Rossing, "Acoustics of carved Baltic psalteries," *Acta Acustica united with Acustica*, vol. 91, pp. 269–276, 2005.
- [26] L. Cremer, M. Heckl, and E. Ungar, *Structure-Borne Sound*, 2nd ed. Springer Verlag, 1988.

THE BOLIVIAN CHARANGO – AN ACOUSTIC STUDY

Owen Woods
University of Cambridge
ukebert@googlegmail.com

Jim Woodhouse
University of Cambridge
jw12@cam.ac.uk

ABSTRACT

The charango, the subject of this paper, is a small stringed instrument from the Andes region of South America. The bridge admittance, mode shapes and impedance ratio were found for a range of charangos, a ukulele and a timple. A graphical method was developed in order to directly compare the results. It was found that the charangos tested were recognisably different from the other instruments. It was also found that the charangos tested were recognisably similar to one another. We can therefore conclude that it is meaningful to consider the charango as a distinct and homogenous instrument. The acoustic characteristics measured were explained using the social circumstances behind their construction, particularly the difference between *campesino* (peasant) and *mestizo* (urban) instruments. The most characteristic feature of the results, the ‘wasted’ Helmholtz resonance, was also explained in this way.

1. INTRODUCTION

After the Spanish invasion, the indigenous Andean peoples adopted and adapted certain Spanish customs, including their musical instruments. The focus of this report is one such instrument – the charango, a small stringed instrument found all over the Andes region. Formally an exclusively *campesino* (peasant) instrument, it has now been adopted by the urban-dwelling *mestizo* as a symbol of identity [1]. As such, they have altered the design of their instrument to suit their own musical preferences.

This paper presents the results of a series of investigations into the acoustics of the charango, to determine whether the instrument has an identifiable acoustic signature that is internally consistent and distinct from other similar-looking instruments.

2. INSTRUMENTS

The charango is a small guitar-like instrument, with a scale length of around 350mm. Traditional (*campesino*) examples have a back carved from a single piece of wood, or else are formed from a gourd or armadillo shell.

Copyright: © 2013 Owen Woods et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Mestizo instruments are generally made of wood throughout.

Figure 1 below shows the rough proportions of a typical instrument.

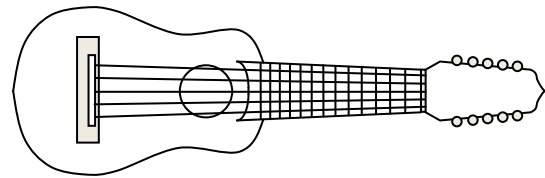


Figure 1. A typical charango shape.

Five charangos were tested during this investigation. Three were typical mass-produced *mestizo* instruments and two were of a more traditional construction. One of the latter had an armadillo back; all the other instruments were wood.

A ukulele, from Hawaii, and a timple, from the Canary Islands, were also tested. These were intended as a comparison, since they were made in similar circumstances and had similar physical characteristics. Results will be compared with previously collected data from a high quality classical guitar, as the guitar has been extensively studied in the past.

3. METHODS AND THEORY

3.1 Bridge Admittances

Plucked stringed instruments generate sound through physical displacement of air by the front plate and body and through resonance of the air inside the cavity. The input to the system is the force from the strings, when caused to vibrate by the player. This vibration is passed to the rest of the instrument through the bridge, so it can be argued that the single most useful acoustical measurement on a stringed instrument is the driving-point frequency response at the bridge [2]. Specifically, we use here the bridge admittance (or mobility), the ratio between the response velocity and the force of the excitation. When plotted on a frequency scale, peaks may be identified which correspond to modes of vibration of the system.

In order to measure the bridge admittance, an instrumented tapping hammer was used to excite the bridge close to the lowest string in the direction normal to the plane of the top plate. A Polytec laser vibrometer was

used to measure the velocity of the bridge, as close as possible to the excitation point.

In order to allow useful comparisons between instruments of different sizes, tuning and stringing, the results will be presented in normalised form on both axes. The frequency axis is normalised by the nominal tuning of the lowest note of the instrument, so that results can be shown in “Semitones above and below the lowest note”. This allows the placement of resonant frequencies on instruments with different tunings to be directly compared in a form relevant to the intended use.

The magnitude axis is normalised using the wave impedance of the lowest string. In order to measure this from strings *in situ* on an instrument, the static transverse deflection of the string when loaded at the mid-point by a known mass was measured. The wave impedance, defined as the square root of the product of the tension and the mass per unit length, is given by

$$Z_0 = \frac{Mg}{8df} \quad (1)$$

where Z_0 is the wave impedance, M is the mass of the load, g is the acceleration due to gravity, d is the measured deflection and f is the frequency of the string in Hz. Since impedance is the inverse of admittance, the measured bridge admittance is multiplied by Z_0 to give the non-dimensional impedance ratio.

3.2 Mode Shapes

Each resonance of the front corresponds to a different mode shape of the instrument. Each mode has areas of high movement separated by lines of no movement, known as nodal lines. Nodal lines may be physical constraints, such as the edge of a plate, or else lines arising from the structure and geometry of the system. The modes observed over the course of the project were associated with the direct physical movement of the front, movement of the whole body, or resonance of the air in the cavity.

The laser vibrometer was used to measure the velocity of different points on the front and neck when excited by an impulse at the bridge as before. These results were then collated using the Polytec software to reveal the physical movement of the instrument at each major peak frequency. This gives “operating deflection shapes”, which are equivalent to mode shapes at low modal overlap. Where there is a significant degree of modal overlap, the mode shapes can be deduced by post-processing the measured frequency responses [3].

4. RESULTS - DISTINCTIVENESS

Initially, one charango (labelled Charango M) was tested, along with a timple and a ukulele. As a comparison, data from a high quality classical guitar was also used. A graphical method was developed to compare the acoustical behaviour of the different instruments. The bridge

admittances were normalised by frequency of the lowest note and by the string impedance as above. The mode shapes were then analysed and individual resonances on the bridge admittance plots correlated with different modes.

The bridge admittances for each instrument were then plotted on a single graph, offset on the magnitude axis for clarity. Lines were drawn connecting similar modes. If the instruments were truly similar, then the frequency of each mode on different instruments would be in roughly the same place relative to the lowest note, meaning that lines would be vertical. If the instruments were dissimilar then this would not be the case, with lines at an angle and even crossing.

Figure 2 shows this plot for the four instruments, with the locations of each mode represented by dots. The grey line denotes the Helmholtz resonance (resonance of the air inside the cavity), orange lines denote modes involving the whole body including the neck, and black lines denote modes mainly confined to the front plate. Dotted lines represent modes that are not present in a particular instrument.

It can clearly be seen that none of the instruments show a significant degree of similarity to each other. Modes are in different orders, at different frequencies and are even missing altogether when comparisons between instruments are made. This can be seen at a glance by the tangle of lines.

We can therefore conclude that charango M is distinctly different from the timple, ukulele and guitar in terms of modal frequencies vs. tuning of strings. We can furthermore say that charango M is distinct from classical guitars in general, since guitars have been shown to exhibit a high degree of coherence: see for example [2].

5. RESULTS - COHERENCE

Having shown that charango M is demonstrably distinct from other, generically similar stringed instruments, the question remains whether it is representative of all charangos and whether charangos form a coherent set.

Five charangos were tested, labelled as M, N, Q, O and A. The first three were *mestizo* instruments. All bar charango M were kindly made available for this study from the collection of Dr Henry Stobart of Royal Holloway College.

The same graphical method as in Section 4 was used to compare these five instruments and it is presented in Figure 3. It can easily be seen that the instruments exhibit a far higher degree of coherence than the previous examples, with straighter lines between modes and few modes not shared.

The coherence is sufficient to be able to say that the charango is a distinct and recognizable instrument, however it is not as coherent as the guitar. This is probably due to the guitar’s many years of methodical development in a performance-orientated framework. The charango has not had this background, as for most of its history there has been little contact between builders.

A possible exception to the conclusion presented above is charango A, the *campesino* armadillo-backed charango. This instrument exhibited characteristics that were not shared by the other instruments.

The normalisation process described above is difficult to apply to charango A, given that this instrument does not have a specific tuning for which it is designed. Most stringed instruments (arguably *mestizo* charangos included) have a tuning that is bound up with how the instrument is played and constructed. Traditional charangos however have a number of different tunings, varying with the area, time of year and current fashion trends amongst performers [4].

Charango A is a smaller instrument than the other instruments tested, with a smaller scale length. Crucially, it also has a smaller cavity volume and a smaller front area, meaning that the first two modes occur at higher frequencies.

This instrument is used primarily for accompanying song and for courtship and as such, the sound is required to be light, bright and loud [4]. This sound quality is reflected in the results. The high frequency modes and the metal strings (which have a high wave impedance) will make the instrument loud and piercing. This makes the instrument fit for purpose – accompanying song outside without amplification.

Mestizo charangos evolved from traditional instruments, of which charango A is an example. A family resemblance can be seen, but the *mestizo* charangos have a more complex sound (and hence a more complex bridge admittance), more suited to playing indoors and being accompanied by other instruments. This reflects the changing use of the instrument by the *mestizo* as opposed to the *campesinos*.

6. ETHNOMUSICOLOGY

The most characteristic aspect of the charango analysis is that the frequency of the Helmholtz resonance is below that of the lowest note of the instrument. This means that its excitation will be limited when the instrument is played. This property can be explained using ethnomusicology. The *mestizo* required an instrument that was quieter and had greater depth than the *campesino* original [1]. They wanted an instrument that facilitated fast finger movement and complex strumming patterns, as well as one which would blend well with traditional guitars, guitar making and playing techniques. This led them to increase the scale length of the instrument and to use of nylon strings.

Strings made of different materials have different damping ratios, stiffnesses and tensions. All of these affect the sound that a string will make through the soundboard. Compared to steel strings, nylon strings have higher damping and lower tension, which means that the higher harmonics are excited less. The use of metal strings in the *campesino* instruments gave them a bright, trebly tone, with little bass response. However, the use of nylon strings on the *mestizo* instruments immediately meant that the bass started to play more of a role, thus

preventing the instruments from sounding like charangos. This is a plausible explanation for why the cavity is oversized: with the Helmholtz resonance being excited by the strings, the bass may have dominated too much. From the author's own empirical observations, stringing charango M with strings below the Helmholtz resonance resulted in the loss of the characteristically focused sound of the original instrument and it became uncomfortably bass dominant.

In conclusion then, the *mestizo* charango is an effort to retain some of the sound of the traditional instrument, whilst attempting to make it more acceptable to the musical elite. This conclusion matches well with Turino's interpretation of the motives behind the popularity of the charango amongst *mestizo* populations [1].

7. CONCLUSIONS

This study has shown that the set of charangos tested is both coherent and distinct. This means that the charangos tested look similar – such that they are recognisable as charangos from their acoustic response – and that as a group they look different from the other instruments tested. This study has therefore shown that it is meaningful to talk about charangos as a unique, homogenous set of instruments, even when taking into account their chequered history. Indeed, it may also be concluded that the armadillo charango (charango A) is demonstrably an antecedent to the *mestizo* charangos, as it has similar characteristics but is designed for a different time and place. The physical characteristics of the charangos tested have been examined and it has been concluded that the history of the *mestizo* charangos may be seen in their construction. These are derivative instruments, made for a purpose alien to the originals and this can clearly be seen in the way that they are built. This is typified by the soundbox being too large to excite the Helmholtz resonance at the lowest note of the instrument.

Acknowledgments

Many thanks to Dr Henry Stobart of Royal Holloway in the University of London, who through his generosity in lending instruments allowed this project to happen.

8. REFERENCES

- [1] T. Turino, "The Urban-Mestizo Charango Tradition in Southern Peru: A Statement of Shifting Identity", *Ethnomusicology*, vol. 28, pp.253-270, 1984.
- [2] J. Woodhouse, "Plucked Guitar Transients: Comparison of Measurements and Synthesis", *Acta Acustica united with Acustica*, vol. 90, pp.945-965, 2004.
- [3] D.J. Ewins, *Modal Testing: Theory, Practice and Application*. Research Studies Press Ltd, Taunton, 2000
- [4] H. Stobart, *Music and the Poetics of Production in the Bolivian Andes*. MPG Books Ltd, Bodmin, 2006.

9. APPENDIX

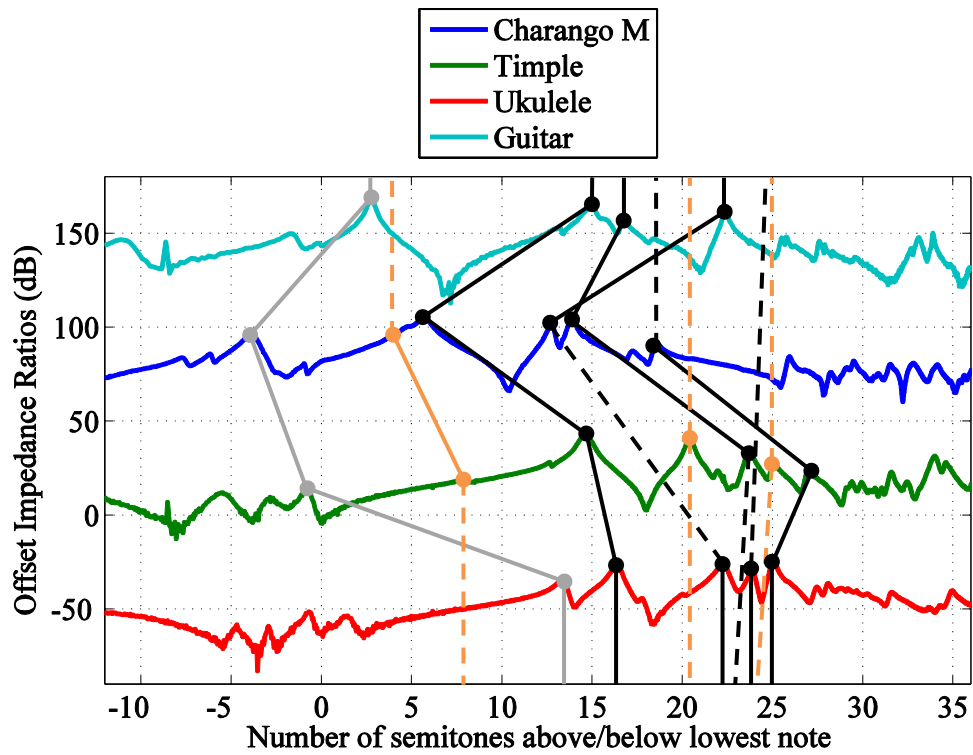


Figure 2. Plot of similar modes of four dissimilar instruments.

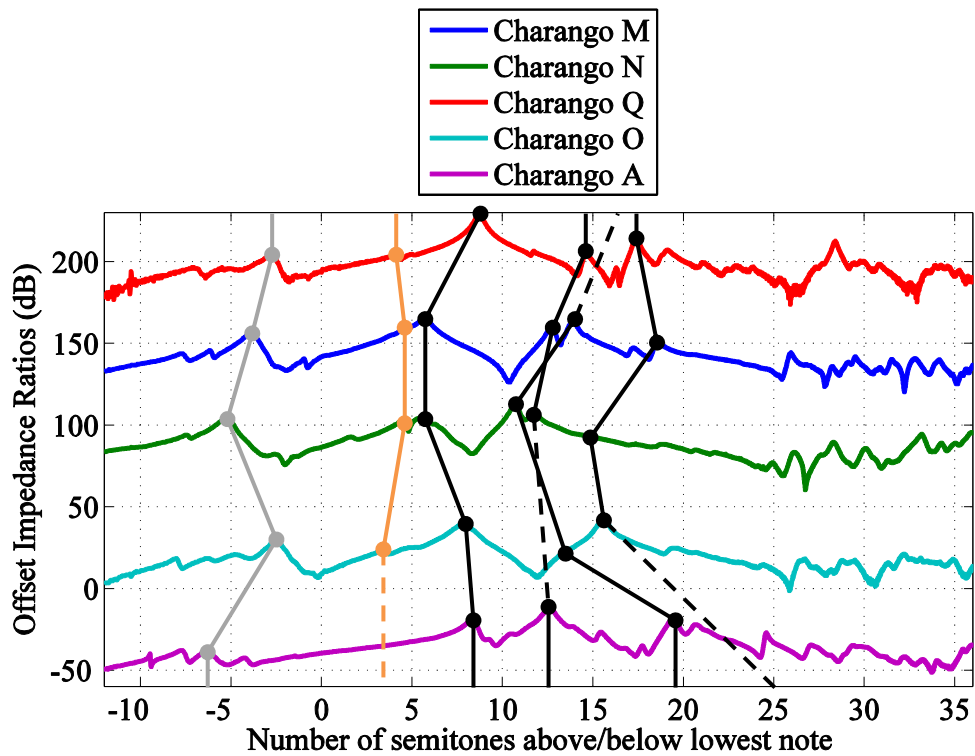


Figure 3. Plot of similar modes of five Charangos.

Singing voice

SUPRANORMAL VOICES IN SINGING

Ken-Ichi Sakakibara

Health Sciences Univ. of Hokkaido, Japan
kis@hoku-iryo-u.ac.jp

ABSTRACT

In this paper, the author surveys various supranormal voices and discusses how to synthesize such supranormal voices based on physiological observations. In particular, the author describes physiological aspects about abrupt register changing, vocal-ventricular phonation, vocal-arytenoid phonation, which are widely found in ethno or primitive music and recently also found in pop music. To synthesize such supranormal voices, the author proposes two different methods: physical modeling and signal modelling.

ACOUSTIC CHARACTERISTICS OF VIBRATO IN DIFFERENT SINGING STYLES

Noam Amir

Dept. of Communication Disorders,
Tel Aviv University
noama@post.tau.ac.il

Irit Ronen

Dept. of Communication Disorders,
Tel Aviv University
iritune@gmail.com

Ofer Amir

Dept. of Communication Disorders,
Tel Aviv University
oferamir@post.tau.ac.il

ABSTRACT

Vibrato is a commonly found characteristic of the singing voice, adding a sense of flexibility and richness. Acoustically, it is a periodic variation of frequency and amplitude, characterized by several parameters: *rate* – The numbers of cycles per second; *extent* – The degree of modulation of F0; *amplitude* – The degree of modulation in intensity. The purpose of the present study was to compare characteristics of vibrato across different singing styles. 429 samples were taken from commercial CDs, sung by thirty singers in three different singing styles: opera, jazz and pop. Vibrato *rate* tended to be slightly higher in the opera style, with borderline significance. Vibrato *extent* was significantly greater in the opera style than in the other two. These results imply that vibrato varies with musical style and support the idea that vibrato is a learned vocal skill rather than a natural vocal behavior.

1. INTRODUCTION

Vibrato is a widespread phenomenon found in playing of various musical instruments and in the singing voice. Several definitions of vibrato have been proposed in the literature. Titze, for example defined it as a modulation – a systematic change in a periodic variable [1]. Sundberg defined vibrato as a regular change in frequency, timbre and intensity [2]. Nair defined it as a periodic variation in fundamental frequency, found in the relaxed vocalization without unnecessary tension [3].

Vibrato is not present in all cultures, but where it is found there have been many different descriptions of how it is produced. In western cultures, where it is widespread, it is regarded as an important feature of classical singing styles. Howard, for example, regards vibrato as instrumental in expressing a range of emotions [4].

Acoustically, vibrato is usually characterized by several parameters:

Rate – the number of vibrato periods per second. According to Howard, for example, this varies between 5 and 9 Hz.

Extent – in technical terms this is the "modulation depth", i.e. the range of F0 values in a single period of vibrato. Values of extent quoted in the literature are varied. Horii, for example reports values of 50 to 150 cents [5].

Amplitude – Similar to extent, but referring to the range of intensities rather than frequencies.

1.1 Vibrato in different singing styles

Very few previous studies have examined the differences in vibrato acoustics between singing styles. In one such study, Rothman et al. compared singing of opera singers and cantors in the past and present, in order to ascertain whether singers from different cultures and periods employ similar aesthetic considerations [6]. Their study showed significant differences between periods but not between different singing styles taken from the same period. Historical singers were found to have higher rate and lower extent and amplitude than contemporary singers.

Bezerra et al. compared vibrato characteristics of 10 male opera singers to 10 male singers in a traditional Brazilian style known as Sertanejo [7]. Singers in each style sang the same piece, expressly for the purpose of this research. A significant difference in rate was found between the two groups, but not in extent.

A previous study by the present authors compared vibrato between singers with different degrees of proficiency rather than different singing style [8], using a novel method to analyze the F0 contour, based on the Fourier Transform.

Finally, several additional studies [9,10] compared different singing styles from the point of view of acoustic characteristics other than vibrato.

1.2 Paradigms in vibrato research

Previous studies of vibrato appear to be divided quite evenly between studies based on commercial recordings and studies based on laboratory recordings. On the one hand, comparing the performance of the same musical pieces by different singers could bias the range of available notes [11]. On the other hand, recording of isolated notes out of context could inadvertently affect singing characteristics [12].

1.3 Rationale and objectives

In view of previous studies on this subject, it appears that there is very little data on differences in vibrato between singing styles. This is an interesting question in itself, in addition to which it can shed light on the longstanding controversy whether vibrato is a learned ability or naturally occurring. To eliminate any bias that may be created by singing in laboratory conditions, we found it most interesting to study these differences on

singing taken from commercial recordings. The objectives of this study were therefore:

1. To define several quantitative measures of vibrato.
2. To find a method to extract these measures from commercial recordings, in the presence of accompanying music
3. To compare these measures across three singing styles: Opera, Jazz and Pop.

2. METHODS

2.1 Participants

Vibrato samples were taken from 30 female singers, taken from commercially released albums, in three different musical styles: 10 opera singers, 10 jazz singers and 10 pop singers. To limit influencing factors of age and musical fashions, ages were limited between 23 and 50, and all excerpts were taken from disks recorded between 1990 and 2010. Opera singing was taken from recordings of Arias: 6 in Italian, 3 in French and one in German. All jazz singers performed in English. Four of the pop singers sang in Hebrew and 6 in English.

2.2 Musical Excerpts

To avoid bias, the performances of each singer were scanned manually for sustained singing of the vowel /a/ lasting more than 800 ms. The first 20 such excerpts were taken from each singer. In some cases fewer excerpts were found, with the smallest number being seven. Each excerpt was judged perceptually by an experienced singer for presence or absence of vibrato.

A certain difference between styles was observed already at this stage. While a sufficient number of excerpts was easily obtained from opera singers, 10 pop singers were discarded for lack of sufficient occurrences of sustained notes lasting more than 800 ms, as well as 5 jazz singers.

2.3 Extracting F0

A major challenge in analyzing the recordings was reliable extraction of F0 in the presence of background music. The method employed here was evaluating the target pitch of each sung note, then narrow-band bandpass filtering the excerpt around either the first or second harmonic. This eliminated background music to a sufficient degree. The filtered signals were then analyzed with Praat software [13]. In cases where the second harmonic was analyzed, the results of the Praat analysis were divided by two. A representative example of pitch analysis before and after filtering is presented in Figure 1, on an excerpt taken from a pop singer.

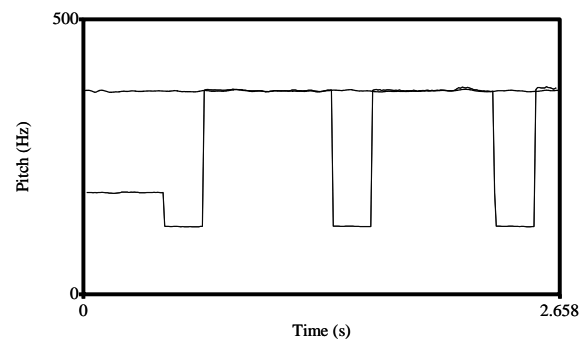


Figure 1. Pitch contour of a sustained note, before filtering (dashed line) and after filtering (solid line). Filtering greatly improves the resultant contour, to the point it can be analyzed for vibrato.

The pitch contour obtained from the raw signal is erratic, whereas after filtering the pitch contour is almost a straight line on this scale. The details of the filtered pitch contour are more obvious when zooming in on the y-axis, as in Figure 2. Slight variations in pitch are obvious, and while they are somewhat irregular, they are quite small. This recording, in fact, was judged not to contain vibrato at all.

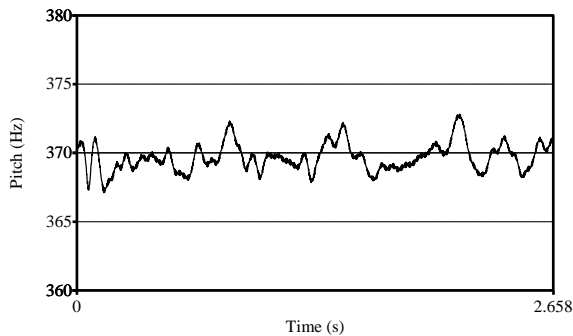


Figure 2. Pitch contour of the sustained note in Figure 1, after filtering. Small variations in pitch are now evident.

2.4 Analyzing the F0 contours

The pitch contours for every filtered excerpt were calculated with Praat and inspected visually for any artifacts, and then saved to files. A custom-written Matlab GUI was used to extract a set of relevant vibrato related features from these files, as detailed below, in two stages.

In the first stage, all excerpts were scanned visually by an experienced research assistant, to determine the presence of delayed vibrato onset. In such cases, the GUI allowed the user to mark the region containing vibrato manually, using the mouse. In the second stage, the software looped over the entire database of excerpts, performing two operations: 1) subtracting a parabolic regression from the contour, to remove slow drifts in pitch; 2) extracting the following features:

- **Rate:** this was established by performing a Fourier transform on the pitch contour (which we term the "pitch spectrum") and finding the frequency of the most prominent peak.
- **Extent:** to reduce bias by spurious pitch values, extent was estimated by interpercentile range in semitones.

A typical screenshot of the Matlab GUI is shown in Figure 3. As a side note, we remark that intensity related features could not be extracted, due to the background music. Whereas the effect of this music on F0 could be removed by the filtering procedure mentioned above, its effect on intensity could not.

2.5 A note concerning intensity modulation

The methodology described above demonstrates how the F0 contour can be extracted from the recordings reliably. While the filtering process enhances extraction of F0, it

clearly removes a large amount of intensity information. On the other hand, the intensity of the raw signal is strongly affected by the presence of the accompanying instruments. Therefore, though intensity modulation is interesting to vibrato research, this information could not be recovered within this research framework.

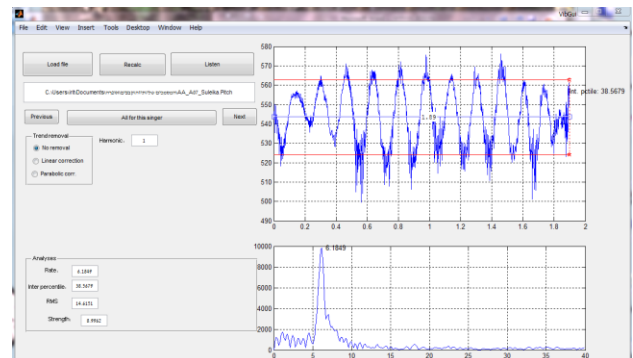


Figure 3. A screenshot of the Matlab GUI used to analyze the pitch contours. Top pane is the pitch contour, bottom pane is its Fourier Transform, showing a prominent peak near 6 Hz.

3. RESULTS

3.1 Distribution of excerpts

Overall, 429 excerpts were extracted, distributed in an approximately even manner over the different singing styles. Vibrato, however, was not present equally in all singing styles. Table 1 presents a summary of this information, along with a count of the number of excerpts with delayed-onset vibrato.

	Total	With vibrato	Without vibrato	Delayed vibrato
Opera	160	160 (100%)	0	12 (8%)
Jazz	130	97 (74%)	33	51 (53%)
Pop	139	89 (64%)	50	5 (6%)

Table 1. Distribution of excerpts with vibrato, without vibrato and with delayed vibrato.

Some differences between the singing styles already become evident. Opera singers employed vibrato throughout, with less than 10% using delayed vibrato. Only about three quarters of the jazz excerpts contained vibrato, with about half of them being delayed-onset. Finally, little more than half of the pop excerpts contained vibrato, with a very small percentage of them being delayed-onset.

3.2 Rate

A box and whisker plot for vibrato rate is presented in Figure 4. Mean rates for the three singing styles were in a relatively narrow range of approximately 0.7 Hz. The

figure shows a trend of opera having the highest mean rate, pop having the lowest, and jazz in between. A one-way ANOVA showed these differences to be borderline significant ($F(2,25)=2.715, p=0.085$).

3.3 Extent

A box and whisker plot for vibrato extent, in semitones is presented in Figure 5. Similarly to rate, Figure 5 shows opera having the highest mean extent, pop having the lowest, and jazz in between. A one-way ANOVA showed these differences to be highly significant ($F(2,25)=15.91, p<0.001$). Posthoc analysis revealed significant differences between opera and jazz and between opera and pop, but no significant difference between jazz and pop

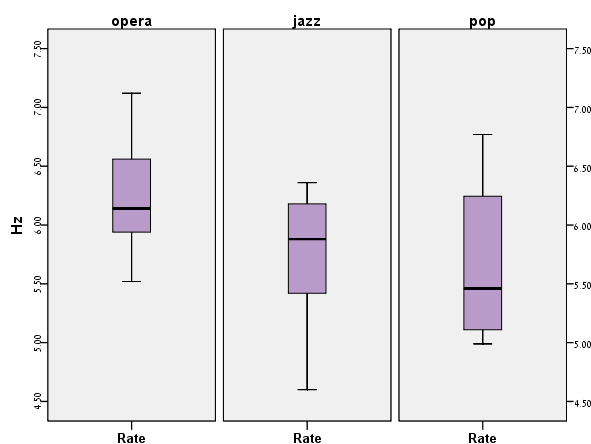


Figure 4. A box and whisker plot of vibrato rate in each singing style.

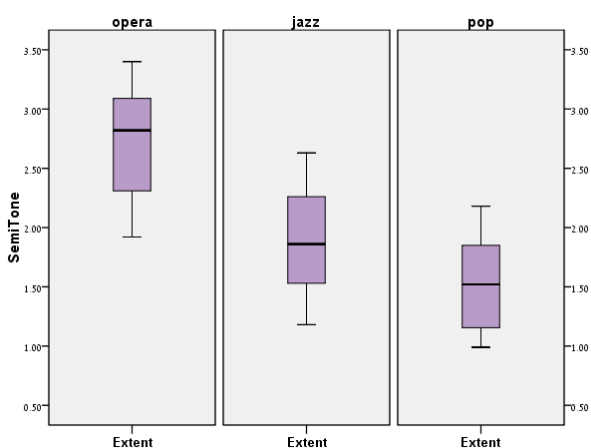


Figure 5. A box and whisker plot of vibrato extent in each singing style.

4. DISCUSSION AND CONCLUSIONS

The results above lead to several interesting conclusions.

Presence of vibrato: First, it is clear that the opera singers sang sustained notes exclusively with vibrato, and almost exclusively from the beginning of the note. As found for other parameters, pop singers used it least, and jazz singers were in between. Interestingly, both pop singers and opera singers had few instances of delayed onset vibrato, while jazz singers employed delayed onset in about half the notes that did have vibrato.

Vibrato rate: variation in mean vibrato rate between singers was quite small, and differences were only borderline significant. This implies some kind of global constraint on vibrato rate, though it is hard to say if this is a physiological constraint or an aesthetic one. However it is probably safe to say that rate is not a distinguishing factor between singing styles.

Vibrato extent: vibrato extent of the opera singers was found to be significantly different from the extent of both jazz and pop singers, though no significant difference was found between the last two. This implies that extent, as opposed to vibrato rate, is more easily controlled, and appears to be quite a strong characteristic in distinguishing operatic singing from other styles.

In summary, the analysis of commercial recordings can be used as a solid basis for studying the characteristics of vibrato in different singing styles. The differences found here indicate that there is a greater difference between opera singing and other styles, though it would be worthwhile basing this on a larger sample in future studies.

5. REFERENCES

- [1] I. R. Titze, *Principles of Voice Production*. Prentice – Hall Inc. Chapter 8: Control of fundamental frequency, pp.191-197. Chapter 11: Fluctuations and perturbations in vocal output, pp. 289-292, 1994
- [2] J. Sundberg, M. Hirano, P. H. Dejonckere, *Vibrato*. San Diego publishing group, 1995
- [3] G. Nair, *Voice tradition & technology: A state-of-the-art studio*, 1999
- [4] E. Howard, E. Sing! *The vocal power method*. Alfred publishing co.,Inc. chapter 6. 2006
- [5] Y. Horii, Y. Acoustic analysis of vocal vibrato: A theoretical interpretation of data. *Journal of Voice*, 3, 36-43, 1989
- [6] H. B. Rothman, J. A. Diaz, K. E. Vincent, Comparing historical and contemporary opera singers with historical and contemporary Jewish cantors. *Journal of Voice*, 14, 205-214. 2000

- [7] A. Bezerra, S. Cukier-Blaj, A. Duprat, A., Camargo, L. Granato. The Characterization of the Vibrato in Lyric and Sertanejo Singing Styles: Acoustic and Perceptual Auditory Aspects . *Journal of Voice*, 23, 666-670. 2009
- [8] N. Amir, O. Michaeli, O. Amir, Acoustic and perceptual assessment of vibrato quality of singing students. *Biomedical signal processing and control* 1, 144-150. 2006
- [9] M. Thalen, J. Sundberg, Describing different styles of singing: a comparison of female singer's voice source in "classical", "pop", "jazz" & "blues". *Logopedics, Phoniatrics, Vocology*, 26, 82-93. 2001
- [10] C. J. Butte, Y. Zhang, H. Song, J. J. Jiang, Perturbation and nonlinear dynamic analysis of different styles. *Journal of Voice*, 23, 647-652. 2009
- [11] I. R. Titze, B. Story, M. Smith, R. Long, A reflex resonance model of vocal vibrato. *Journal of Acoustic society academy*, 111, 2272-2282. 2002
- [12] D. M. Hicks, E. Teas,. An electroglottographic study of vocal vibrato. *Journal of voice*, 1, 142-147. 1987
- [13] P. Boersma, D. Weenink, Praat: doing phonetics by computer [Computer program]. Version 5.3.43, <http://www.praat.org/>

	Diane Schuur	Midnight
	Natalie Cole	Unforgettable: With love
	Stacey Kent	Close your eyes
	Karrin Allyson	Wild for you
	Susie Arioli	Night Lights
Pop	Shiri Maimon	A moment before
	Ninette Tayeb	Barefoot
	Maya Buskila	Breaking the silence
	Miri Messika	Miri Messika
	Madonna	Something to remember
	Mariah Carey	Day dream
	Kylie Minogue	Light years
	Jennifer Lopez	On the 6
	Tina Arena	Don't ask
	Celine Dion	These are special

APPENDIX 1 – LIST OF EXCERPT SOURCES

Style	Singer	Album
Opera	Cecilia Bartoli	An Italian Songbook
	Angela Gheorghiu	Verdi Heroines
	Sumi Jo	Virtuoso Arias
	Arleen Auger	Schubert Lieder
	Renee Fleming	Night Songs
	Olga Borodina	Arias
	Sylvia McNair	Reveries - Melodies
	Vesselina Kasarova	Mozart arias
	Anna Netrebko	Netrebko & Villazon
	Veronique Gens	Melodies Francaises
Jazz	Rachelle Ferrell	First Instrument
	Dianne Reeves	That day
	Jeri Brown	Fresh start
	Madeleine Peyroux	1.Careless love
		2. Half the perfect world
	Malia	Young bones

EVALUATION OF PITCH DETECTION ALGORITHMS: CASE OF MONOPHONIC VOCAL PERFORMANCE

Robertas Budrys

Faculty of the Humanities,
Kaunas University of Technology, Lithuania
budrys@super.lt

Rytis Ambrazevičius

Faculty of the Humanities,
Kaunas University of Technology, Lithuania
rytisa@delfi.lt

ABSTRACT

Evaluation of musical scale in monophonic vocal performance is a complicated task, and some issues causing inaccurate results might occur. For instance, one issue is the selection of algorithm which detects the fundamental frequency of performance as precisely as possible. If precision of the algorithm is worse than jnd or if the algorithm detects unsteady pitch track inaccurately, its data is often not suitable for further analysis. In this paper, three different pitch detection algorithms are discussed: auto-correlation, YIN and SWIPE¹. The algorithms were applied to generate pitch tracks of synthesized sine tones and examples of natural voice, and the results were compared.

1. INTRODUCTION

Acoustical measurements of musical pitch are of great importance for various fields including ethnomusicological studies among others. One of the purposes of acoustical measurements is to identify the intervals in the traditional musical scales. This allows not merely an estimation of their deviations from twelve-tone equal temperament, but also to reveal their essential properties, to classify them, and to seek for the groundings of the ascertained phenomena in music perception.

For instance, some recent studies of traditional Lithuanian vocal music show that attempts to transcribe and classify these songs in terms of Western musical system ("Ancient Greek modes", chromaticisms, modulations, etc.) could result from mere misperception of different modal thinking ("the collision of two emic scale systems") [1: 1817]. The main evidences of this kind of studies are based on acoustical measurements.

1.1 Issues of Musical Scale Evaluation

It is well known that quite free and elaborated intonation is characteristic of a vocal performance. Lithuanian folk song performance is not an exception. In terms of fundamental frequency, individual tones are not very stable, some glides and embellishments between them are typical. Therefore acoustical measurements and scale evaluations of this kind of music becomes sufficiently complex.

Copyright: © 2013 Robertas Budrys, Rytis Ambrazevičius. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

While estimating the scale of a particular vocal solo performance, some issues causing inaccurate results might occur. Three main categories of the issues can be distinguished: (1) extraction of fundamental frequency (and other sound properties) from the recording; (2) chunking of continuous pitch track¹ into separate tones and the estimation of their pitches (and other properties); (3) choice of the method for evaluation of the musical scale based on the collected data.

The first issue is related to selection of software implementing certain pitch detection algorithm (PDA) which extracts the track of fundamental frequency from the audio recording. This issue will be discussed in the paper. The other two issues are outside the scope of the current research².

1.2 Accuracy of Pitch Detection

If one presumes that any PDA detects fundamental frequency with some inaccuracy, the question arises about how precise the output of algorithm should be and what is the tolerable limit of error. At least, two problems related to this question should be mentioned.

The first problem is pitch jnd (just noticeable difference), the smallest change in pitch perceptible by a listener. It is shown that pitch jnd can be as low as a few cents (e.g., [4, 5: 863, 6: 185]). However, these results are only valid under ideal listening conditions and only for successive tones. In the case of real performances, under influence of various factors (for instance, if tones of different ranges of intensities, durations, timbres, or separated by a disturbing information, are considered), pitch jnd can be much higher [7: 391–396, 8: 127–128]. Because of jnd, a performer may sing slightly "out of tune" (here we do not mean larger and perceivable deviations) and adequate precision of PDA should be linked to this fact. Consider the error of ± 10 cents (i.e., the difference between PDA readings and actual pitch) as the satisfactory limit of accuracy, although lower values would be highly preferable.

The second problem comes from fast changes of fundamental frequency. Fast glides, vibrato, U-shaped and other fast modulations of fundamental frequency are perceived as integral pitches. Some studies show that the perceived pitches correspond to the averaged f0s [9, 10]

¹ To be precise, here and hereafter the notion of pitch track actually means logf0 track, as traditionally accepted by the authors of PDAs.

² Concerning the software-aided semi-manual method for pitch and scale estimation, see [2: 66–67] and [3: 59].

while others additionally include the effect of memory decay [11, 12]. In addition, realizations of brief unsteady tones can significantly differ from intentions of a performer, therefore such tones should be considered with caution.

2. METHOD

2.1 Algorithms

Three PDAs were chosen for the evaluation: auto-correlation (AC), YIN and SWIPE³. These algorithms are quite popular, widely used, and can be downloaded freely from the web. The authors of the algorithms claim that their algorithm is more accurate compared to others [13, 14, 15]. The gross error rate technique is applied to evaluate the performance of the competing algorithms, where gross error is stated if the pitch estimated by PDA deviates from the reference pitch by more than 20% [14, 15], i.e., about 4 semitones below or 3 semitones above the actual pitch. Usually PDAs make octave errors, i.e., fundamental frequency is detected at half of cycle or at double cycle. Our studies deal with Lithuanian monophonic folk songs which usually do not exceed the interval of octave, therefore initial aural evaluation of approximate pitch range in the performance and the possibility to adjust the analysis range of PDAs are employed. This prevents the algorithms making most of octave errors.

The AC algorithm is implemented in PRAAT software³ as the main PDA (other algorithms are also possible). YIN⁴ and SWIPE⁵ are toolboxes for MatLab software. We will discuss these algorithms not as mathematicians, but as end-users who need reliable results from the software.

2.2 Material

Twenty two sine wave examples of constant and changing f0s were synthesized to test the performance of each algorithm: steady tones (4 examples), tones with (“normal”) vibrato (8 examples), tones with “slow vibrato” (8 examples), and glides (2 examples). The steady and vibrato tones have constant or mean f0s of 110, 220, 440, and 880 Hz and durations of 5 s. Frequency of the vibrato tones is 6.5 Hz (the typical value of natural vocal vibrato [e.g., 16: 139, 17, 18: 142]) and the amplitudes are ± 48 and ± 96 cents. Frequency of the “slow vibrato” tones was set to 1 Hz and amplitudes of ± 96 and ± 204 cents were chosen⁶. Starting and ending points of the glides are respectively 110 and 880 Hz, and durations are 2 and 5 s. (See Table 1 for details.)

In addition, four examples of natural voice were prepared from the recordings of monophonic Lithuanian folk songs (performed by four female singers). The examples contain short segments (app. 0.5–0.7 s) of typical per-

formance situation: steady tone, vibrato tone, f0 track of inverted U shape, and tone with embellishments.

2.3 Procedure

Based on our previous considerations, the certain PDA was evaluated as adequate if it met the following requirement: in the case of sine wave examples, the deviation between the detected and actual pitch tracks did not exceed ± 10 cents. Additionally, the PDA was preferred if it detected unsteady fundamental frequency as accurate as possible (or more exactly, if it performed better than other PDAs).

The pitch tracks of all sound examples were generated using all three PDAs. Default settings of the algorithms were used. For AC algorithm, pitch analysis range was adjusted depending on the certain sound examples to prevent octave errors. For YIN algorithm applied to some examples, the size of analysis window was trimmed (compared to default size) to achieve significantly better results⁷. Pitch detection was sampled at every 10 milliseconds.

3. RESULTS

3.1 Structure of Results

The output from each algorithm consists of a set of pitches (samples) corresponding to the fundamental tone of a sound example at every 10 milliseconds. Pitch is represented in semitones relative to A4 (440 Hz). Generated pitch tracks of sine tone examples were aligned with the actual ones in time, so the difference between the two pitch tracks was the smallest possible. The actual pitch tracks were known in advance as they were determined in the synthesis of the sound examples. In contrast, it is impossible to obtain the actual pitch tracks of the natural voice examples⁸, therefore the generated pitch tracks of the natural performances were compared only between each other.

3.2 Absolute Deviation

We will start with a discussion of the sine wave examples. The certain method to be chosen for the evaluation of deviation between the detected and actual pitch tracks depends on the type of sound example (steady tone, glide, etc.). In the case of sine wave examples, the method of averaged absolute deviation was applied: the absolute difference between each pair of samples in the two pitch tracks (generated and actual) was found and then the mean was calculated. Table 1 shows the results for all sound examples in the case of each PDA.

SWIPE⁹ performs the worst: its output differs from the actual pitch app. from 2 to 19 cents, on the average. Other two algorithms perform well enough (the deviation is mostly considerably less than 1 cent).

³ <http://www.fon.hum.uva.nl/praat/>

⁴ <http://audition.ens.fr/adc/>

⁵ <http://www.cise.ufl.edu/~acamacho/english/>

⁶ WaveLab 6 software was used to synthesize all sine wave examples. The software has some limitations on generating waveforms, therefore the closest numbers to 50, 100 and 200 cents were chosen.

⁷ SWIPE⁹ did not make any octave errors; therefore analysis range was not defined for this algorithm. Changes in other SWIPE's settings did not result in significant improvement of results.

⁸ Probably the only way to get the pitch track very close to the actual one would be the laryngograph-aided measurement procedure during the recording session of performance.

Sound example		AC	YIN	SWIPE'
ST	110 Hz	.22	0	6.60
	220 Hz	.23	0	4.22
	440 Hz	.25	0	2.14
	880 Hz	.27	0	3.15
VT	110 Hz	±48 ct	.79	.70*
		±96 ct	1.18	1.42*
	220 Hz	±48 ct	.27	.30*
		±96 ct	.38	.62*
	440 Hz	±48 ct	.26	.27*
		±96 ct	.33	.53*
	880 Hz	±48 ct	.26	.22*
		±96 ct	.34	.45*
	110 Hz	±96 ct	.27	.15
		±204 ct	.27	.40
SVT	220 Hz	±96 ct	.28	.12
		±204 ct	.22	.27
	440 Hz	±96 ct	.28	.10
		±204 ct	.24	.22
	880 Hz	±96 ct	.26	.10
		±204 ct	.31	.21
	2 s		.04	.21
	5 s		.04	.09
G				19.40
				9.69

Table 1. Averaged absolute deviations for sine wave examples in the case of each PDA; in cents. ST – steady tone, VT – vibrato tone, SVT – “slow vibrato” tone, G – glide. Bold numbers highlight the best results. Here and hereafter, the asterisk indicates the results for the adjusted sizes of analysis window (YIN).

3.3 Averaged Pitch

The method discussed above allows identifying the overall discrepancies between the actual and generated pitch tracks. However, these discrepancies may have only negligible influence on final results in certain situations. For example, the integral (perceived) pitches in a natural vocal performance are evaluated based on segments of quasi-stationary fundamental tone or on several full cycles of vibrato; the pitch average is calculated⁹. Small discrepancies in the output of PDA may occur due to the digital signal processing. Quite possibly, significant cancellation of the averaged deviations can take place if sufficiently long segment of pitch track is averaged. If it is not the case, then the mean of the detected pitch will differ from the actual pitch. The method of comparison of the two averaged pitches of the actual and generated pitch tracks was applied for the steady and vibrato sine wave examples. Table 2 shows the results for these sound examples in the case of each PDA. Again, with a few exceptions, SWIPE' performs the worst while other two algorithms are adequate enough.

Let's take one more look at SWIPE'. Figure 1 shows two pitch tracks of steady sine tone, the actual one and the one generated by SWIPE'. It is clearly seen that, to say nothing of some minor bumps, the SWIPE's track is app. 6 cents sharp compared to the actual pitch track. Table 2 shows that SWIPE' behaves quite similar in many cases: the output is sharp or flat app. from 2 to 6 cents. While this accuracy does not exceed the desirable

limit of ± 10 cents (except for one case), this is far away from the results AC and YIN possess.

Sound example		AC	YIN	SWIPE'
ST	110 Hz	-.22	0	6.28
	220 Hz	-.23	0	4.00
	440 Hz	-.25	0	2.00
	880 Hz	-.27	0	-3.00
VT	110 Hz	±48 ct	-.26	.03*
		±96 ct	-.09	.11*
	220 Hz	±48 ct	-.22	.02*
		±96 ct	-.16	.06*
	440 Hz	±48 ct	-.25	.02*
		±96 ct	-.19	.05*
	880 Hz	±48 ct	-.25	.02*
		±96 ct	-.21	.06*
				.01

Table 2. Comparison of averaged pitches for steady and vibrato sine tones (ST and VT respectively) in the case of each PDA; in cents. The pluses and minuses show whether mean of detected pitch is sharp or flat compared to the actual pitch.

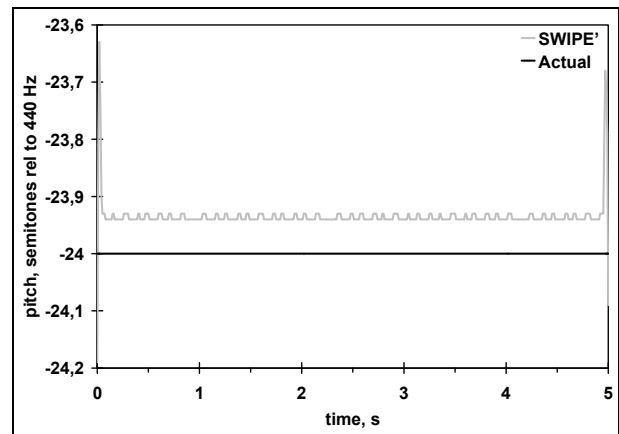


Figure 1. Steady sine tone (110 Hz). The actual pitch track and output of SWIPE'.

3.4 Shape of Pitch Track

Many research projects on vocal or instrumental performances investigate the phenomena of vibrato and other types of rapid changes of fundamental tone (e.g., various embellishments). Thus it is important for PDA to generate the correct shape of pitch track. The extreme points of pitch track (e.g. minima and maxima of each vibrato cycle) should be detected as accurate as possible. Since all vibrato cycles are identical in the synthesized sine tones with vibrato and “slow vibrato”, the method of comparison of the generated single vibrato cycle was applied for these two types of sound examples¹⁰. Table 3 shows the results.

SWIPE' shows deviations exceeding the limit of ± 10 cents in many cases and the discrepancies are very irregular, i.e., this algorithm detects wrong pitch in unpredictable way. This PDA performs poorly even in the case of much slower and deeper vibrato. Both AC and YIN perform quite well with both (“normal” and “slow”) vibrato.

⁹ This is the second issue of musical scale estimation; see footnote 2.

¹⁰ In the case of vibrato tones, one cycle was 16 samples long, and in the case of “slow vibrato” tones, one cycle was 100 samples long.

Another way to judge the quality of algorithm output is simple visual evaluation. Figure 2 shows the actual pitch track and outputs of all PDAs for single vibrato cycle. The output of AC is slightly shifted back in time axis, while the output of YIN is slightly shifted forward. Yet the pitch tracks given by the algorithms are very similar to the actual one. The SWIPE's pitch track is slightly shifted forward in time axis and much wider than the actual one. Time shifts occur due to the peculiarities of different algorithms; yet they have no influence on pitch estimation.

Sound example			AC		YIN		SWIPE'	
			min	max	min	max	min	max
VT	110 Hz	±48 ct	-.12	-.90	.84*	-.92*	4.00	-.64
		±96 ct	1.61	-1.95	1.80*	-1.58*	20.99	-12.09
	220 Hz	±48 ct	.32	-.48	.48*	-.44*	-4.22	17.84
		±96 ct	-.39	-.63	.96*	-.98*	3.57	9.71
	440 Hz	±48 ct	0	-.17	.39*	-.39*	-15.56	20.40
		±96 ct	-.94	-.54	.69*	-.73*	-.95	7.97
	880 Hz	±48 ct	-.61	.05	.33*	-.33*	-16.68	17.31
		±96 ct	.56	-.45	.65*	-.66*	-8.34	10.64
SVT	110 Hz	±96 ct	-.21	.17	.10	-.22	-6.03	14.02
		±204 ct	-.07	.03	.33	-.33	-5.09	10.08
	220 Hz	±96 ct	-.14	.24	.12	-.14	-5.01	14.00
		±204 ct	-.43	.12	.36	-.31	-8.03	10.02
	440 Hz	±96 ct	-.16	.36	.14	-.14	-5.01	13.00
		±204 ct	-.16	.12	.31	-.30	-10.01	7.00
	880 Hz	±96 ct	-.27	.55	.14	-.12	-9.00	10.00
		±204 ct	-1.06	.24	.31	-.35	-14.00	3.00

Table 3. Comparison of minima and maxima of single vibrato cycle for sine tones with vibrato and “slow vibrato” (VT and SVT respectively) in the case of each PDA; in cents.

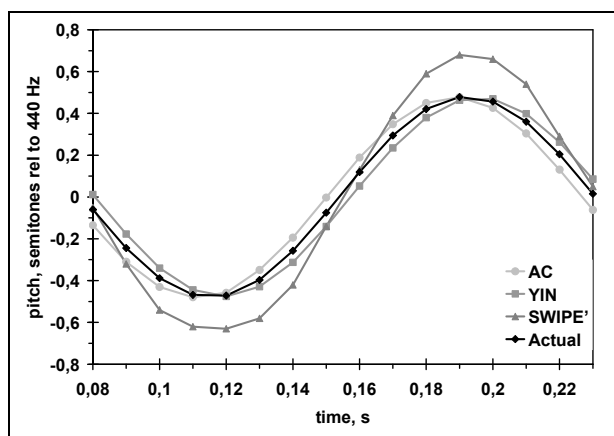


Figure 2. Single cycle of sine vibrato tone (averaged fundamental frequency 440 Hz, vibrato frequency 6.5 Hz, and amplitude ±48 cents). Actual pitch track and outputs of all PDAs.

3.5 Glides

Table 1 shows that AC and YIN detect pitch tracks of glides almost perfectly; therefore there is no need to investigate the performance of these two algorithms further in the case of (long perceived) glides. However, SWIPE' generates quite different pitch track, so visual evaluation is helpful for the judgment of the algorithm. Figure 3 shows small portion of the actual pitch track and SWIPE' output for gliding sine tone. Though the fundamental tone

risers evenly, SWIPE' detects strange bumps which differ from the actual pitch up to app. 21 cent (up to app. 29 cents for points not shown in the Figure 3).

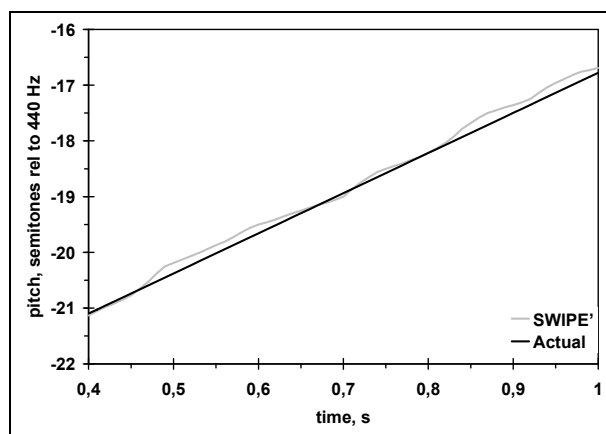


Figure 3. Beginning of gliding sine tone (length of 5 s). Actual pitch track and output of SWIPE'.

3.6 Examples of Natural Voice

Figures 4–7 show the outputs of all algorithms for four typical situations occurring in the vocal performance. Note that the range of pitch axis varies from 30 to 450 cents in the figures. Therefore, seemingly the huge difference between the pitch tracks in Figure 4 is actually negligible in comparison to the one observed in Figure 6.

In general terms, the pitch tracks generated by AC and YIN are more similar in between than to those of SWIPE'. However, YIN and SWIPE' give generally smoother tracks than AC. These differences could be possibly attributed to the different levels of smoothening resulting from the different approaches and default settings of the PDAs.

Despite obvious visual differences, the results for the examples of natural voice can be examined applying some methods used in the case of sine tones; see the results in Table 4. Except for the case of the quasi-steady tone, the algorithms present quite different fundamental frequencies, on the average. Nevertheless, the performances of AC and YIN differ the least.

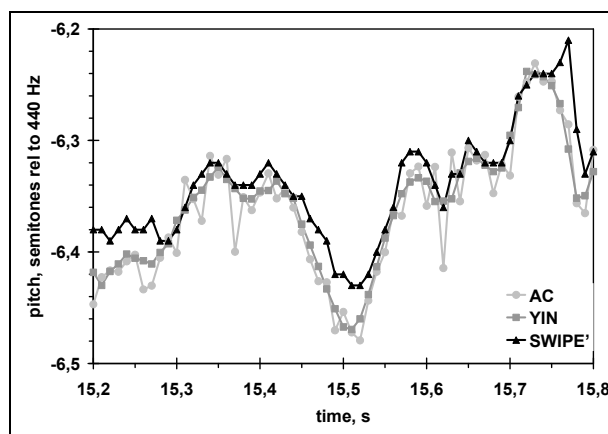


Figure 4. Example of natural voice, quasi-steady tone. Comparison of outputs from all PDAs.

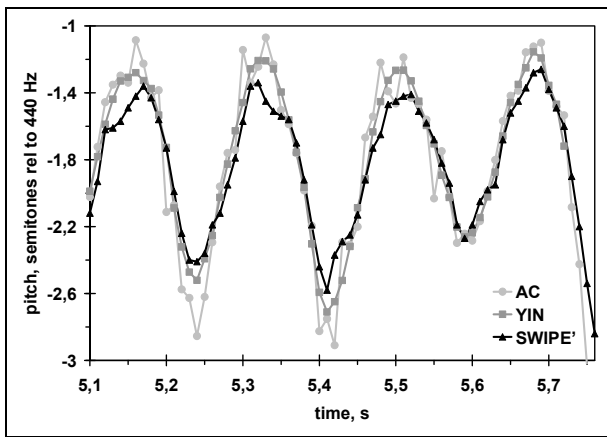


Figure 5. Example of natural voice, vibrato tone. Comparison of outputs from all PDAs.

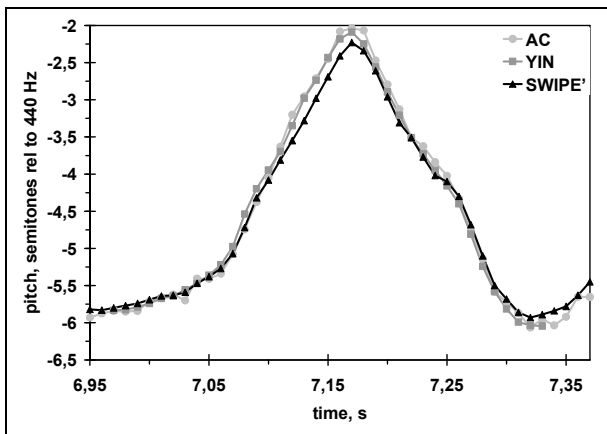


Figure 6. Example of natural voice, pitch change of inverted U shape. Comparison of outputs from all PDAs.

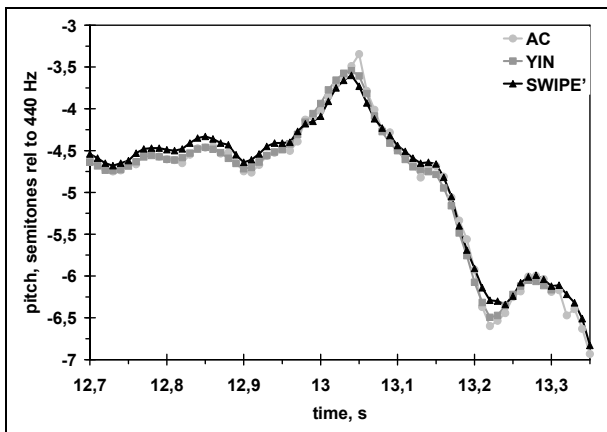


Figure 7. Example of natural voice, tone with embellishments. Comparison of outputs from all PDAs.

Sound example	AC-YIN	AC-SWIPE'	YIN-SWIPE'
Steady	1.53	2.64	1.92
Vibrato	10.49	15.38	9.79
Inv. U shape	7.84	11.33	10.31
Embellishments	5.41	9.46	9.39

Table 4. Averaged absolute deviations between pairs of outputs from all PDAs for the examples of natural voice; in cents.

In the cases of quasi-steady and vibrato tones, the averaged pitches of these tones given by the PDAs can be compared. In both cases, the difference between AC and YIN is as small as app. 0.4 cents, while the difference between these algorithms and SWIPE' varies from 1.8 to 3.1 cents.

4. DISCUSSION

Although pure sine tones do not exist naturally, they could serve as good reference points for testing PDAs, and then applied in the study of natural sounds. Obviously, even under ideal conditions, neither algorithm can perform without small errors. Nevertheless, many of them are much smaller than pitch jnd and therefore practically do not have any influence on the results if phenomena based on perception are then studied. Importantly, testing PDAs for sine wave examples showed some fault tendencies which, probably, can manifest even in larger degrees when estimating pitch in natural sounds. Among PDAs examined, SWIPE' is distinguished by poor and unpredictable performance (with errors exceeding ± 10 cents) with sine tones. AC and YIN perform quite well with all types of sine waves making no significant pitch detection errors. However, both algorithms need certain minor adjustments, i.e., defining analysis range (AC) and the size of analysis window (YIN), for the best results.

As for the natural vocal sounds, some comparisons showed that AC and YIN generated relatively similar pitch tracks, and they both are suitable for the estimation of mean pitch of quasi-stationary and vibrato tones of natural voice. However, these algorithms should be used with caution for detailed research of vibrato and rapid pitch changes. For the examples of natural voice, the output from SWIPE' lacks sufficient accuracy. First, the inaccurate results for sine tones mean that this algorithm would probably perform even worse with more complicated waveforms. Second, it is less possible that similar pitch estimates given by the other two algorithms are more inaccurate than those of SWIPE'.

Some software packages make use of PDAs discussed here: AC is implemented in PRAAT software, YIN and SWIPE' – in Tarsos software¹¹, AC and SWIPE' – in PsySound3 software¹², SWIPE' – in NoteView software¹³, etc. It is worth mentioning that after some results are obtained employing certain software, the accuracy and reliability of the algorithm used should be taken into account before making scientific conclusions. For instance, the authors of NoteView software demonstrate its possibilities by analyzing performances of horn player, and claim that “The median of the within-note SD across the two performances was 6 cents with a maximum of 18 cents and a minimum of 2 cents” [19: 28]. While keeping in mind the accuracy of SWIPE' (which is implemented in NoteView), it is hard not to doubt the accuracy of these results as well.

¹¹ <http://tarsos.0110.be/>

¹² <http://psysound.wikidot.com/start>

¹³ <http://sam.arts.unsw.edu.au/research-and-creative-practice/research-projects/empirical-musicology/>

The test of PDAs presented in the current paper demonstrates only some features of the discussed algorithms. Extension of the research would include larger samples of synthesized and natural sound examples, more PDAs, statistical verification of the results, and consideration of each PDA from acoustical and mathematical viewpoints.

Acknowledgments

The study is supported by European Social Fund (Global Grant).

5. REFERENCES

- [1] R. Ambrazevičius, “Pseudo-Greek modes in traditional music as result of misperception,” in *9th International Conference on Music Perception and Cognition*, Bologna, 2006, pp. 1817–1822.
- [2] R. Ambrazevičius, “Modelling of scales in traditional solo singing,” *Musica Scientiae*, Special Issue “Interdisciplinary Musicology”, pp. 65–87, 2005–2006.
- [3] R. Ambrazevičius and R. Budrys, “Pitch Evaluations in Traditional Solo Singing: Comparison of Methods,” in *12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music*, Thessaloniki, 2012, pp. 58–63.
- [4] W. Hess, *Pitch determination of speech signals. Algorithms and devices*. Berlin: Springer-Verlag, 1983.
- [5] R. Parncutt and A. J. Cohen, “Identification of microtonal melodies: Effects of scale-step size, serial order, and training,” *Perception and Psychophysics*, vol. 57, no. 6, pp. 835–846, 1995.
- [6] E. Zwicker and H. Fastl, *Psychoacoustics. Facts and models*. Berlin, Heidelberg: Springer-Verlag, 1999.
- [7] D. Deutsch, “The processing of pitch combinations,” in *Psychology of Music*, 2nd ed., D. Deutsch, Ed. San Diego, London: Academic Press, 1999, pp. 349–411.
- [8] B. Snyder, *Music and Memory: An Introduction*. Cambridge, MA, London: The MIT Press, 2000.
- [9] I. V. Nábělek, A. K. Nábělek, and I. Hirsh, “Pitch of tone bursts of changing frequency,” *Journal of the Acoustical Society of America*, vol. 48, no. 2, pp. 536–553, 1970.
- [10] J. C. Brown and K. Vaughn, “Pitch center of musical sounds with vibrato,” *Journal of the Acoustical Society of America*, vol. 94, no. 3, pp. 1860–1860, 1993.
- [11] C. d’Alessandro and M. Castellengo, “The pitch of short-duration vibrato tones,” *Journal of the Acoustical Society of America*, vol. 95, no. 3, pp. 1617–1630, 1994.
- [12] C. d’Alessandro and P. Mertens, “Automatic pitch contour stylization using a model of tonal perception,” *Computer Speech and Language*, vol. 9, no. 3, pp. 257–288, 1995.
- [13] P. Boersma, “Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound,” *Proceedings of the Institute of Phonetic Sciences*, vol. 17, pp. 97–110, 1993.
- [14] A. de Cheveigne and H. Kawahara, “YIN, a fundamental frequency estimator for speech and music,” *Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1917–1930, 2002.
- [15] A. Camacho and J. G. Harris, “A sawtooth wave form inspired pitch estimator for speech and music,” *Journal of the Acoustical Society of America*, vol. 124, no. 3, pp. 1638–1652, 2008.
- [16] D. E. Hall, *Musical Acoustics*. Pacific Grove, CA: Brooks/Cole, 1990.
- [17] P. Desain et al., “Rhythmic aspects of vibrato,” in *Rhythm Perception and Production*, P. Desain and W. L. Windsor, Eds. Lisse: Swets & Zeitlinger, 1999, pp. 203–216.
- [18] T. D. Rossing, F. R. Moore, and P. A. Wheeler, *The Science of Sound*. 3rd ed. San Francisco, Boston, New York [...]: Addison-Wesley Publishing Company, 2002.
- [19] D. Gunawan and E. Schubert, “NoteView: a computer program for the analysis of single-line musical performances,” *Acoustics Australia*, vol. 38, no. 1, pp. 25–30, 2010.

FORMANT TUNING IN BYZANTINE CHANT

Georgios Chrysochoidis

Georgios Kouroupetroglou

Dimitrios Delviniotis

Sergios Theodoridis

National and Kapodistrian University of Athens, Department of Informatics and Telecommunications,
Panepistimioupolis, Ilisia, GR-15784 Athens, Greece
{geoxry, koupe, ddelvis, stheodor}@di.uoa.gr

ABSTRACT

We present an investigation of formant tuning in the context of the Byzantine Ecclesiastic chant voice. The recordings selected for the analysis are part of the DAMASKINOS prototype acoustic corpus of Byzantine Ecclesiastic voice. More specifically, we analyzed recordings from 10 different chanters in ascending musical scales of the diatonic genre, for the /a/ vowel. The method of analysis included a semi-automatic segmentation of the audio material, extraction of the measurements in PRAAT and the final post-processing in MATLAB. Results show clear evidence of formant tuning in at least six of the chanters, proving that the technique is in use by the modern Byzantine music performers.

1. INTRODUCTION

In 1960, Gunnar Fant presented the theory of a source-filter production model for vowels, in his work “Acoustic Theory of Speech Production” [1]. According to this model, the voice source produces a harmonic series, consisting of the fundamental frequency f_0 and a large number of harmonic frequencies, the partials [1].

Specifically, when applied to vowel production, the speech signal could be thought as the result of the source signal, produced by the glottis, and the resonator or vocal tract filter. A linear mathematical model supports this theory, which allows for relatively simple handling of calculations. The vocal tract filter can be further considered as a linear time-invariant filter for very short periods of time, making the system even more mathematically tractable.

The resonance frequencies of the vocal tract are called formants, designated as F1, F2, F3, etc. in ascending order, and they can be displayed as spectral peaks in the frequency response of the vocal tract filter [1]. The vocal tract has four or five important formants that are used to amplify and dampen certain frequencies. The length and shape of the vocal tract determine the formant frequencies, resulting in the production of the different vowel sounds of the radiated speech signal. The lowest two formants F1 and F2 largely determine the vowel [1],

while the remaining higher order formants are related to the quality of tone [2].

Adjusting the vocal tract in order to align formants with harmonics, thereby amplifying certain portions of the vocal spectrum, is known as *formant tuning* [3]. This intuitive act, that is highly dependent upon the vertical laryngeal positions, has been used from trained singers in the past, in cases where the singing voice should be heard across large spaces along with other sound sources, like for example music orchestras [3].

In the past, research concentrated on the relation between the quality of the voice and the formants. Later works [4,5], revealed the existence of the singer’s formant, which can be explained acoustically as a clustering of formants, F3 and F4, or F4 and F5, and even in some cases F3, F4, and F5. The singer’s formant enables a singer to be heard over an orchestra, since there is little competition from the orchestra near the frequency range of the singer’s formant [4]. This is a way for the singer to save some vocal effort, in other words it results in “vocal economy” [3].

Formant tuning is considered as another vocal strategy used by trained singers when trying to produce the ideal voice and economizing on vocal effort. Several works have been published concerning formant tuning strategies applied by classically trained Western operatic voices [6,7,8,9], as well as contemporary [10] and traditional [11] ones. Recent works [8,9] give a description of the formant tuning literature with details about the different methods and their limitations [9]. Three trends seem to be the dominant ones: a. F1 and F2 are tuned to a partial, b. F1 and F2 are not related anyhow to harmonics of the f_0 and c. F1 and F2 are tuned just above their nearest partial in a way so that they don’t coincide.

While literature on formant tuning continues to grow for other types of vocal music genres, Byzantine Ecclesiastic chant voice hasn’t been studied in the same context. Current work tries to fill this gap by providing an initial approach, along the path followed by other recent investigations on this subject, and presents the first results of this preliminary analysis.

2. BYZANTINE CHANT MUSIC

Byzantine Chant Music (BCM) is a religious type of monophonic vocal performance practiced mainly in churches. Its main purpose is to serve the religious needs of the Orthodox Christian worship, providing a musical accompaniment for the ecclesiastical poetry [12,13]. BCM follows aesthetical rules formed over the course of

centuries, traditionally transferred from master performers to apprentices. It is a microtonal music since it contains intervals smaller than the conventional contemporary Western theory semitone.

The temperament used in modern Byzantine chant is based on a 72 equal divisions of the octave, called *moria* (plural form of *morio*) [13]. Compared to the 12-tone equal temperament followed by the Western theory, a Western semitone would equal to 6 *moria* in this 72-tone equal temperament system. Byzantine music uses the term *echos* to refer to a specialized type of musical mode [12,13], denoting not only the musical scale being applied in a melody with a definite “tonic” or main note called *vasi* (Greek word for “base”), but also specific musical phrasing [13].

There are three musical genera in Byzantine chant: the diatonic, the chromatic and the enharmonic [12,13]. The fundamental intervallic differences between the three genera, apart from musical aesthetics, serve different musical meanings, when considered in a specific lyric context. This leads to another classification of the three genera and its subdivisions, based on the *ethos* of the music genre, as Chrysanthos of Madytos, one of the main three reformers of the modern Byzantine music theory, named it [12,14].

The musical scale used by the chanters in this work belongs to the diatonic genre and it is shown in Figure 1.

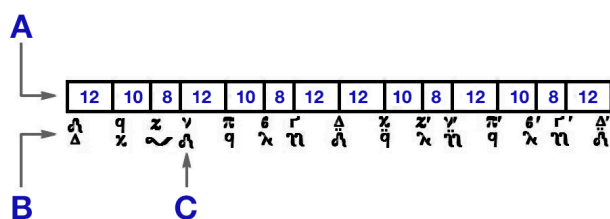


Figure 1. Diatonic genre scale used in Byzantine chant.

A: Scale intervals measured in *moria* (72-tone equal temperament). B: Symbols used for the scale degrees. C: Scale's tonic.

3. MATERIALS

The materials selected to be analyzed in the present work are part the DAMASKINOS prototype acoustic corpus of Byzantine Ecclesiastic voice [15,16]. The DAMASKINOS corpus was designed and developed in the Department of Informatics and Telecommunications, University of Athens. Our aim was to create a standard, tagged corpus of BCM, from a statistically representative sample of modern performers. The entire corpus consists of the recordings of twenty chanters carefully selected from entire Greek territory, recorded strictly under the same controlled conditions of digital recording, with a second channel for the electroglottograph signal (EGG), as well as other measurements like glottal flow and subglottal pressure. The musical performance of each subject followed a structured protocol and included: a) specific music sections of all the musical genera from the entire repertoire of Byzantine chant (*papadika*, *stehera*, *irmologika*) and b) musical exercises in all the musical scales per genre. All the sections were recorded, except from being chanted, at their spoken and musical

recitation as well, in normal intensity of voice and in the basic octave of C3 [15]. The contents of the corpus were tagged using multiple layers of metadata [17]. These data mainly consist of the lyrics syllables, the musical notation symbols of the Byzantine music score and the musical intervals. Each of these types of data forms a tagging layer on its own. A couple of extra layers are used for parts of the recordings where the chanter overlooked the written score, by means of musical expression (intervals, score) [17].

In this work we used a subset of the DAMASKINOS corpus, which consists of ascending musical scales for the vowel /a/, performed by a total of 10 chanters. During recordings, the chanters were asked to perform the intervals of the diatonic scale presented in Figure 1 for the specific vowel. Chanters were free to choose the performed pitch span that suits best to their vocal range.

4. METHODS

The measurements used in our analysis were acquired using the PRAAT software. PRAAT is a valuable software tool in the field of phonetics and voice analysis in general [18]. It is a flexible tool, which provides functionality that could handle most of the tasks needed in this work. It handles all the widely used audio file formats by importing the audio files as objects on which various operations can be applied. Apart from sound files, PRAAT can create and make use of tagging files in the form of layers called *tiers*. Tiers consist of boundaries and intervals between them. Labels can be added at these intervals making it easy to tag sound recordings. Several tiers can form a TextGrid object. To summarize the main PRAAT features we used in our analysis: a. it handles large audio files, b. it extracts measurements of the vocal parameters using its built-in functions, c. it uses tagging layers for audio file annotation, and d. it uses a scripting language for automating processing. Indeed, PRAAT can manipulate, edit and analyze long stereo audio files. Annotation of the files is done with tagging layers, using boundaries to mark time exact points in the recording, and the intervals in-between them to insert the metadata. One of the most powerful features of PRAAT is inevitably its scripting language. This language includes variables, loops, jumps, formulas, procedures, arrays, etc. which provide the flexibility to implement complex algorithms in combination with the ready-to-use analysis commands found in PRAAT.

The functions that were mostly used in our measurements are the ones for pitch and formant analysis. For pitch analysis PRAAT uses an algorithm that performs an acoustic periodicity detection on the basis of an accurate autocorrelation method [19]. Formant values are calculated with the algorithm by Burg [20,21]. PRAAT is a commonly used speech analysis tool and its accuracy has been thoroughly tested in several related papers [22,23,24]. Furthermore, PRAAT measurements are used in many investigations as the baseline for accuracy comparisons [25,26,27,28].

We started processing the audio files by using a series

of PRAAT scripts that analyzed intensity and pitch, in order to annotate the audio file, by marking the basic musical units, which correspond to the scale's degrees [17]. Our final purpose was to extract separate audio files for each scale's degree. This was a three-step process: 1. voiced – unvoiced parts of the audio file were labeled using a tag layer, 2. for each voiced part boundaries were placed at pitch transitions between the scale's degrees, and 3. intervals in-between boundaries were labeled according the audio segment's average pitch. Before moving forward to the next step we were able to fine-adjust the placement of the boundaries. The transition zones between notes have been excluded. Each audio segment between boundaries was extracted to a separate audio file with average time duration of 1 sec.

Next, we analyzed each of the extracted files, using PRAAT's readily available functions, in order to acquire the actual measurements we needed. Three vocal parameters were measured every 10 msec: pitch value for f_0 , frequency values for the formants F1 and F2. Other measurements, like sound intensity level, formant bandwidths, formant levels and partials levels were also extracted, although not used in the current work. All data were stored in tab-delimited text files. Final processing was done in Matlab. First the data files were imported, the mean values were calculated for pitch and formant frequencies and the corresponding graphs were plotted.

5. RESULTS

Displaying the actual measurements for each chanter is the next step in our investigation. This involves plotting the frequency tracks of the two lowest formants F1 and F2, along with the partials h_2 - h_8 , where $h_n = n \cdot f_0$ and f_0 is the fundamental frequency of the vowel. In Figure 2 we present these measurements for each chanter. Both axes represent frequency values in semitones, measured from D2 (73 Hz).

Results for chanter 1 show a clear tendency of the F1 and F2 formants to follow the slope of the h_2 and h_4 respectively, above the C4 note. This is clearly depicted by a simultaneous characteristic bend in the track of both F1 and F2 around the pitch of C4.

Formants F1 and F2 of chanter 2 follow a similar pattern. The simultaneous bend at about the same pitch, as the one noticed for chanter 1, is found in F1 and F2 of chanter 2, although F2 jumps over to the h_3 partial at the scale's peak note.

Chanters 3 and 4 are two more complicated cases. While both of them show signs of formant tuning, it is not clear if this is done intentionally. Specifically, chanter 3 seems to tune F1 to h_3 , then to h_2 , while getting higher in the pitch range between C3 and D#4. On the other hand, chanter 4 shows a quite similar F1 track to the one described for chanter 3. Indeed F1 tunes initially to h_3 from C3 to F3, then tunes down to h_2 from A3 to D4. In both 3 and 4 chanters F2 tunes to h_3 at the highest part of the scale, namely between E4 and A4.

Formants F1 and F2 for the chanters 5 and 6 seem to be unrelated to the partials. Despite the fact that for some

tones they seem tuned on the harmonics, the overall image of the two formants tracks, leads us to believe that this was a coincidence and not the result of an intentional tuning effort. An exception seems to occur at the top 3-4 notes where F2 is tuned to h_3 or h_4 .

Chanter 7 tunes its F1 to the h_2 partial starting approximately at D4 and keeps it tuned up until the ultimate note of G#4. Formant F2 for the same chanter tunes to h_4 at about C4 then jumps to h_3 at F4.

Chanter 8 seems to start tuning its F1 formant to h_3 at about E3 and then tunes it to h_2 at G#3. Its F2 formant remains relatively constant until A3 where it tunes to h_4 .

Chanter 9 mainly tunes F1 to h_2 at E4, but also shows signs of tuning F1 to h_3 at the region between F#3 and A3. Its F2 formant tunes to h_4 above E4 and then jumps to h_3 at the last note of its range.

Chanter 10 is the clearest evidence of formant tuning among the sample of 10 chanters. It keeps its F1 tuned to h_2 throughout the range of B3-G4 that is 6 consecutive notes. In the same pitch region F2 is tuned to h_4 for the first 5 notes then finally tunes to h_3 .

6. DISCUSSION

Previous research has shown that singing voice classification could be based on the formant frequency values and vocal range [29,30,31]. This way a psaltic voice can be classified in one of the classical music voice types. Moreover, although BCM performers have not been exposed, traditionally, to classical vocal training techniques, formant tuning can be considered as an intuitive act [3], and therefore be performed by non-classical trained voices as well.

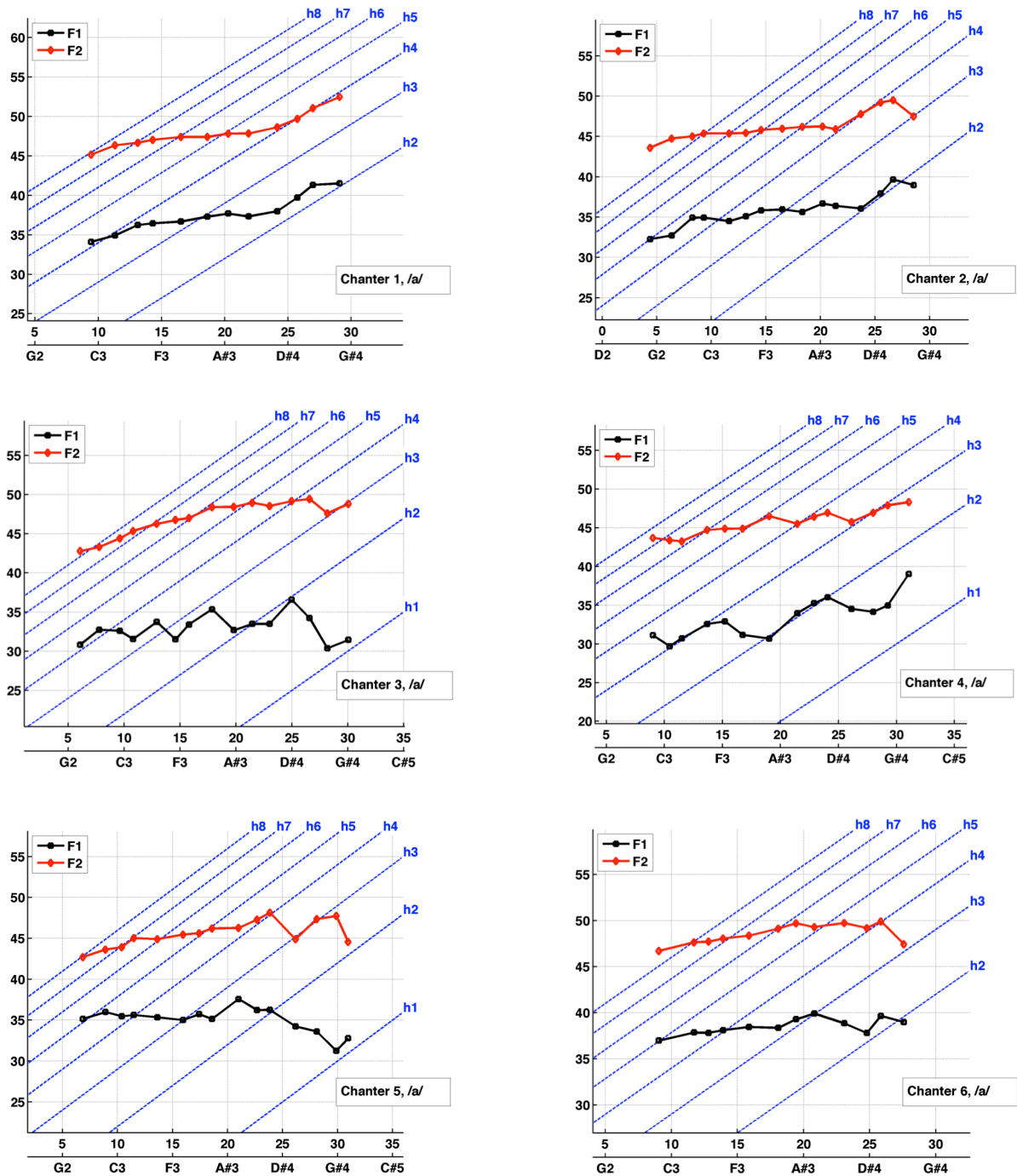
The pitch range that draws our attention is the one referred to as the *passaggio*, which is D4-G4 for tenors and B3-E4 for baritones [6,9,29]. It has been found that formant tuning is mostly observed in and above this range [9]. This can be easily seen in our results. Most chanters tend to tune their F1 and F2 to partials at this pitch range, although there were cases of formant tuning at lower scale degrees.

Considering an accuracy margin of about 20 Hz for the tuning of the formant frequencies [32,9], as well as the maximum distance criterion of 50 Hz between the formant and its nearest harmonic, used in similar works [9], a 2 semitone approximate distance between the formant F1 and the partial h_2 , in the frequency range of E4, could still be counted as tuning [9]. This could easily explain the distance between the F1 and h_2 , found in the results for chanter 1, since it can be considered inside the formant tuning tolerance limits.

The distance in semitones between each formant F1, F2 and their closest partial for the vowel /a/, is displayed in Figure 2. As before, the data presented in the graph are for the ascending diatonic scale, for all ten chanters (Ch1-Ch10). In order to draw safer conclusions, only notes above C4 (22 semitones above D2) have been accounted for. Indeed, the particular pitch range is near the *passaggio* region, where formant tuning is more likely to occur.

By examining the F1 graph of Figure 3 we notice that at

Formants F1 & F2 (semitones above D2)



Fundamental frequency (semitones above D2)

Figure 2. Frequencies, in semitones from D2 (73Hz), of the formants F1 and F2 for the vowel /a/. Results are presented for each chanter for the ascending diatonic scale. Harmonics h2-h8, where $h_n = n \cdot f_0$ and f_0 is the fundamental frequency of the vowel, are displayed by the diagonal blue lines.

least 6 of the total of 10 chanters, analyzed in the current investigation, tend to tune their F1 to the h2 harmonic, at the pitch range above C4.

Formant tuning was most apparent in cases where the F1 and F2 remained relatively constant throughout the scale, before reaching a breaking point near the beginning of the passaggio region, like in the cases of chanters 1, 2, 8, 9 and 10. Another observation is that F2 formant of the ultimate scale's note was tuned to either h3 or h4 in all

chanters. In fact the F2 graph of Figure 3 shows that, above C4, all chanters seem to have their F2 tuned to one of the h3, h4 partials.

Regarding the question as to whether there is a common tuning strategy followed by most chanters, the answer is not obvious. Although in many cases we observed similarities between the chanters for F1 and F2, we tend to believe that each chanter follows his own personal strategy to achieve the aesthetic result he desires.

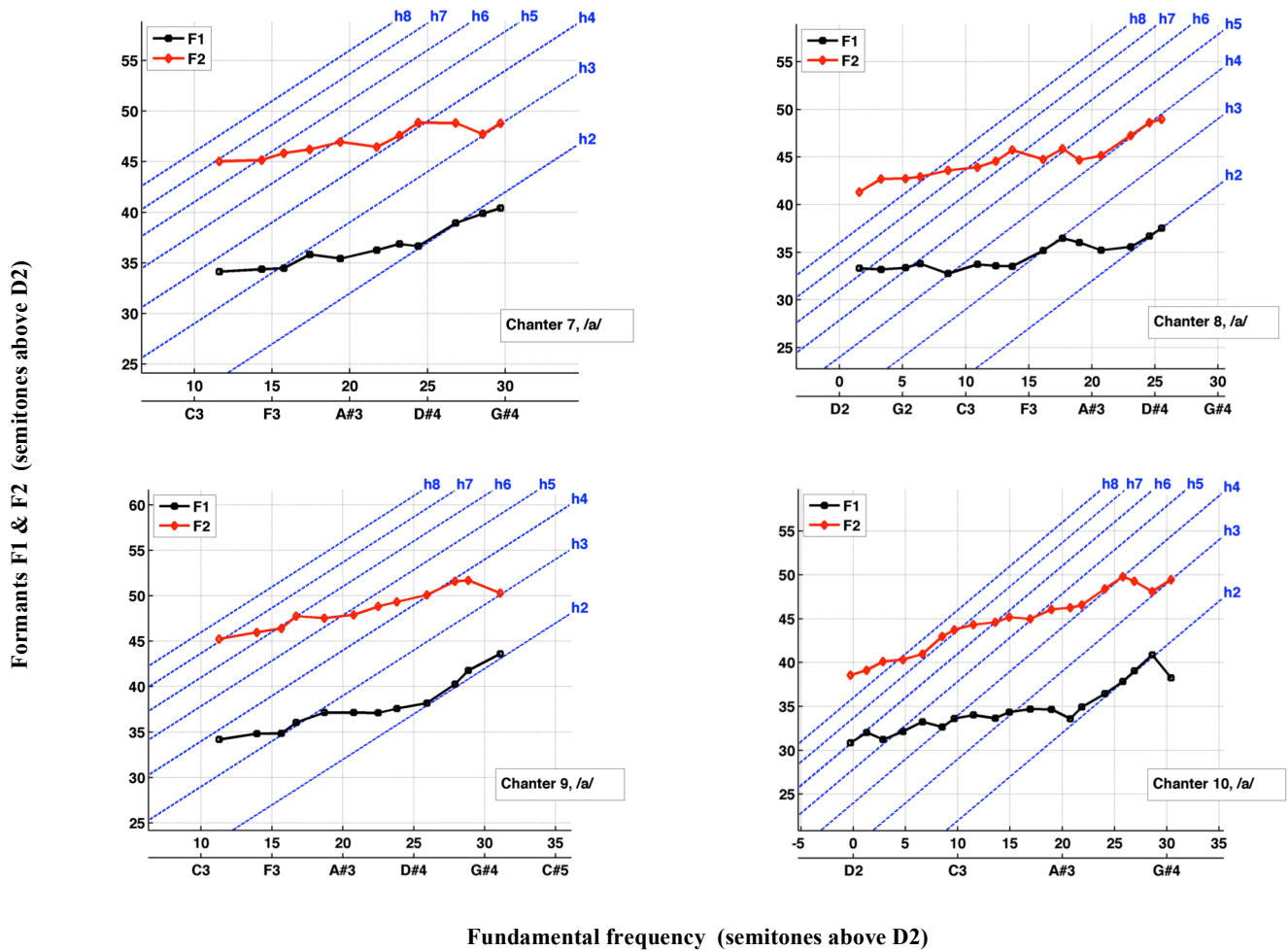


Figure 2. (Continued)

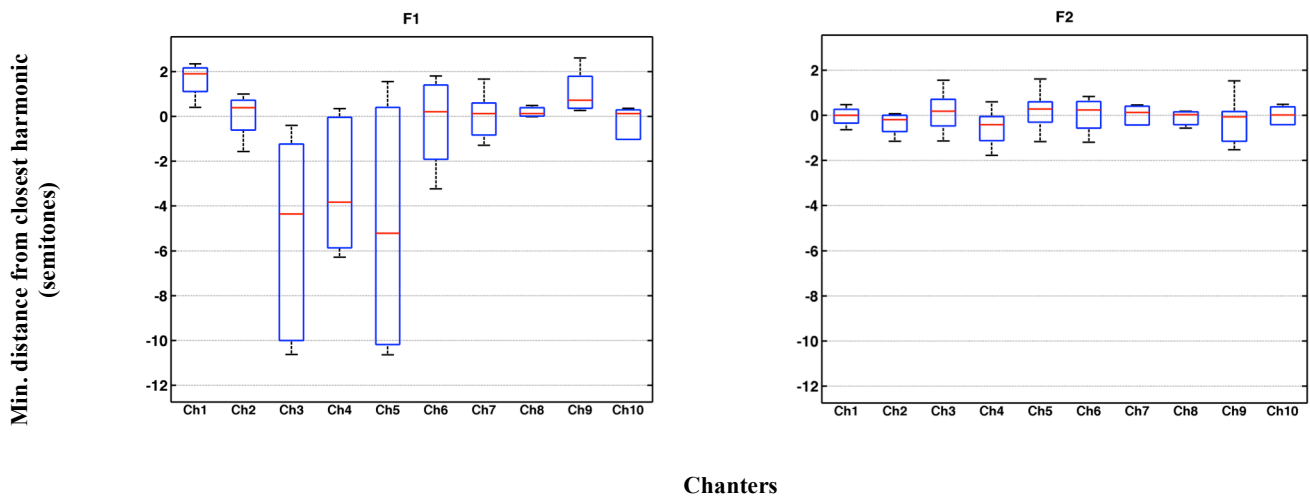


Figure 3. Distance in semitones between the first formant F1 and its closest harmonic and between the second formant F2 and its closest harmonic, for the vowel /a/. Results of all ten chanters (Ch1-Ch10) are presented for the ascending diatonic scale, for pitches above C4.

This last assumption agrees with the answer given by Sundberg et al. [9] and Henrich et al. [32] to a similar question, for the passaggio transition. Our opinion is also supported by the fact that most professional Byzantine chanters usually have not been given a systematic phonetic training as part of their vocal pedagogy that could lead in standardizing the Byzantine chant vocal

technique.

7. CONCLUSIONS

It is evident from both Figure 2 and Figure 3 that formant tuning is in use by the modern Byzantine chant

performers. Not only at high f_0 but also at the lower and middle scale regions, formant tuning has been observed in most of the total 10 chanters analyzed in the current investigation.

Going through the examples of formant tuning found in our results, formant F1 coincided, in most cases, with the partial h2, while formant F2 was tuned either to h3 or h4, in all cases.

A closer examination of the vocal range above C4 reveals that all chanters, in our sample, had their F2 resonance tuned to the partial h3 or h4.

In contrast to other types of singing voice, our results cannot be directly compared to other findings for the BCM in the field of formant tuning, since, to our knowledge, there aren't any similar publications on this subject. However when compared to results of professional male opera singers [9] from recent investigations, one might conclude that there is a stronger formant tuning tendency in BCM.

8. FUTURE WORK

Future work will mainly cover extracting and displaying more data from the DAMASKINOS corpus. For a start, this would include retrieving data from all the chanters of the corpus. Analyzing data from other vowels as well, would give us a clear picture as to whether chanters show a preference in formant tuning for specific vowels.

Analyzing and comparing results obtained from ascending and descending scales of the same genre could provide good insights into the intentionality of formant tuning. While Miller and Schutte [6] found no evidence of formant tuning in descending scales for baritone singers, considering our assumption that Byzantine chanters follow personal formant tuning strategies, as well as the opinion expressed earlier about the intuitive character of formant tuning, this comparison could give strong evidence on the intentionality of the technique in Byzantine chant.

Another important and interesting part of a future investigation would be the comparison of formant tuning strategies followed by BCM performers who have also been trained as professional opera singers and vice versa. The design of the DAMASKINOS corpus has included singers that meet the above conditions.

Finally, searching for a possible relation between formant tuning and different genera in BCM could also lead to important facts, both about the modern practice of Byzantine chant performance and also about the intentionality factor stated before.

Acknowledgments

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: "THALES - Investing in knowledge society through the

European Social Fund", under the project: "ERASITECHNIS: Processing, Analysis, Synthesis and Context-Based Retrieval for Multimedia Music-Related Data Bases of Traditional Music and Dancing Recordings", MIS 375435.

9. REFERENCES

- [1] G. Fant, *Acoustic Theory of Speech Production*. Walter de Gruyter, 1970.
- [2] G. Carlsson and J. Sundberg, "Formant frequency tuning in singing," *J. of Voice*, vol. 6, no. 3, pp. 256–260, 1992.
- [3] J.A. Stark, *Bel canto: a history of vocal pedagogy*. University of Toronto Press, 2003.
- [4] J. Sundberg, "The acoustics of the singing voice," *Sci. Am.*, vol. 236, no. 3, pp. 82–91, 1977.
- [5] J. Sundberg, *The Science of the Singing Voice*. Northern Illinois Univ, 1987.
- [6] D.G. Miller and H.K. Schutte, "Formant tuning in a professional baritone," *J. of Voice*, vol. 4, no. 3, pp. 231–237, 1990.
- [7] G. Carlsson and J. Sundberg, "Formant frequency tuning in singing," *J. of Voice*, vol. 6, no. 3, pp. 256–260, 1992.
- [8] J. Sundberg, F.M.B. Lã, and B.P. Gill, "Professional male singers' formant tuning strategies for the vowel /a/," *Logoped Phoniatr Vocol*, vol. 36, no. 4, pp. 156–167, 2011.
- [9] J. Sundberg, F.M.B. Lã, and B.P. Gill, "Formant Tuning Strategies in Professional Male Opera Singers," *J. of Voice*, Dec. 2012.
- [10] M.E. Bestebreurtje and H.K. Schutte, "Resonance strategies for the belting style: Results of a single female subject study," *J. of Voice*, vol. 14, no. 2, pp. 194–204, 2000.
- [11] N. Henrich, M. Kiek, J. Smith, and J. Wolfe, "Resonance strategies used in Bulgarian women's singing style: A pilot study," *Logopedics Phoniatrics Vocology*, vol. 32, no. 4, pp. 171–177, 2007.
- [12] D.G. Panagiotopoulos, *Theoria kai Praxis tis Byzantinis Ecclesiastikis Mousikis*, 6th ed. Athens: Sotir, 1997.
- [13] D. Delviniotis, G. Kouroupetroglou, and S. Theodoridis, "Acoustic analysis of musical intervals in modern Byzantine Chant scales," *J. of the Acoustical Society of America*, vol. 124, no. 4, pp. EL262–EL269, 2008.
- [14] M. Chrysanthos and P.G. Pelopides, *Theoretikon mega tes mousikes*. En Tergeste : ek tes typographias Michael Vais (Michele Weis), 1832.
- [15] D. Delviniotis and G. Kouroupetroglou, "DAMASKINOS: The Prototype Corpus of Greek Orthodox Ecclesiastical Chant Voices", in *Proc. Int. Conf. Crossroads | Greece as an Intercultural Pole of Musical Thought and Creativity*, Thessaloniki, Greece, pp.1-14, 2011.
- [16] G. Kouroupetroglou, D. Delviniotis, and G. Chrysochoidis, "DAMASKINOS: The Model Tagged Acoustic Corpus of Byzantine Ecclesiastic

- Chant Voices,” in *Proc. ACOUSTICS Conf.*, Heraclion, Greece, 2006, pp. 68–76.
- [17] G. Chrysochoidis, D. Delviniotis, and G. Kouroupetroglou, “A semi-automated tagging methodology for Orthodox Ecclesiastic Chant Acoustic corpora,” in *Proceedings of the Int. Conf. Sound and Music Computing*, Lefkada, Greece, 2007, pp. 126–133.
- [18] P. Boersma, “Praat, a system for doing phonetics by computer,” *Glott International*, vol. 5, no. 9/10, pp. 341–345, 2002.
- [19] P. Boersma, “Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound,” in *IFA Proceedings 17*, 1993, pp. 97–110.
- [20] N. Andersen, “On the calculation of filter coefficients for maximum entropy spectral analysis,” *Geophysics*, vol. 39, no. 1, pp. 69–72, Feb. 1974.
- [21] D. G. Childers, *Modern spectrum analysis*. IEEE Press: sole worldwide distributor (exclusive of IEEE), Wiley, pp. 252–255, 1978.
- [22] P. Boersma and G. Kovacic, “Spectral characteristics of three styles of Croatian folk singing,” *The Journal of the Acoustical Society of America*, vol. 119, no. 3, p. 1805, 2006.
- [23] P. Escudero, P. Boersma, A. S. Rauber, and R. A. H. Bion, “A cross-dialect acoustic description of vowels: Brazilian and European Portuguese,” *The Journal of the Acoustical Society of America*, vol. 126, no. 3, p. 1379, 2009.
- [24] T. Wempe and P. Boersma, “The interactive design of an F0-related spectral analyser,” in *Proc. 15th ICPhS*, 2003, pp. 343–346.
- [25] I. Jemaa, O. Rekhis, K. Ouni, and Y. Laprie, “An evaluation of formant tracking methods on an Arabic database,” in *10th Annual Conference of the International Speech Communication Association-INTERSPEECH 2009*, 2009.
- [26] H. Boril and P. Pollák, “Direct time domain fundamental frequency estimation of speech in noisy conditions,” in *Proceedings of EUSIPCO 2004 (European Signal Processing Conference)*, 2004, vol. 1, pp. 1003–1006.
- [27] D. G. Silva, L. C. Oliveira, and M. Andrea, “Jitter Estimation Algorithms for Detection of Pathological Voices,” *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, p. 567875, 2009.
- [28] K. Bunton and B.H. Story, “A test of formant frequency analyses with simulated child-like vowels,” Presented at the 161st Acoustical Society Meeting, 129(4), pt. 2 of 2, 2011.
- [29] D. Delviniotis, “Byzantine chant analysis by using signal processing techniques,” PhD Theses, National and Kapodistrian University of Athens, Department of Informatics and Telecommunications, 2002.
- [30] L. Dmitriev and A. Kiselev, “Relationship between the Formant Structure of Different Types of Singing Voices and the Dimensions of Supraglottic Cavities,” *Folia Phoniatrica et Logopaedica*, vol. 31, no. 4, pp. 238–241, 1979.
- [31] T.F. Cleveland, “Acoustic properties of voice timbre types and their influence on voice classification,” *J. of the Acoustical Society of America*, vol. 61, p. 1622, 1977.
- [32] N. Henrich, J. Smith, and J. Wolfe, “Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones,” *J. of the Acoustical Society of America*, vol. 129, p. 1024, 2011.

Glitch Free FM Vocal Synthesis

Chris Chafe

Center for Computer Research in Music and Acoustics, Stanford University
cc@ccrma.stanford.edu

ABSTRACT

Frequency Modulation (FM) and other audio rate non-linear modulation techniques like Waveshaping Digital Synthesis, Amplitude Modulation (AM) and their variants are well-known techniques for generating complex sound spectra. Kleimola [1] provides a comprehensive and up-to-date description of the entire family. One shared trait is that synthesizing vocal sounds and other harmonically-structured sounds comprised of formants can be problematic because of an obstacle which causes distortions when intensifying time-varying controls.

Large deflections of pitch or phoneme parameters cause jumps in the required integer approximations of formant center frequencies. Trying to imitate human vocal behavior with its often wide prosodic and expressive excursions causes audible clicks. A partial solution lay buried in some code from the 80's. This, combined with a phase-synchronous oscillator bank described in Lazzarini and Timoney [2] produces uniform harmonic components which ensure artifact-free, exact formant spectra even under the most extreme dynamic conditions. The paper revisits singing and speech synthesis using the classic FM single modulator / multiple-carrier structure pioneered by Chowning [3]. The revised method is implemented in Faust and is as efficient as its predecessor technique. Dynamic controls arrive multiplexed via an audio rate "articulation stream" which interfaces conveniently with sample-synchronous algorithms written in Chuck. FM for singing synthesis can now be "abused" with radical time-varying controls. It also has potential as an efficient means for low-bandwidth analysis – resynthesis speech coding. Applications of the technique for sonification and in concert music are described.

1. INTRODUCTION

Synthesis of singing voice by computer has a history which begins in the very first years of computer music. The song *Daisy Bell (Bicycle Built for Two)* was sung by a computer in 1961 in an arrangement by Max Mathews and Joan Miller with vocal synthesis by John Kelly and Carol Lochbaum when the Bell Telephone Laboratories experiments with digital music synthesis were only 4 years old. It was an early case of analysis – resynthesis speech coding providing a means for singing synthesis. Over the decades,

most synthesis technique have been applied to emulate the singing voice (additive, subtractive, physical model, FOF, etc.). The quest continues more than fifty years later with composers attracted to vocal synthesizers like Yamaha's Vocaloid¹ where they can explore a fascination with musical personalities of singers which never existed. This paper joins a thread which began with Chowning's work in the late 70's, early 80's involving FM for vocal synthesis and which has been virtually languishing since it's early use in a few musical works.

John Chowning's FM singing voice method was first described in his 1980 article [3] prior to completing *Phonē* at IRCAM (1981). The multi-channel tape piece features a wide variety of singing voices and morphing of vocal timbres with other FM-generated timbres such as gongs. The technique creates multiple formants with independent tunings using multiple carriers and a shared modulator. Two formants are used for his version of a soprano voice "eee" and three formants for his spectrally-rich basso *profondissimo*. A later version adds a third formant to the soprano model in a synthesis of the vowel "ahh" [4]. Pitch vibrato which causes synchronous spectral modulation is especially effective and Chowning has often demonstrated how crucial this is to rendering vowels convincingly. "It is striking that the tone only fuses and becomes a unitary percept with the addition of the pitch fluctuations, thus *spectral envelope does not make a voice!*" [3].

The method has an inherent shortcoming which limits the amount of vibrato excursion and limits phoneme transitions to nearby phonemes. Beyond these limits noticeable artifacts occur which are caused by discrete shifts of formant center frequency. Discontinuities are perceived as clicks and result from integer shifts in the carrier to modulator ratio $c : m$ which are required in order to track a desired formant center frequency f_c for a given pitch f_p . The modulating oscillator is always set to f_p , so $m = 1.0$. The carrier ratio c is an *integer approximation* and quantization of the actual *real* ratio f_c/f_p .

Formant synthesis with FM is essentially contradictory to the physics. The harmonic nature of voiced sound allows only harmonic number ratios $c \in \mathbb{N}_{\geq 1}$ for the carrier. Where physical sound production is an excitation – resonance mechanism with independent tuning of both elements, FM models can only approximate the resonance frequencies of the latter when constrained to produce harmonic spectra. The inherent problem is that these approximations are discontinuous in frequency. In practice, this

Copyright: ©2013 Chris Chafe et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ Vocaloid3 uses a triphone frequency-domain concatenative synthesis engine.

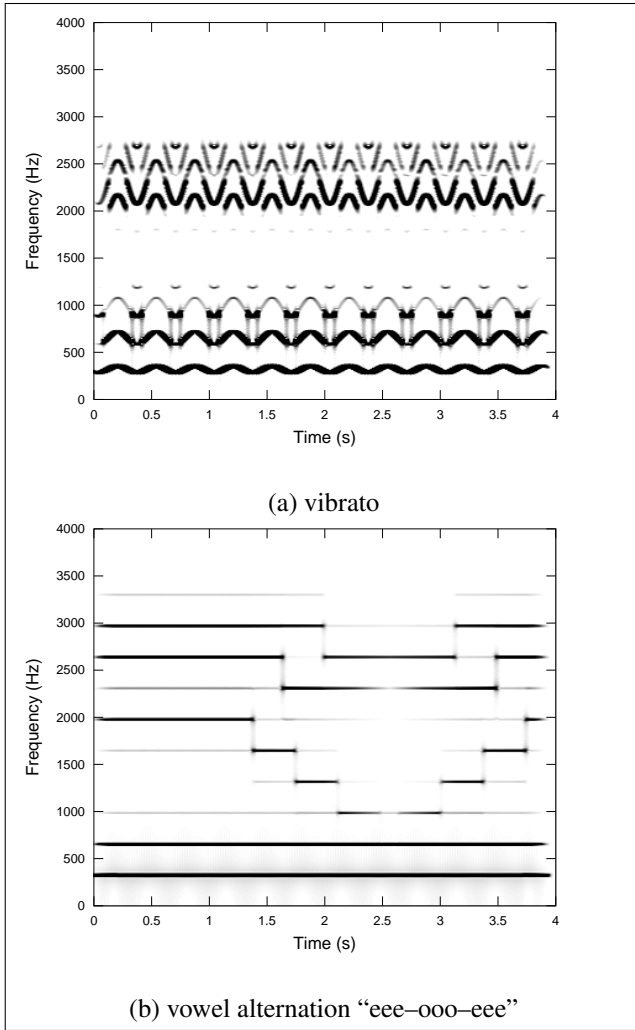


Figure 1. Clicks always occur when transitions to a new formant center frequency f_c forces a carrier oscillator to change its harmonic ratio. FM vocal formants use a $c : m$ ratio where $c \in \mathbb{N}_{\geq 1}$ and $m = f_p$, where f_p is the desired pitch.

severely limits the amount of pitch skew not only constraining vibrato, but also portamento and glissando to small ranges. In the method's original form, it is impossible to shift ratios without causing glitches like those shown in the spectrograms of Fig. 1.

Finding a solution became necessary in order to use FM as the synthesis engine for a sonification project involving brain signals. FM singing offers advantages for this body of work which attempts to fashion a singing choir direct from the mind. The goal isn't what one probably first imagines e.g., an ensemble of mind-controlled voices. Instead, this is a technology for auditory display of the rapid fluctuations of EEG and electrocorticography (eCog) recordings. Singing voice synthesis has its attractions in that it can allude to imagery of "inner voices" but it is also particularly apt because of the ease with which listeners lock on to patterns of phonemic and other voice-like timbral transitions. The range of data encountered in brain recordings (from quiescent to seizure) and a desire to have a very flexible mapping strategy have been moti-

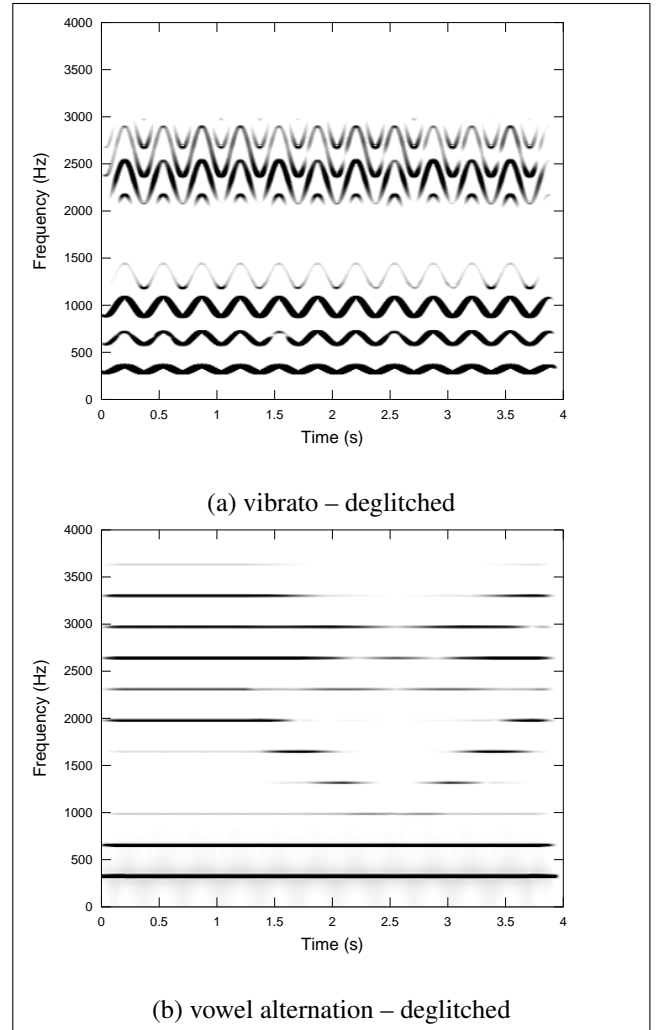


Figure 2. Result of applying the solution adopted from Le Brun to the synthesis shown in Fig.1.

variations for the present investigation into solving the discontinuity problem. The completed work will reach the public as a gallery installation (exploring recorded data) and as a medical monitoring device for detecting seizures (with the singing voice controlled directly from electrodes in real time).

2. EARLY SOLUTION

Marc Le Brun described digital waveshaping synthesis in 1979 as a generalized paradigm for non-linear modulation synthesis [5]. FM is a special case of waveshaping synthesis and in devising a way to avoid the discontinuity problem for waveshaping, Le Brun also solved it for the FM case. Le Brun's solution remains unpublished (until now) with one exception: Bill Schottstaedt has preserved it as a synthesis instrument in the Common Lisp Music (CLM) project [6]. From the code comment, "**Vox**, an elaborate multi-carrier FM instrument is the voice instrument written by Marc Le Brun, used in *Colony* and other pieces."

Vox avoids the integer ratio shift discontinuities by implementing a cross-fading solution. Two carriers corresponding to even and odd harmonic numbers are assigned

to each formant “bracketing” the true formant center frequency. Their assignments are made from the two nearest harmonics $f_{lower} = \lfloor f_c/f_p \rfloor$ while the other is the nearest upper harmonic $f_{upper} = \lceil f_c/f_p \rceil$. The assignment of harmonics to individual oscillators is dynamic and depends on whether they are even numbered or odd numbered. When an oscillator is required to change its harmonic number the other will be approaching the actual target f_c/f_p . The two carrier oscillators’ amplitudes sum to unity in a mixture whose gains are complementary and linearly determined by proximity to the target. The key feature which makes this work is that it ensures that the oscillator which is having its frequency changed will be muted. As a nice side-effect, it also sharpens the accuracy with which the target formant center frequency is being synthesized.

Le Brun’s paper describes “a unified conceptual framework for a number of nonlinear techniques, including frequency-modulation synthesis. Both the theory and practice of the method are developed fairly extensively, beginning with simple but useful forms and proceeding to more complex and richer variations.” The cross-fade solution however only existed in code from the same era. To detail the historical record precisely, its first implementation was written in the MUS10 compiler (Stanford Artificial Intelligence Laboratory’s version of Bell Laboratories’ MusicN compilers). Later, it was ported to CLM as **pqw-vox** a “translation from MUS10 of MLB’s waveshaping voice instrument (using phase quadrature waveshaping).” Today, both **pqw-vox** and the FM version **vox** can be found translated to Scheme in Schottstaedt’s Snd project [6] as instruments defined the file *clm-ins.scm*.²

The cross-fade solution has not been incorporated in common FM vocal synthesis implementations. Today, the most notable is the **FMVoice** instrument included in the Synthesis Tool Kit (STK) [7]. The class *FMvoices.cpp* can be freely downloaded as part of STK’s source code and has been ported to various platforms e.g., Chuck [8] and Max / MSP / PeRColate [9]. In porting this class to Faust [10] and dealing with the discontinuity problem, I subsequently “rediscovered” for myself Le Brun’s early solution. The same cross-fade solution also appears in Lazzarini and Timoney [2].

3. NEW PROBLEM

Lazzarini and Timoney also describe a method for generating formants with phase-synchronous oscillators. Its importance will become apparent. After adopting Le Brun’s code in my own work and verifying that the discontinuity’s clicks were gone (Fig. 2) I noticed that the fix introduced a new problem. This was again an audible artifact plaguing vibrato. Not clicks, but a new kind of artifact. Where perfectly periodic vibrato should elicit perfectly periodic spectral modulation it in fact, didn’t. From one vibrato cycle to the next an overlaid pattern of spectral modulation is heard. The problem arises from phase mismatches in the pair of formants (even and odd harmonic numbers) being

mixed for each formant. These are the pair being cross-faded to combat the clicks in what I will now label as the “first-order problem.”

The cross-fade technique assumes that the energy of all coincident pairs of spectral lines will sum arithmetically. However, this assumption does not take phase into account. A “second-order problem” is caused by phase interference of coincident spectral lines. These are the spectral lines (carrier and sideband frequencies) of the two overlapping formant generators which fill out the spectral envelope of the formant. They are identical spectra which are shifted relative to one another by one harmonic number. All phases are generated relative to their respective carrier oscillators rather than to the ensemble of frequencies as a whole. And without phase-synchronous oscillators, these phases are arbitrary in time since they are independently determined by control changes.

As discussed in Sec. 2 the first-order artifact is only apparent under changing conditions of pitch and phoneme target. Similarly, the second-order effect may remain unnoticed under steady-state conditions. With no change to carrier frequencies, carrier phases will also be constant and so will the resultant spectral mix. Nevertheless, interference between unrelated sets of phases can have an effect which alters the static spectral envelope and is perceived as a quality mismatch away from a target steady-state vowel. The problem becomes more apparent when carrier frequencies are being shifted dynamically, especially if these changes are happening periodically. Vibrato is a good way to emphasize the problem. Spectral distortions which may be imperceptible under other conditions are easier to hear with control changes which are repeating. The ear can pick out the distortion effect as a kind of spectral “isorhythm” or aliased pattern which is superimposed. Vibrato with a given period will generate a longer-period pattern of spectral modulation as seen in Fig.3. If you study the regions around 700 Hz and 1200 Hz, you will notice patterns in which phase-related Moiré fringing is inscribed on the amplitudes of the harmonics.

3.1 Minimizing Fringing

An initial attempt to minimize the audible effect of phase fringing is worth mentioning even though it isn’t ultimately the solution being adopted. It exploits the fact that phase interference is most notable when the cross-fade mix of the two carriers approaches equal portions (when interaction will be greatest). This is the point at which the carriers are equidistant from the target center frequency. Conversely, the least interference occurs when one of them is closest to the target and the other is essentially muted. Taking advantage of this proportion where one oscillator dominates, by expanding its time in the (vibrato-related) duty cycle, is one way to minimize fringing.

In listening tests, it was found that the cross-fade ramp can be made non-linear and still mask the first-order discontinuity perfectly. By using a power law for the ramp slope, fringing is reduced by causing less time to be spent in the portion of the duty cycle with the problematic mix ratio. Initial experiments under periodic vibrato condi-

² One caution: some implementation versions belonging to this family have mistakenly labeled carrier oscillators as “modulators” and the reverse: their “carrier” is actually the modulator.

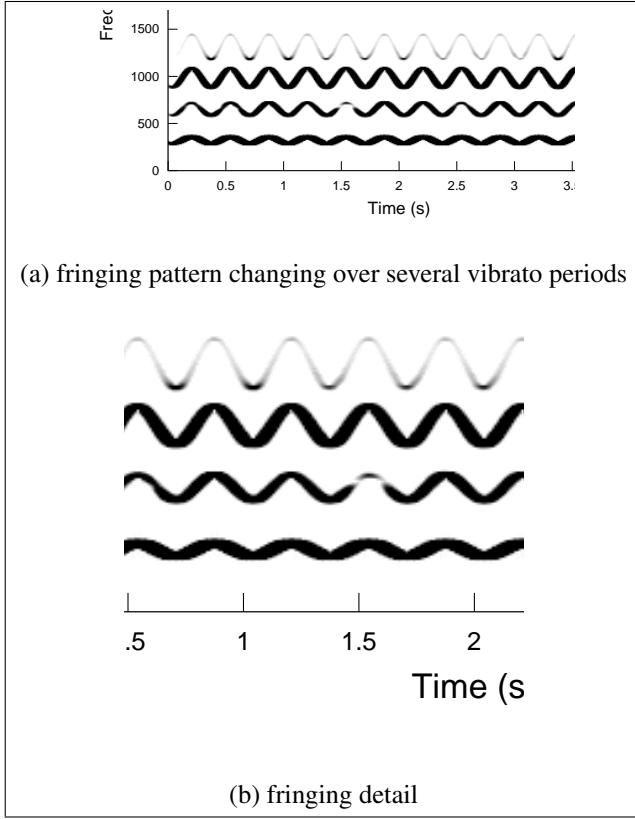


Figure 3. Phase-related Moiré fringing, zoomed in from Fig. 2(a). When either of the two cross-faded carriers resets its frequency because the harmonic ratio needs to shift, it also resets its phase with respect to the other carrier. The result is a Moiré fringing effect visible in spectrograms.

tions indicated that even a very significant exponent can be used e.g., $f(x) = x^7$ and still mute the first-order artifact smoothly enough to avoid a click. This greatly reduces the time spent in “phase interference mode” and all but eliminates the audible effect of the second-order problem.

This fix has its drawbacks. Some fringing still remains during the portion of the duty cycle where the cross-fade mix briefly crosses through the equal portion region. More significantly though, is that altering the temporal dynamics of the mix distorts the mix away from the target i.e., away from the mix which best approximates the intended formant center frequency.

3.2 A Better Way

The root of the problem is in the use of independent oscillators. The way around this is to employ a bank of linked oscillators such as the phase-synchronous oscillators described in [2]. Here, a single phasor is shared by the modulator and all carriers. In the present implementation the bank can be constructed with any number of harmonic outputs and all will be tapped off of a single common phasor.

The familiar sampled sinusoid generates the fundamental (pitch) frequency signal for the bank to be constructed:

$$x(t) = A \sin(\omega t + \phi) \quad (1)$$

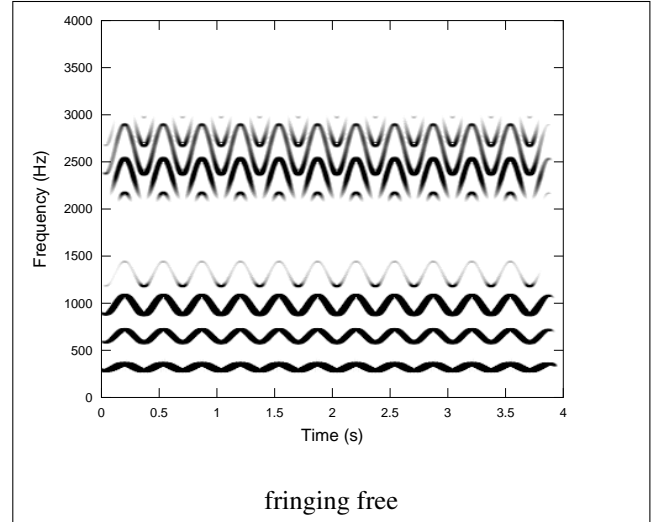


Figure 4. Re-rendering the vibrato example from 2(a) with phase-synchronous oscillators eliminates the fringing artifact.

where A is oscillator amplitude and ω in rad/s is calculated as $2\pi f$, where f is frequency.

Expressed in pseudo-code, Eq. 1 can be implemented with the modulo function:

```
w = f / SR
mp = 0.0
for n = 0 to N
    y[n] = a * sin(2pi * mp)
    mp = (mp + w) mod 1.0
end
```

The constant SR is the sample rate and the variable mp is the fundamental’s instantaneous phase.

The key to the next step is sharing the instantaneous phase mp with any other oscillators, where o specifies oscillator number, cp_o its instantaneous phase and h_o its harmonic number:

```
cp[o] = (h[o] * mp) mod 1.0
y[o][n] = a[o] * sin(2pi * cp[o])
```

and since we’re interested in doing FM

```
m[o] = y[n] * i[o]
cp[o] = (h[o] * mp) mod 1.0
y[o][n] = a[o] * sin(2pi * cp[o] +
    m[o])
```

The above pseudo-code implements one simple FM pair consisting of an independent carrier and shared modulator which produces a formant centered at harmonic h of pitch frequency f with modulation index i . The latter coefficient determines formant bandwidth and is typically used in a low range (< 2.0). In practice, a bank of six (or more) carrier oscillators of this kind will be used to generate a vocal sound. These will create phonemes of 3 (or more) formants represented by a time-varying distribution of h , a , and i coefficients.

The completed glitch-free method consists of Chowning FM singing voice + Le Brun cross-fade algorithm (from Sec. 2) + Lazzarini phase-synchronous oscillator bank (from Sec. 3). Fig.4 displays a spectrogram of vibrato rendered using the fully-realized solution (Faust and Chuck code to generate the example are included in the program appendices following). Classic phoneme table synthesis using Chowning's method can now be extended to arbitrary dynamic behavior.

4. APPLICATIONS

The singing voice technique has been used in three projects. Sound examples for each of these can be found online [11].

4.1 Converting eCog Signals to Music

Electrocorticography (eCog) registers brain electrical activity directly from inside the skull. Using sensors placed in regions suspected of giving rise to epilepsy, eCog arrays provide precise diagnostic monitoring as well as signals of great importance for studying the brain itself. In a current sonification project, the data is sung by a digital chorus. Each singer is a vocal simulation synthesized by the present technique. Where arrays have been implanted for therapeutic reasons, a large number of sensors is available (> 50) and the chorus can be made equally large. The aim of the work is to create a music directly from these sensors.

A correspondence exists between the temporal structures of music and the dynamics of brain activity monitored by eCog. Musical time has its notes, rhythms, phrases and epochal structures. Brain signals are marked by structures on the same time scales. Translation to music requires no modification of time base. In fact, the present approach avoids "re-composing" or altering the data in any way. The individuals who have contributed data to this project share our interest in discovering the potential of music as a different, new way of comprehending the complexity of brain dynamics.

Seizures we have listened to have a characteristic progression. They arise with a light, fast modulation "aura" not unlike a super-fast vibrato or tremolo which gives way to a nearly regular strong march of pulses. Multiple trains of pulses play against each other polyrhythmically. This crescendoes to an almost unbearable climax, the apex of the seizure. When it seems impossible for it to grow further, there is an abrupt cessation revealing a state of nothing. The paroxysm has switched off and the music is a quiet, calm, sustained chord. Motion is regained after this repose, but it is in a new world. Long, slow, undulations characterize the postictal phase whose affect is troubling, almost nauseous. Typically, this can last 45 minutes until normal brain activity is regained.

The method for translating brain signals into music is to have them modulate synthesized tones. The choice of singing synthesis makes an aesthetic connection to the "human-ness" of the data. Our chorus of eCog channels "performs" via modulations of pitch, loudness, vocal qualities and spatial location.

4.2 Speech synthesis

Speech synthesis, with its widely varying pitch and phoneme transitions, provides a good "real-life" test of the formant synthesis technique. The test has been created with a "toy" analysis – resynthesis platform driving synthesis from digitized singing and speech. The formant tracking analyzer is written in Chuck and the formant synthesizer is a Chuck UGen (unit generator) written in Faust. The analysis portion is FFT-based and uses a relatively long (4096 sample) window for formant accuracy (at 48 kHz sample rate). An example speech input fragment and the method's resynthesized output are compared in the spectrograms of Fig.5. Signal coding in this version consists simply of recording formant parameter updates which are relatively sparse (and could be greatly optimized). The results are promising for developing this into an FM-based speech coder – the example consists of two different speakers in a heated, emotional dialog. Their voices and identities are preserved, as is their expressive prosody and intelligibility. The analysis tracks populations of short-lived formants which in the example are limited to 4 at a time (using 9 oscillators total).

4.3 Near the Inner Ear

The formant generation technique was applied in a recent composition for orchestra, computer music and video premiered in 2013 by the Stanford Symphony Orchestra, Jindon Cai, conductor. The orchestral score by Dohi Moon was recorded and analyzed with a formant tracking algorithm. Analyzed formant tracks were also obtained from a recording of the first movement of Beethoven's Ninth Symphony (the two pieces were performed together for the conclusion a full Beethoven cycle). *Near the Inner Ear* incorporated resynthesized clips from both analysis data sets.

The composition used formant-based resynthesis to generate a novel timbral identity whose music is tied to the orchestral writing but whose comportment is different enough to constitute a kind of musical "alter ego." The work opens with a 90 s section in which an antiphonal exchange exploits such contrasts. The resynthesis instrument returns at various moments during the piece with musical gestures which reinforce the accompanying video composition by John Scott.

5. CONCLUSION

One goal of the analysis – resynthesis system going forward is to create a large database of acquired vocal sounds in order to structure more complex phoneme-based singing synthesis from a vastly expanded table. Applications described above have demonstrated a potential to create rich reservoirs of articulations and timbral identities. The hope is tap into greater timbral variety for the sonification work. Possibly also to acquire listeners' own voice traits for EEG-driven synthesis.

Resynthesis in the service of music projects like *Near the Inner Ear* can also take liberties with acoustic structures in order to create new instrumental identities. In *Near the Inner Ear*, the formant tracker operating on orchestral sound

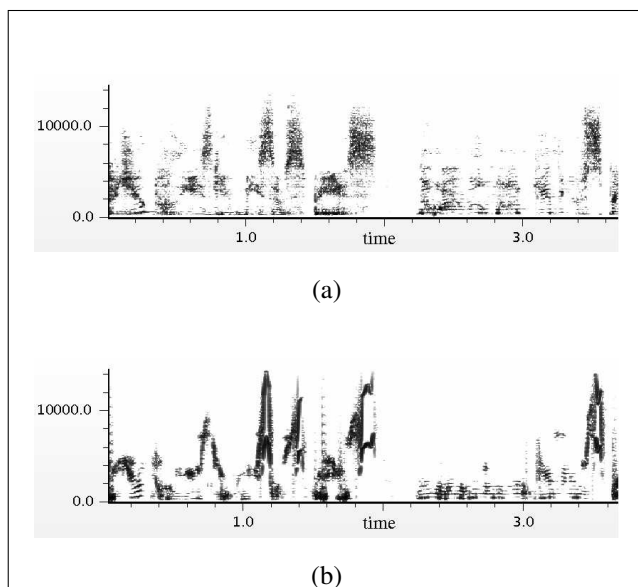


Figure 5. Analysis – resynthesis of dialog: (a) is the input from an argument between a teenage daughter and her mother, “can’t you please give me some space,” and “no, I will not give you some space.”, (b) FM resynthesis

inferred formants where no model would predict them to exist. Experimenting with resynthesis, it was found that often the tracker would emphasize pitches of inner voices in the analyzed recordings. Rather than impose an f_0 fundamental pitch, formant frequencies themselves were allowed to become the f_0 . The remaining formants would then create a kind of “instrumental singing” whose resonances mapped to the pitch structures from the original recordings.

The present synthesis method can be used for non-vocal sounds whose acoustic structures are also represented with formant-like resonances. Horner has explored timbre matching for a sampled trumpet using a genetic algorithm to find suitable FM formant parameters [12].

The improvements to FM vocal synthesis detailed in this paper can be extended to other audio rate modulation schemes, in particular those which also employ single modulator / multiple carrier structures. A glitch-free AM vocal synthesis “cousin” has also been implemented in Faust. AM has the advantage of simplicity in prediction of dynamic sideband behavior (AM sidebands are free of the Bessel function which determines FM sidebands).

Acknowledgments

Many thanks to John Chowning for his inventions and encouragement, musical and technical. Bill Schottstaedt continues to passage into the future comprehensive sets of synthesis instruments and analysis tools. His Snd project preserves and provides essential computer music algorithms without which much of the present work would not have been possible.

6. REFERENCES

- [1] J. Kleimola, “Nonlinear abstract sound synthesis algorithms,” Ph.D. dissertation, Aalto University, Helsinki, Finland, 2013.
- [2] V. Lazzarini and J. Timoney, “Theory and practice of modified frequency modulation synthesis,” *J. of Audio Eng. Soc.*, vol. 58, no. 6, pp. 459–471, 2010.
- [3] J. Chowning, “Computer synthesis of the singing voice,” in *Sound Generation in Winds, Strings, Computers*, J. Sundberg, Ed. Royal Swedish Academy of Music, 1980, pp. 4–13.
- [4] —, “Frequency modulation synthesis of the singing voice,” in *Current Directions in Computer Music Research*, M. Mathews and J. Pierce, Eds. MIT Press, 1989, pp. 57–64.
- [5] M. L. Brun, “Digital waveshaping synthesis,” *J. of Audio Eng. Soc.*, vol. 27, no. 4, pp. 250–266, 1979.
- [6] “Scheme, Ruby, and Forth Functions included with Snd,” last viewed 29 Mar. 2013. [Online]. Available: <https://ccrma.stanford.edu/software/snd/snd/sndscm.html>
- [7] “FMVoices Class Reference, in The Synthesis ToolKit in C++,” last viewed 29 Mar. 2013. [Online]. Available: <https://ccrma.stanford.edu/software/stk/>
- [8] “Chuck : Strongly-timed, concurrent, and on-the-fly audio programming language,” last viewed 29 Mar. 2013. [Online]. Available: <http://chuck.cs.princeton.edu/>
- [9] “PerColate, A collection of synthesis, signal processing, and image processing objects for Max/MSP,” last viewed 29 Mar. 2013. [Online]. Available: <http://music.columbia.edu/percolate/>
- [10] “FAUST (Functional Audio Stream),” last viewed 29 Mar. 2013. [Online]. Available: <http://faust.grame.fr/>
- [11] “Supporting online materials.” [Online]. Available: <http://ccrma.stanford.edu/~cc/vox/smac2013som/>
- [12] J. B. A. Horner and L. Haken, “Machine Tongues XVI: Genetic algorithms and their application to FM matching synthesis,” *Computer Music J.*, vol. 17, no. 4, pp. 17–29, 1993.

7. PROGRAM A

The Faust program, FMVox.dsp [11], implements an FM-voice algorithm with four formants consisting of uniform phase table-lookup harmonic oscillators. A multiplexed audio rate coefficient stream drives the synthesis. For each formant a demuxer is included which extracts its coefficients from the stream. Formants are instantiated in a parallel composition by the Faust process which outputs the sum of their signals. The resulting unit generator compiled by Faust can have a scalable number of formants. A texture

of multiple voices can be created from multiple, independent voice control streams which flow to independent unit generators. Using this architecture a choir of voices can be distributed across multiple sample-synchronous threads and / or multiple cores.

```
declare name "FMVox";
import("filter.lib");
ts = 1 << 16;
fs = float(ts);
ts1 = ts+1;
ct = (+ (1) ~ _ ) - 1;
fct = float(ct);
sc = fct*(2*PI)/fs:sin;
sm = fct*(2*PI)/fs:sin/(2*PI);
dec(x) = x-floor(x);
pha(f) = f/float(SR):(+:dec) ~ _;
tbl(t,p)= s1+dec(f)*(s2-s1)
with {
    f = p*fs;
    i = int(f);
    s1 = rdtbl(tbl,t,i);
    s2 = rdtbl(tbl,t,i+1); };

fupho(f0,a,b,c) = (even+odd):*(a)
with {
    cf = c/f0;
    ci = floor(cf);
    cil = ci+1;
    isEven= if((fmod(ci,2)<1),1,0);
    ef = if(isEven,ci,cil);
    of = if(isEven,cil,ci);
    frac = cf-ci;
    comp = 1-frac;
    oa = if(isEven,frac,comp);
    ea = if(isEven,comp,frac);
    ph = pha(f0);
    m = tbl(sm,ph):*(b);
    even = ea:*(tbl(sc,(dec(ef*ph))+m));
    odd = oa:*(tbl(sc,(dec(of*ph))+m));};
frame(c) = (w ~ _ )
with {
    rst(y)= if(c,-y,1);
    w(x) = x+rst(x); };
demux(i,ctr,x) = coef
with {
    trig = (ctr==i);
    coef = (*(1-trig)+x*trig) ~ _;};
formant(f_num,ctlStream) = fsig
with {
    ctr = frame(ctlStream<0);
    co(i) = demux(i,ctr,ctlStream);
    f0 = 1;
    a = f0+1+f_num*3;
    b = a+1;
    c = a+2;
    fsig = fupho(co(f0), co(a),
        co(b), co(c)); };
nf = 4;
process = _<:par(i,nf,formant(i)):>_;
```

The **fupho** function implements one formant (of uniform phase harmonic oscillators) which is free of glitches caused by dynamic behavior. Its inputs are the fundamental frequency, formant amplitude, bandwidth and center frequency (respectively, a, b, c). A **fupho** receives demuxed controls via its calling function **formant**.

8. PROGRAM B

FMVox is used in the Chuck program, **FMVoxVib.dsp** [11], to produce Fig.4. The example defines a master “shred” which executes for 4 s. It sets up a DSP graph in which one **FMVox** instance receives a sample-synchronous control stream and sends its output to the dac. The master shred “sporks” child shreds for vibrato and a multiplex control stream. The **data** float array holds “aaa” vowel coefficients for four formants described by their amplitude and center frequency. For this test code formant bandwidths are set the same globally.

```
Step stream => FMVox fmv => dac;
4 => int nFormants;
1::ms => dur updateRate;
SinOsc vibLFO => blackhole;
vibLFO.freq(3);
vibLFO.gain(0.1);
Std.mtof(64) => float p => float f0;
fun void vibrato() {while (true){
    ((vibLFO.last()+1.0)*p) => f0;
    1::ms => now;
}}

[ // "aaa"
  [ 349.0, 0.0],[ 918.0,-10.0],
  [2350.0,-17.0],[2731.0,-23.0]
] @=> float data[][];

fun void mux(float val) {
    stream.next(val);
    1::samp => now;
}

-1 => int startFrame;
95 => float db;
fun void muxStream() {
    updateRate-14::samp => dur padTime;
    while(true){
        padTime => now;
        mux(startFrame);
        mux(f0);
        for (0 => int f; f<nFormants; f++){
            mux(Math.dbtorms(db+data[f][1]));
            mux(0.2);
            mux(data[f][0]);
        }
    }
}

spork ~muxStream();
spork ~vibrato();
4::second => now;
```


Testing a new protocol to measure tuning response behaviour in solo voice ensemble singing

Helena Daffern

University of York

helena.daffern@york.ac.uk

Jude Brereton

University of York

jude.brereton@york.ac.uk

ABSTRACT

Tuning within choirs when performing *a capella* continues to be a subject of interest to researchers and practitioners in terms of theoretical preferred tuning systems and the instinctive tuning behaviour of choral singers. Although more studies are exploring the issues surrounding tuning within choirs, most research on tuning in choirs focuses on the mean fundamental frequency of tones, either through analysis of performance, or perceptual listening tests. This paper considers a specific protocol and tools for analysis to assess the tuning characteristics of individuals singing in a choral texture. The results illustrate the tuning features throughout single notes and the extent to which singers adapt their tuning to the surrounding chord in equal temperament and just tuning. To test the effectiveness of this new methodology, data collected in a previous experiment is reconsidered in which singers were recorded performing a specially written choral exercise where the other parts were synthesised and heard over headphones. Overall the analysis proved a successful and efficient process. The results of the analysis suggest that further research into the tuning tendencies of choral singers, particularly in terms of fundamental frequency trajectories would prove valuable as a research area. Modifications to the protocol are considered for an upcoming study, which is planned for next year, and a revised hypothesis for the tuning strategies in the set exercise is presented.

1. INTRODUCTION

Tuning behaviour in ensemble singing has long been a topic of interest to practitioners and researchers. Analysing choral singing is problematic because of its very nature of creating multi-layered sounds. From an analytical perspective, considering the vocal practice of individual singers within a chorus based on the acoustic waveform is problematic, due to the need to separate the signal to identify the contribution of each singer. Using close proximity microphones allows the singer of interest to be prevalent but will not eliminate bleed in the signal from the other performers and therefore isn't reliable for in depth acoustic analysis. Certain characteristics, such as fundamental frequency, can be isolated between singers in a choir by using electroglottography or

electrolaryngography [1] but this doesn't allow consideration of other acoustic factors of voice quality which could affect the perceived intonation. The influence of individual singers on each other when singing as a chorus is also an important factor to consider: when analysing vocal characteristics it is difficult to ascertain whether features are idiomatic of an individual's singing tendencies or reactive to their choral companions.

Although it is thought to have been employed by certain instruments since the sixteenth century, the standardized use of equal temperament in Western music performance is a relatively new convention established alongside the development of the modern piano and the compositional desire to exploit potential modulations across all keys within one piece of music [2]. The tuning decisions made by choirs when performing *a capella* have been a point of dispute in literature since the early writings of Zarlino in the 1560s and continues today [2, 3].

Professional choral performance practice often involves a close focus on intonation with conscious consideration of tuning systems and factors impacting on tuning such as vibrato and spacing. Whilst amateur choirs may be less aware of the principles of tuning systems, they still consider tuning a vital aspect of good choral singing. When writing about just tuning Mayer asserts; 'when realized, this discipline can produce that ethereal beauty of tone ... that distinguishes the really fine choir' [4]. Choirs are advised to practice without piano accompaniment so that they 'are able to sing in 'just' intonation based on intervals derived from the harmonics series' [5]. Common issues arising in choral training, not attributed to issues of temperament due perhaps to a lack of understanding, could still be the result of their cause: for example when concentrating on 'finding and correcting the low minor thirds and other faulty intervals' [6].

Academic opinion is divided as to the tuning preferred by *a capella* groups, with evidence of inclinations towards various tuning systems including equal temperament and just intonation. Based on the practice-led theories of a preference towards just intonation in choirs such as those given above [3,4,5], Howard produced a set of purpose written choral exercises to examine pitch drift and tested the prediction with a vocal quartet [7]. The results showed a tendency towards just rather than equal temperament as the exercise drifted very flat of equal temperament when performing the set exercises and these results were supported by a further experiment [1]. Analysing the tuning of the quartet singing "The First Nowell", a popular Christmas carol, rather than the be-

spoke exercises the quartet were still found to drift in pitch, although the verse of the hymn drifted sharp rather than flat [8]. A study involving a perceptual test with sixteen expert listeners conducted by Nordmark and Ternström found no preference for just intonation, with participants choosing thirds tuned closer to equal temperament [9], drawing on previous research which also found a preference for large major thirds in choral singing, although a close match of fifths and octaves to just intonation. Considering the tuning preferences of violinists, pianists and non-musicians, Loosen concluded that the concept of ‘accurate tuning’ is connected to musical experience rather than innate auditory function, noting that pianists tended to prefer equal temperament [10].

As mentioned above, many factors are thought to contribute to intonation in choral singing and attempts are being made to identify, isolate and analyse these variables many of which are touched upon by Ternström in his overview of research into choir acoustics [11]. Ternström and Sundberg found loudness to be connected to the precision of tuning, with deterioration at louder levels [12]. In the same study spectral factors were proven important to tuning accuracy with three-tone complex rather than two-tone complex reference tones significantly improving intonation. These findings support previous studies conducted by Ternström and Sundberg in which vibrato, vowels, and higher frequency partials were shown to contribute to tuning precision [13].

A recent experiment conducted by Howard et al. [14] considered the mean fundamental frequency of tones sung by singers performing Exercise 3 (see Figure 1) when tuning to a synthesised chorus, repeated with the chorus parts tuned in both just and equal temperament. These exercises were composed for the investigations in [1,8]. The use of synthesized parts rather than recording a live quartet removed the issue of inter-singer influence, as well as the acoustical issue of capturing individual voices without bleed from other singers. The purpose of the study was to ascertain whether singers are naturally more inclined towards just tuning or equal temperament: the hypothesis being that choral singers prefer just intonation based on the opinions of practitioners outlined above. The findings showed that the singers (all females singing soprano and alto) tended to tune sharp of the reference note when singing with the just tuned chorus and flat of the reference note when singing with equal temperament, although with higher deviations from the reference note in the just tempered condition.

Calculating the mean fundamental frequency value of a tone is a common method used to consider the tuning behaviour of singers in ensemble situations. Especially considering the assertion that perceptually, the mean fundamental frequency is the important pitch reference to a listener this seems a useful analysis [15]. However, it doesn’t take into account the ‘adapting’ that takes place when a singer first sings a note and then ‘tunes’ it to the surrounding sounds.

In the practice of choral singing the fast adaption of singers’ tuning is well known by singers and conductors. Studies have demonstrated both the sensitivity of singers to tuning and the rapid response of singers to fine tune notes [16]. The analysis of mean fundamental frequency

of tones, therefore, whilst demonstrating that singers tune whole notes to the surrounding tuning system, does not reveal the tuning of the note the singer is ‘expecting’ to sing. In other words the note they have established in their mind before singing may differ from the note they adapt to, albeit very quickly. Considering the tuning behaviour of each note will illustrate emerging patterns of adapting tuning between temperaments and therefore give an insight into the tuning systems that allow for most ‘ease’ of tuning.

The present study re-considers the data collected by Howard et al. [14], whilst testing a new analysis method to extract the features of the tuning characteristics throughout each tone to assess the tuning transitions in the two tuning conditions. A pilot analysis was conducted to test the methodology of recording individuals when singing with a synthesized choral exercise played in just tuning and equal temperament. To do this, tools were employed that had not been used in the previous paper including AMPACT [17] and Sonic Visualiser [18]. The effectiveness of using these analysis tools to assess temporal aspects as well as fundamental frequency trajectories are considered. The results are compared to the findings of the previous study to consider the effectiveness of this new protocol and the relevance of tuning features within tones when studying tuning characteristics in choirs. Finally, an amended protocol for analysing the tuning behaviour of individual singers in a choral setting is presented.

2. METHOD

2.1 Musical Exercise

The music the subjects performed was taken from a set of exercises written by Howard to explore pitch shift in choral music (Figure 1). This exercise has been used in several papers to illustrate expected pitch shift when employing just intonation rather than equal temperament [1,7,8]. A synthesised version of the exercise was created in PureData in both equal temperament and just intonation. The piece was tuned such that the note which is held across the beat-long rest in other parts was treated as the stable tone for tuning purposes, with each new chord being tuned to the frequency of the tied note or ‘reference frequency’. All tones were free of vibrato, consisted of a larynx source based on the LF model with three formants mapped to the vowel /a/ [14]. The introductory first bar, providing a reference to middle C (C4= 261.626Hz was always included whichever part the singer sang. The subjects listened to three parts of the synthesized exercise over headphones, singing the missing part.



Figure 1: Exercise sung by subjects tuned to just and equal temperament [14]

2.2 Experiment

Four female subjects took part in the experiment and sang both the soprano and alto parts of the exercise. The subjects were aged between 19 – 21 and were all studying music at undergraduate level with singing as their first study instrument. The recordings took place in the Rymer Auditorium in the Music Department at the University of York, which seats 150 and was designed as a listening space for reproduced sound and therefore has a very dry acoustic and low background noise level.

Subjects were recorded using a DPA 4060 head-worn microphone connected to a Zoom R4 digital recorder set to record in uncompressed PCM .wav format at 48 kHz and 16 bit resolution. The synthesised choral parts were heard over DT990 headphones and the subjects controlled the volume so that they could hear themselves and the recording at a ‘natural’ level. The subjects were warmed up and reported no vocal problems on the day of the recordings.

Each subject was given time to practise the exercise before being recorded. The subjects were recorded singing the exercise four times, twice with the equal temperament version and twice with just temperament. Subjects were not informed of the purpose of the experiment or that there were variations in the tuning in different versions. The versions were randomly ordered to the participants.

2.3 Analysis

AMPACT (Automatic Music Performance Analysis and Comparison Toolkit) is a toolkit developed by Devaney and Mandel [17,19] for the MATLAB programming environment. It automates the process of analysing musical performance data from recorded audio by using information available from a MIDI score [20]. Many blind onset detection algorithms struggle to correctly identify notes without percussive onsets, such as in vocal performance. However, the AMPACT toolkit has been shown to be a robust method of onset detection for sung notes by estimating note onsets and offsets through matching the recorded performance to a MIDI score [21].

The AMPACT toolkit makes an initial estimation of the temporal location of each of the sung notes in the audio file, which can first be checked by the user, and any errors corrected. Once the notes have been correctly segmented, fundamental frequency estimation is carried out in the toolkit by implementation of Kawahara and de Cheveigne’s YIN algorithm [22]

AMPACT allows various performance parameters to be analysed, such as inter-onset intervals between notes, tempo information, vibrato rate and depth. In the present study fundamental frequency values for each note were calculated and compared to the expected frequency values in the Just Tuning (JT) and Equal Temperament (ET) versions of the sung exercise.

In addition, fundamental frequency trajectories are characterised by Direct Cosine Transforms using a method which has been evaluated by Devaney et al. [21]

First the steady-state portion of the note is found and in order to minimize the effect of starting and ending points

in the portion, and vibrato, the signal is smoothed through the connection of the midpoints between the peaks and troughs. The Direct Cosine Transform of the steady state portion is then computed and approximated using the first three coefficients. The approximated fundamental frequency trajectory for each note in the exercise is plotted and a visual comparison made between each recorded exercise.

The aim of this study is in part to test the effectiveness of using an automated processing technique such as AMPACT in order to facilitate the analysis of sung data in the context of an experiment to investigate tuning in choral singing.

Each sung exercise is approximately 40 seconds long, and once the protocol for analysis had been established the analysis of each exercise took approximately 5 – 10 minutes depending on the number of errors in the initial MIDI alignment that needed to be corrected.

3. RESULTS

3.1 Analysis method reviewed

AMPACT [17] and Sonic Visualiser [18] were found overall to be effective as tools to analyse the voice recordings. Devaney states that AMPACT analysis takes ten times real-time length of recorded data [19]. The authors found this generally to be the case once the protocol was established in MATLAB, and a much quicker process than analysing the recordings in a programme such as PRAAT to then extract and manipulate the data by hand. However, the alignmentVisualiser process in AMPACT which automatically isolates individual notes in a recording based on the reference data of a MIDI file, was not entirely accurate, requiring careful edits by hand in another programme (in this case Sonic Visualiser) before continuing with the AMPACT analysis. In spite of this, overall the AMPACT analysis was found to be robust in terms of fundamental frequency extraction and approximated fundamental frequency trajectories when checked against long-hand analysis which also confirmed its time efficiency.

Allowing the singers to adjust the volume of the playback themselves resulted in the synthesised exercise being picked up on the recording for subject three despite being played through headphones. The data collected for subject three was therefore disregarded for analysis in AMPACT because the bleed from the synthesised tones was too prominent on the recordings for the fundamental frequency to be extracted accurately.

3.2 Fundamental frequency trajectories

Fundamental frequency trajectories were obtained for each rendition of the exercise for subjects 1, 2 and 4 using the method described above. Each subject sang both the alto and soprano parts four times, twice with the synthesized chorus programmed to JT and twice when programmed to ET. This provided six versions of the alto and soprano line in each tuning condition for analysis.

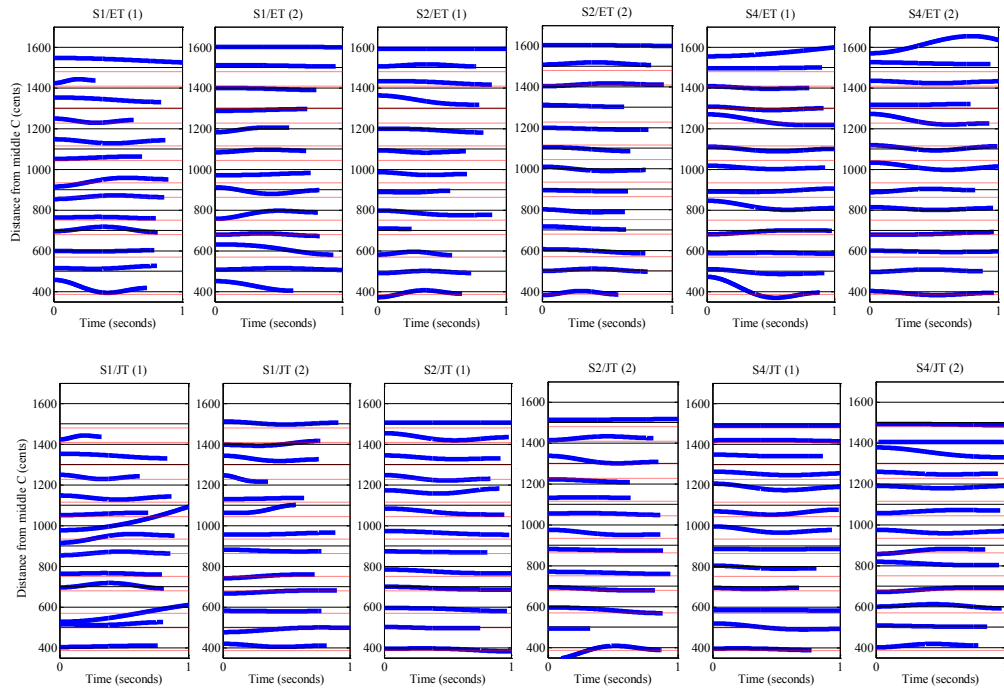


Figure 2. The fundamental frequency trajectories of each note of the soprano part sung by each singer, twice in the Equal Temperament condition (ET) and twice in the just tuned condition (JT). The expected frequency values for ET are shown by the solid black lines, the expected frequency values for JT are shown with red lines.

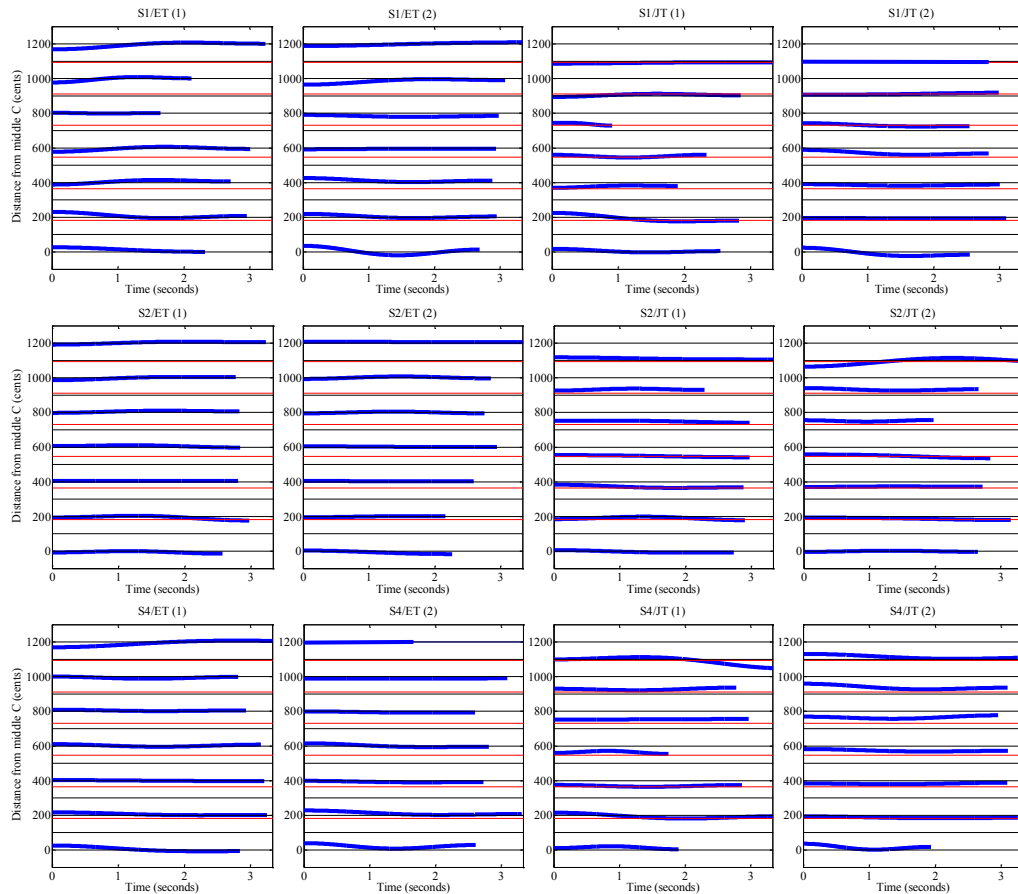


Figure 3. The fundamental frequency trajectories of each note of the alto part sung by each singer, twice in the Equal Temperament condition (ET) and twice in the Just Tuned condition (JT). The expected frequency values for ET are shown by the solid black lines, the expected frequency values for JT are shown with red lines.

Figures 2 and 3 show the fundamental frequency trajectories for each sung example with reference to the expected frequency in each tuning condition. The results are displayed in cents and represent the distance of each sung tone from middle C which was the original reference tone of the exercise.

All the singers (whether usually performing soprano or alto in choirs) produced more stable tones when singing alto, whilst the soprano examples have considerable variability of fundamental frequency trajectory within tones. In some tones in the soprano examples the variability can be up to a semi-tone displacement, for example, in the case of the first tone of subject four in equal temperament and the tenth tone of subject one also in ET, both of which start very sharp of their target frequency (Figure 2).

Overall, particularly in the alto recordings, there is a tendency for singers to begin the note sharp of the expected frequency and adapt closer throughout the tone. On many tones in which the singer adapts the tuning, there is a tendency to ‘overshoot’ the pitch correction before correcting once again towards the initial fundamental frequency, this is observed in the curves of the fundamental frequency trajectories such as tone one of the soprano for subject 4 in equal temperament.

3.2.1 JT compared to ET conditions

The fundamental frequency trajectories are also less consistent for most subjects singing both parts in the just tuned condition compared to equal temperament. This supports the findings of the previous study which found a larger standard deviation in the just tuned condition when analysing the mean fundamental frequency of tones.

3.2.2 Cross-subject patterns

Certain patterns emerge across specific notes as sung by all subjects. Excluding the JT versions for subject one, the seventh note sung in the soprano line is consistently adapted downwards with no ‘overshoot’ by all subjects in both tuning conditions, usually after beginning the note sharp of the expected value (see, for example, subject four, Figure 2).

The first note in both the soprano and alto often displays the most variability of all the tones in the exercise for all subjects. In the case of subject four the first note sung in ET appears to be an anomalous result as in all other examples of that note for that subject the trajectory is stable. For other subjects, however, a second attempt at each condition doesn’t necessarily stabilise the trajectory of the first note, or indeed the tuning of any of the notes. In the alto recordings the second attempt is often the least stable trajectory for the first note, implying it isn’t an issue of practise or familiarity with the exercise.

Whilst certain notes of the exercise display noticeable instabilities, others are characteristically accurate and stable. The second note of the soprano part is relatively stable and accurate across all subjects compared to the trajectories of other tones. It is worth noting that the tuning predictions for this note in JT and ET are almost identical and so the ‘choice’ of tuning for that note is reduced.

In ET the fourth note in the alto is highly accurate and stable for all subjects while the same note in the just tuned condition for subjects one and four is less stable and less accurate to the predicted fundamental frequency, tending to be sharp and beginning closer to the equal temperament prediction.

4. DISCUSSION

The results of this short study reveal that the use of AMPACT analysis in MATLAB is a robust and effective tool in considering tuning behaviour in ensemble singing. The results of the analysis also reveal that singers adapt to their tuning environment and further experiments will be worthwhile using the musical exercise and an adapted recording protocol.

4.1 JT compared to ET

The instability in the fundamental frequency trajectories in the JT condition compared to the ET condition suggests that the singers are more ‘comfortable’ or find it easier to tune in equal temperament. The standard deviation results for the mean fundamental frequency values analysed in the previous study supports this theory. However, there are certain factors that could be contributing to the instability in the JT trajectories that do not represent the natural tuning of choirs in specific temperaments, including the method used to predict the JT frequencies, a possibility which is discussed in more detail below.

4.2 Voice parts

The tendency observed for trajectories to be less stable for the soprano recordings than the alto could be attributed to a number of factors, including the different musical conditions of the two parts. For example, the soprano part involves notes of just one beat in length compared to the sustained three beat tones held in the alto part. The relative stability of the final tone, a four-beat note, compared to the other tones in the soprano examples for all but subject four in the equal temperament condition supports this theory. Another consideration is the musical interval being tuned in each part. The alto always moves onto the root note whilst the soprano alternates between providing the third and the root of the chord. The contrast in variation of intervals across the exercise between the two parts as well as the intervals themselves could contribute to more difficulties tuning the soprano line.

The features of specific notes such as note seven in the soprano part could be indicative of this issue, the seventh note being tuned as a sixth if tuning to the alto reference pitch but a third if tuning to the root of the chord. The lack of consistently repeated features when tuning these intervals across the exercise (for example notes 3, 5 and 9) implies that another factor is contributing to this result. Another musical consideration could be the scoring of the exercise. The nature of the exercise modulating through many keys gives rise to many accidentals appearing in

the score and enharmonic equivalents being used. The visual impact of the score, therefore, might subconsciously effect the tuning expectations of the singers. Chord seven is the only chord in which every voice-part is sharpened, perhaps encouraging the singers to sharpen their tuning. This is only speculation and further research using a protocol specifically designed to consider this matter would need to be employed.

In the protocol for the recordings the tuning conditions were randomised and the singers sang all four renditions of the exercise in quick succession. This could have had a significant impact on the results with singers trying to adapt to dramatically different tuning systems, especially without conscious understanding of the purpose of the experiment and so the differences could be confusing. Trained singers are likely to remember the tuning system they have just been exposed to, having already adapted (to some extent) to that system. Subject one singing alto received the examples in alternated tuning conditions, whilst the other subjects tuned twice consecutively in each condition. This could explain the general insecurity of the trajectories for this exercise compared to the other subjects.

The insecurity of the first note in the alto recordings cannot be explained either by intervallic tuning issues (in terms of a preference to tuning certain intervals) or perception of the written score, as the first note is the root note of the chord and the first chord to be tuned. The first note for the alto part was given as the reference tone through headphones to begin the exercise. It is therefore surprising that this note lacks accuracy in terms of the starting fundamental frequency, often beginning sharp of the given note. One reason for this could be the use of synthesised tones as the reference: there is evidence to suggest that it is more difficult to replicate pitches that are presented as ‘unnatural’ sounding [23]. In order to reduce the possibility of the unnaturalness of the synthesised parts influencing the pitching choices of the singers, a different approach will be considered, such as using a live recording of a single vowel and pitch shifting it to create the exercise, such as Ternström et al [12].

4.3 Comparison to previous study

In the original study the mean fundamental frequency was obtained for each note sung and the results compared to the expected values for each temperament. The analysis revealed that overall, singers tended to sing flat of the expected values in the equal tempered condition, and sharp of the expected values when the exercise was justly tuned. There was greater deviation from expected tuning in the just tuned condition than the equal tempered version [14]. The fundamental frequency trajectories support the mean values, having increased stability in the equal temperament condition.

4.4 Predicted tuning of the exercise

The nature of just temperament requires that when predicting pitch drift certain decisions have to be made, including which note of the chord the singers are using as

their reference in tuning. In the current exercise the note which was tied across a chord change was used as the reference note against which to calculate the just tempered chord. This results in considerable overall pitch drift in the piece. In practice, this may not be the natural tendency of singers in a choral context. For example, there is often encouragement amongst singers to tune to the root note in a given chord. In the exercise the reference note is the fifth of the chord, expecting the singers to tune to as a third inversion rather than the root. If the exercise were re-tuned to represent just temperament in relation to the root of each chord, rather than the tied note, this might reveal different results and give further insight into the tuning behaviour of choirs. Certainly, exploring the possibilities of the predicted tuning for the exercise in just temperament would be worthy of further study. The role of the singer maintaining the tied note would be of particular interest, as they would need to adapt their tuning once the new chord entered in order to keep the chord ‘in tune’. The tendency for the fundamental frequency trajectories to alter in the alto when the chord enters (in spite of its role as a reference chord) suggests that this could be indicative of good choral singing.

5. CONCLUSIONS

The results suggest that the singers were most comfortable / accurate when tuning to the equal tempered condition rather than the just tuned version of the exercise. However, decisions made about the predicted fundamental frequency values in just tuning may be responsible for this result rather than reflecting the tuning preferences in choral singing in general. The limited data set collected from only three singers makes any general conclusions as to tuning in choral practice difficult to conclude. The analysis protocols employed in this study were valuable and will prove effective tools in future research.

6. FURTHER RESEARCH

This paper was designed to assess the effectiveness of a protocol for analysing tuning behaviour of singers in a choral situation. A longitudinal study of tuning precision in choral singers is planned for next year which will assess the tuning characteristics of individual singers in ensembles as they are trained as part of the MA in Solo Voice Ensemble Singing in the Music Department at the University of York. The protocol from the current study will be adapted to vary the tuning predictions of the exercise in just temperament, and improve the naturalness of the set choral accompaniment, whilst maintaining the omission of vibrato in the first instance. A larger data set collected using the revised protocol will allow for valuable statistical analysis to be carried out on the results and an insight into the tuning behaviours of singers in ensembles to be made.

7. REFERENCES

- [1] D. Howard, "Equal or non-equal temperament in a capella SATB singing", *Logopedics Phoniatrics Vocology*, vol. 32, no. 1 pp. 87 – 94, 2007.
- [2] M. Lindley, "Just Intonation" in *New Grove Online* <http://www.oxfordmusiconline.com/subscribe/article/grove/music/14564#S14564.2> (last accessed 7th April 2013).
- [3] B. Haynes, "Beyond temperament: non-keyboard intonation in the 17th and 18th centuries", *Early Music*, vol.19, no.3, pp. 357-382, 1991.
- [4] F.Mayer, "The relationship of blend and intonation in the choral art", *Music Educators Journal*, vol. 51, no. 1, pp. 109-110, 1964.
- [5] J.Potter, *The Cambridge Companion to Singing*. Cambridge 2001, p. 160
- [6] C. Hansen, "Tuning the Choir", *Music Educators Journal*, vol. 51, no. 2, pp. 85-89, 1964.
- [7] D. Howard, "Intonation Drift in A Capella Soprano, Alto, Tenor, Bass, Quartet Singing With Key Modulation", *Journal of Voice*, vol. 21, no. 3, pp. 300 – 315, 2007.
- [8] D. Howard, "A capella SATB quartet in-tune singing: evidence of intonation shift", in *Proceedings of the Stockholm Music Acoustics Conference*, 2003.
- [9] J. Nordmark and S. Ternström, "Intonation preferences for major thirds with non-beating ensemble sounds", *Quarterly Progress and Status Report*, vol. 37, no. 1, pp. 57-62, 1996.
- [10] F. Loosen, "The effect of musical experience on the conception of accurate tuning", *Music Perception*, vol. 12, no. 3, pp. 291-306, 1995.
- [11] S. Ternström, "Choir acoustics – an overview of scientific research published to date," *Quarterly Progress and Status Report*, vol. 45, no. 1, 2002.
- [12] S. Ternström and J. Sundberg, "Intonation precision of choir singers", *Journal of the Acoustical Society of America*, vol. 84, no. 1, pp. 59 – 69, 1988.
- [13] S. Ternström and J. Sundberg, "Acoustical factors related to pitch perception precision in choir singing" *Quarterly Progress and Status Report*, vol. 23, no. 2-3, pp. 76 – 90, 1982.
- [14] D. Howard, H. Daffern, and J. Brereton, "Four-part choral synthesis system for investigating intonation in a capella choral singing", submitted to *Logopedics, Phoniatrics Vocology*, 2013 in review.
- [15] J.Shonle and K.Horan, "The pitch of vibrato tones", *Journal of the Acoustical Society of America* vol.67, no.1, pp.246-252, 1980.
- [16] A. Grell, J. Sundberg, S. Ternström, M. Ptok, E. Altenmüller, "Rapid pitch correction in choir singers", *Journal of the Acoustical Society of America*, vol.126, no. 1, pp. 407 – 413, 2009.
- [17] J. Devaney, AMPACT: Automatic Music Performance Analysis and Comparison toolkit [Computer software] www.ampact.org (last accessed 4/4/2013)
- [18] C. Cannam, C. Landone, and M. Sandler, "Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files", in *Proceedings of the ACM Multimedia 2010 International Conference*.
- [19] J. Devaney, M. Mandel, D. Ellis, and I. Fujinaga, "Automatically extracting performance data from recordings of trained singers" *Psychomusicology: Music, Mind & Brain*, vol.21, no.1&No. 2, pp. 108 – 136, 2011.
- [20] N. Orio, and D. Schwarz, "Alignment of monophonic and polyphonic music to a score". In *Proceedings of the International Computer Music Conference*, pp.155–8. 2001
- [21] J. Devaney, M. Mandel, I. Fujinaga "Characterizing singing voice fundamental frequency trajectories" in *IEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New York, 2011,
- [22] H. Kawahara, A. de Cheveigne "Yin, a fundamental frequency estimator for speech and music" *Journal of the Acoustical Society of America* vol. 111, no. 4 pp. 1917–1930, 2002.
- [23] B. Kleber, A. G. Zeitouni, A. Friberg, and R.J. Zatorre, "Experience-Dependent Modulation of Feedback Integration during Singing: Role of the Right Anterior Insula", *Journal of Neuroscience*, Vol 33 pp. 6070-6080, 2013

ASSESSMENT OF THE ACOUSTICAL IMPACT OF PIRIFORM SINUSES IN MRI BASED VOCAL TRACT REPLICAS

Bertrand Delvaux, David Howard

Department of Electronics, University of York, UK

ABSTRACT

Mixed soft/solid models of the Vocal Tract were molded with a 3D rapid prototyping technique based on MRI data obtained from two male singers during the phonation of five English vowels as in hard, stern, neap, port and food. The replicas are used to assess the spectral role of the piriform sinuses. In an anechoic chamber, a sound source producing a sinesweep is connected to the tracts. The acoustical response is then used to compute the transfer functions. In one case, the piriform sinuses remain empty, in the other, they are filled with water. The spectral impact of the piriform sinuses is discussed.

1. INTRODUCTION

The piriform fossae, or piriform sinuses, owe their name to the pear shape they display. These two cavities are located posteriorly at the bottom of the pharynx. Together with the laryngeal vestibule and ventricles, they form the hypopharyngeal cavities (see Fig. 1) who are thought to be responsible for the uniqueness of a voice, by shaping the formants F3, F4 and F5, with large inter-speaker variations and small intra-speaker (i.e., inter-phoneme) variations [1]. In particular, the piriform fossae, as side branches of the Vocal Tract (VT), are thought to produce troughs in the region of 4 to 5 kHz, and play a significant role in the singer's formant between 2 and 3 kHz [2]. The singer's formant is a well-established feature of the acoustic output from the VT of trained opera singers that is independent of the vowel being sung. This suggests that it is related to a region of the VT that does not change greatly in shape with articulation; anatomically, this must be associated with the hypopharyngeal cavities.

Dang and Honda [3] carried out a preliminary study of the piriform fossae on models and humans, injecting water in the sinuses of humans phonating in a supine position and in mechanical models of half VTs.

Here we present a study based on full VT replicas moulded with a 3D rapid prototyping technique and based on MRI data obtained while subjects were singing in a supine position.

2. MOULDING THE VOCAL TRACTS

The VT replicas were molded based on volumetric MRI data collected while a professional male singer was singing 5 En-

glish vowels in a supine position [4], by a 3D fast gradient echo sequence. The relaxation time was 4.8 ms and the excitation time was 1.7 ms. Acquisition is isotropic 2mm in a 192x192 matrix. Output is then interpolated to 512x512 using 50% slice overlap giving an effective anisotropic output of 0.75x0.75x1mm. A stack of 80 images is produced in the midsagittal plane in approximately 16s.

The MRI data were then segmented with the open source code ITK-Snap, to rebuild a 3D Vocal Tract, whose .STL file was then sent to Sirris, an R&D company for 3D rapid prototyping. The material used was TangoPlus 75 and 100 for the flexible and solid parts respectively. The tracts were printed on a Connex 500.

The technical specificity of the vocal tracts consists of an insert to be plugged onto a 952.210UK - driver unit, actual shaped lips, and an opening at half of the VT, just above the valleculae to allow liquid to be poured in the cavity. The tutor is only there to support the upper part of the pharynx which might collapse because of the gravity. The replicas are made of flexible TangoPlus 70 except on a section of about 1 cm from each side of the opening of the VT made of solid TangoPlus 100. The thickness of the shell of the VT is of 2 mm.

3. ACOUSTIC MEASUREMENTS

Three Vocal Tract replicas for the English vowels as in *port*, *neap* and *food* were used for the experiment.

The acoustic effects of the piriform fossae on the transfer function of the full and half tract were investigated by an experiment using the amplitude modulation scheme for exponential sweep [5]. It consists of using an exponential swept signal of the form

$$s(t) = \sin[\theta(t)] = \sin \left[K \left(e^{\frac{t}{T}} - 1 \right) \right] \quad (1)$$

where

$$K = \frac{T\omega_1}{\ln \frac{\omega_2}{\omega_1}}, L = \frac{T}{\ln \frac{\omega_2}{\omega_1}} \quad (2)$$

with t being the time, T the duration of the sweep and $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$ the lower and higher extremities of the frequency range swept by the sine.

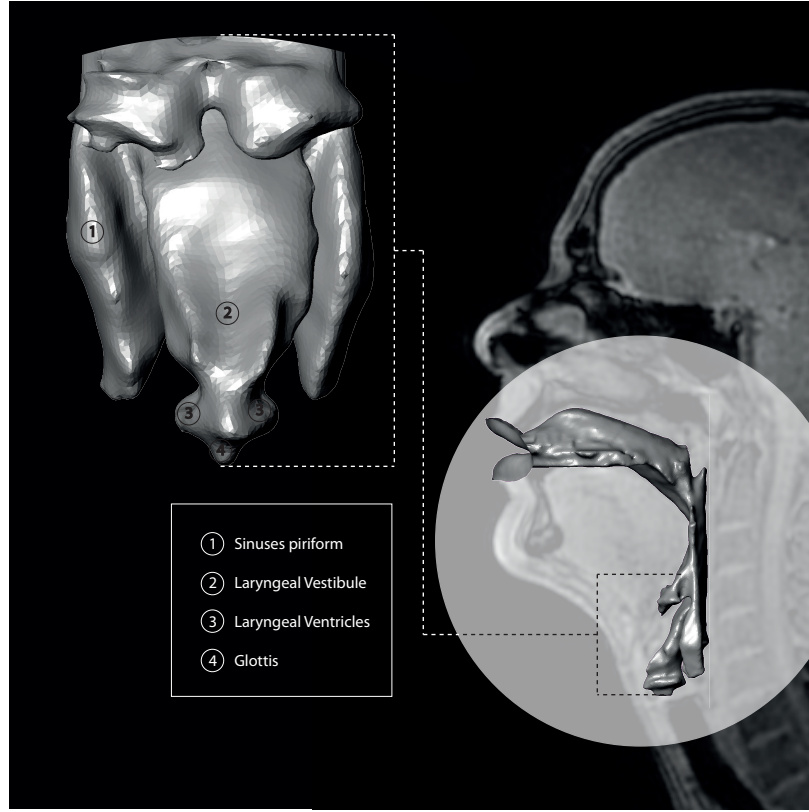


Fig. 1: Details of the hypopharyngeal cavities, from 3D segmentation of a Vocal Tract whose subject is singing the English vowel as in *stern*, superimposed on the MRI slice whose it is extracted.

The amplitude of this exponential swept signal is then modulated by the factor

$$m(t) = \frac{\omega_1}{\omega(t)} \quad (3)$$

with

$$\omega(t) = \frac{d[\theta(t)]}{dt} = \frac{K}{L} e^{\frac{t}{L}} \quad (4)$$

being the instantaneous frequency, as developed in [5].

This amplitude modulation allows the sine sweep to sweep all the frequency range of interest while having a flat spectrum.

The experiment was carried out in a 6-sided anechoic chamber. A probe microphone type 40SA was located 3 cm far from the radiating end. The signal was preamplified by a power Module type 12AA before being written on a USB type device with a 96 kHz sampling rate on a 16 bits WAV file. The VTs were excited at the glottis by a 952.210UK driver unit, with the amplitude modulated sine sweep mentioned before.

The setup had 2 configurations : one with the piriform sinuses filled with water, once without. Accurate water pouring was performed with the aid of a pipette.

An FFT with a hanning windowing and a bilateral zero-padding was performed on the signal to obtain the transfer function of the full/half VTs with/without piriform sinuses filled with water.

4. RESULTS

The results of the full VT transfer functions are displayed on Figs. ??, ?? and ?. The dotted line represents the VT in its normal state whereas the dashed line represents the VT whose piriform sinuses have been filled with water. We can clearly see the effect of the piriform sinuses, creating a trough between 4 and 5 kHz as mentioned in the literature and consistent with the results obtained by Dang and Honda [3], whereas the spectrum remained almost unchanged below 2 kHz. We also observe among the three sung vowels a shift of the singing formant towards higher frequencies when the water is filled in the piriform sinuses. The results of the half VT transfer functions are displayed on Figs. ??, ?? and ? and confirm the observations made before. In conclusion, it can be seen on the graphs that the piriform sinuses emphasize the singer's formant by shifting it to lower frequencies and

creating a trough right after (in the region of 4 to 5 kHz). This might give a clearer perception to the human hearing.

5. CONCLUSIONS

The work presented here has focused on the role of the piriform fossae in a full/half Vocal Tract replica, of equal thickness, based on MRI data collected when a male singer was singing different vowels, whose transfer functions were estimated with a sine sweep method. The work by Dang and Honda [3] considers only half a mechanical Vocal Tract, carved out of a vinyl chloride block, during speech, with a two-point sound-pressure method. The present study allows a special look into the impact of the piriform sinuses on the singer's formant and above frequencies, on a full tract replica with a more recent approach of the sine sweep.

6. REFERENCES

- [1] T. Kitamura, K. Honda and H. Takemoto, "Individual variation of the hypopharyngeal cavities and its acoustic effects," *Acoust. Sci. and Tech.*, vol. 26, no. 1, 2005.
- [2] J. Sundberg, "Articulatory interpretation of the singing formants," *J. Acoust. Soc. Am.*, vol. 55, pp. 838–844, 1974.
- [3] J. Dang and K. Honda, "Acoustic characteristics of the piriform fossa in models and humans," *J. Acoust. Soc. Am.*, vol. 101, no. 1, January 1997.
- [4] M. Speed PhD Thesis *Voice Synthesis Using the Three-Dimensional Digital Waveguide Mesh* Department of Electronics, University of York, UK. March 2012
- [5] Q. Meng, D. Sen, S. Wang and L. Hayes, "Impulse response measurement with sine sweeps and amplitude modulation schemes," in *Signal Processing and Communication Systems*, 2008.

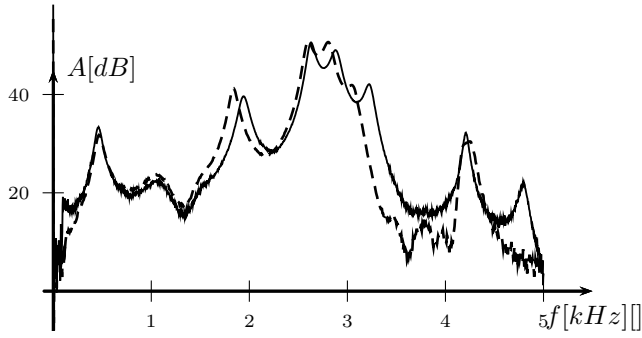


Fig. 2: Transfer function of the Full Vocal Tract of *neap* with water (solid line) and without (dashed line) in the piriform sinuses

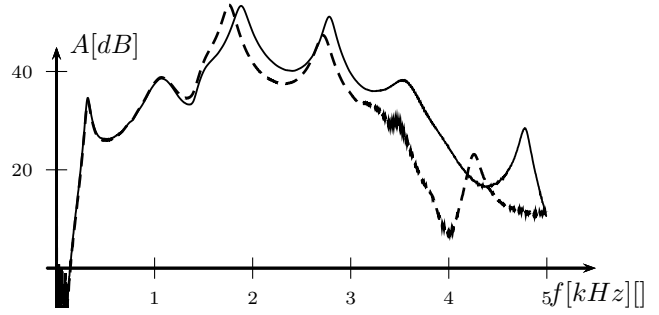


Fig. 5: Transfer function of the Half Vocal Tract of *neap* with water (solid line) and without (dashed line) in the piriform sinuses

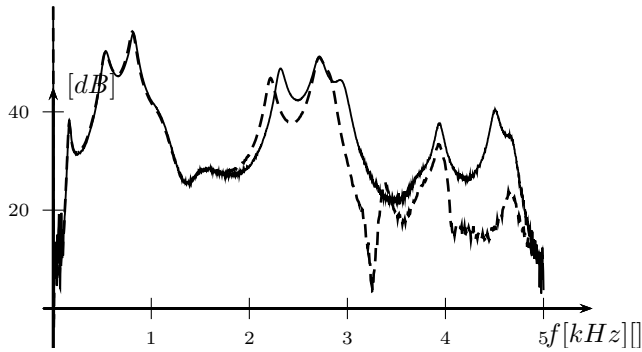


Fig. 3: Transfer function of the Full Vocal Tract of *food* with water (solid line) and without (dashed line) in the piriform sinuses

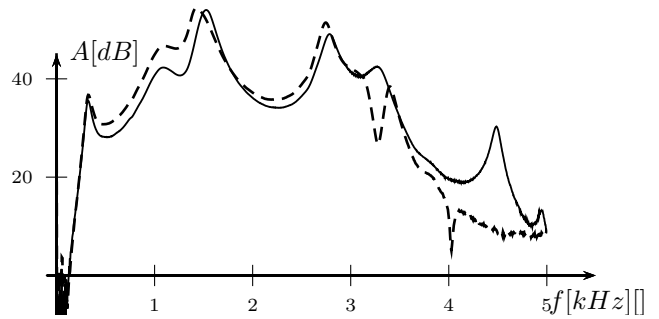


Fig. 6: Transfer function of the Half Vocal Tract of *food* with water (solid line) and without (dashed line) in the piriform sinuses

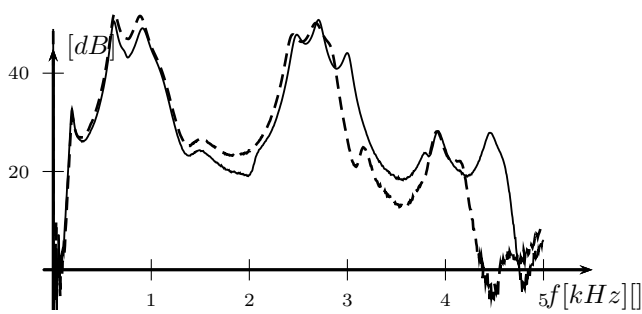


Fig. 4: Transfer function of the Full Vocal Tract of *port* with water (solid line) and without (dashed line) in the piriform sinuses

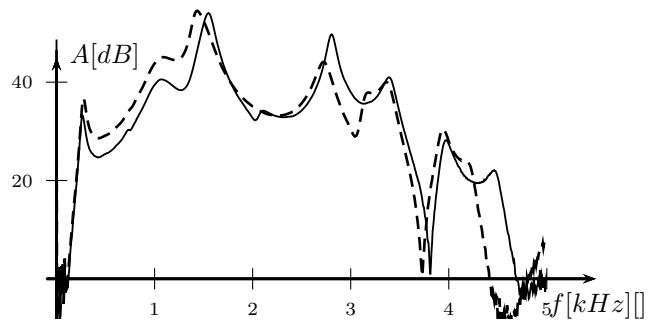


Fig. 7: Transfer function of the Half Vocal Tract of *port* with water (solid line) and without (dashed line) in the piriform sinuses

VIBRATO ANALYSIS IN BYZANTINE MUSIC

Dimitrios S. Delviniotis

Georgios Kouroupetroglou

Sergios Theodoridis

National and Kapodistrian University of Athens, Department of Informatics and Telecommunications,
Panepistimioupolis, Ilisia, GR-15784, Athens, Greece
{ddelvis, koupe, stheodor}@di.uoa.gr

ABSTRACT

In this work we study the vibrato rate, extent and intonation in Byzantine Music. Two methods of analysis have been applied: the first based on the analytical signal and the second on the crests and troughs of the waveform of the vibrato signal. Tones - samples were extracted from ascending and descending music scales, chanted by five famous singers for all the Greek vowels. The two analysis methods produced identical results in the level of significance, $\alpha=0.05$, concerning the mean extent, the mean standard deviation of the rate and the mean intonation, while they differed in the rate (1.9-6.6%), the mean standard deviation of the extent (4-6 cents) and the standard deviation of intonation (0.46-1.20 Hz). Typical values of the average rate within a tone were found to be 5.35 Hz (SD: 0.96 Hz) and 5.13 Hz (SD: 0.95 Hz), while the most frequent values were 4.8 Hz and 4.5 Hz, for the first and second method, respectively. The average extent within a tone was 50 cents (SD: 18 cents). Moreover, the dependence of the vibrato parameters on pitch and sound intensity has been studied; there was no systematic relationship between them.

1. INTRODUCTION

The term “Byzantine Music” (BM) describes the contemporary church music of the Greek Orthodox Church, primarily, but also it refers to both the medieval sacred chant of Christian Churches following the Constantinopolitan Rite and the secular music in the Byzantine and post-Byzantine era. The eight-mode system of BM has affected the modern Greek folk and popular songs. In an effort of a comprehensive study of Greek singing, the study of BM is necessary, fundamental and of primal importance. Although musicologically, BM has been systematically studied [1], its acoustical attributes have not been thoroughly examined [2, 3, 4, 5].

The BM is purely vocal music and is performed without the accompaniment of musical instruments. Therefore, only the voice must satisfy the requirements for artistic

musical ornaments, which embellish the melody in the music performance. Vibrato constitutes a characteristic ornament of the melodic voice [6]. It has been extensively studied in the context of various kinds of vocal music world-wide [7, 8, 9, 10]. Many of these studies had to do with the lyrical songs of Western opera, addressing the basic vibrato parameters rate and extent [10, 11, 12]. Also, dependencies were found between the appropriateness or not of vibrato with its rate, extent, periodicity and onset. Moreover, the relationship of poor or good vibrato with respect to the variability of its rate and extent has been studied [13]. In BM, the vibrato seems [14] to be a rather rare phenomenon, with rate and extent values to be different from those in Western opera. In that preliminary study [14] only average values of the rate, extent and vibrato duration were measured, by analyzing a single BM hymn.

In this work, the characteristics of BM vibrato rate and extent are studied in order: a) to assess their values in detail, and b) to search for any relationships with other voice features, such as fundamental frequency (f_0) and sound intensity.

2. METHOD OF THE ANALYSIS AND MATERIAL

Vibrato analysis methods have been based on: a) studying the spectrogram under manually intervention of the user [11, 12], b) frequency analysis with a sliding window on the vibrato waveform [13], c) calculating the instantaneous frequency, resulting after the application of the Hilbert transform on the vibrato signal [15, 16].

2.1 The analytic signal

According to the last mentioned method above, vibrato is considered to be a time dependent signal of the form:

$$\hat{f}(t) = b(t) + a_v(t)\cos\varphi_v(t) \quad (1)$$

where:

$$\varphi_v(t) = 2\pi \int_{-\infty}^t f_v(\tau) d\tau \quad (2)$$

The signals $a_v(t)$ and $\varphi_v(t)$ correspond to the time varying parameters of the extent and the rate of vibrato,

respectively. $b(t)$ corresponds to the intonation of vibrato. It is known [15] that there is no single solution to the problem of determining these three signals, which constitute the time varying model in (1) and (2). The signal $b(t)$, could be obtained by passing the signal $f(t)$ through a low pass filter with a cutoff frequency less than 3Hz, with the condition the vibrato rate to be equal or greater than 4Hz.

By subtracting the signal $b(t)$ from $f(t)$, we get the pseudo-sinusoidal signal $s(t)$:

$$s(t) = f(t) - b(t) \quad (3)$$

Further analysis of the $s(t)$ estimates the signals $\alpha_v(t)$ and $\varphi_v(t)$ through the analytic signal, which is defined as:

$$z(t) = s(t) + jH[s(t)] \quad (4)$$

where $H[s(t)]$ is the Hilbert transform of $s(t)$.

The analytic signal $z(t)$ can be expressed in polar coordinates as follows:

$$z(t) = A(t)e^{j\theta(t)} \quad (5)$$

In the case $s(t)$ is a band-pass signal, then the amplitude and the phase derivative of the analytic signal define the amplitude and frequency modulation, respectively. The extent of the vibrato signal can be estimated from the amplitude $A(t)$ divided by 2π . The rate of the vibrato signal can be estimated from the derivative of $\theta(t)$ divided by 2π .

2.2 The algorithm of this analysis

2.2.1 The first algorithm (analytic signal)

For the estimation of $b(t)$ we have followed a different analysis method from that in [15]. Specifically, i) instead of low pass filtering the vibrato signal, first, the crests and troughs are located at the waveform and then the center of the vertical distance between two successive extremes is located, for all the maximum–minimum pairs. ii) The number of these central points is augmented via a cubic piecewise interpolation process, which preserves the shape of the data and respects monotonicity [17]. The resulting $b(t)$ is a smoothed line following the slow changes in the vibrato waveform. We have followed the above estimation approach for the intonation for two reasons: a) Due to the rather low values of the vibrato rate in BM [14], a low pass filtering is not possible to be applied without loss of information that the vibrato signal conveys, and b) because part of the low pass filtering resulted in an intonation, which was not passing through the central points of the vibrato waveform, but instead it received values beyond the extremes. This last observation did not agree with the pitch perception of the vibrato tone.

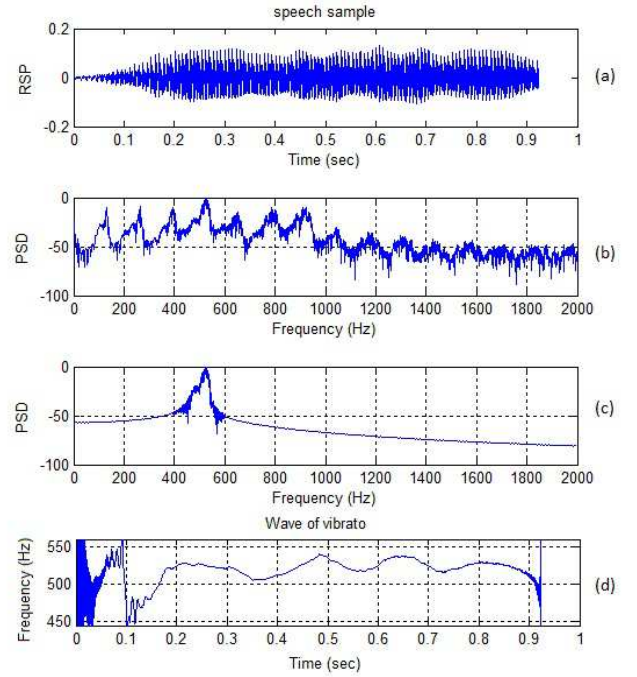


Figure 1. Signals for the first two steps of the analysis: (a-b) tone's waveform and spectrum, (c) fourth partial's selection and (d) vibrato waveform.

The steps of the first algorithm of the analysis are:

- 1) Compute the waveform of a partial of the voice signal, via band pass filtering. Select the partial, which shows the highest signal to noise ratio (Fig. 1a-1c).
- 2) Find the vibrato signal through the analytic signal based on the equations (4) and (5) (Fig. 1d).
- 3) Remove the rapid changes of the vibrato waveform through low pass filtering with a 20 Hz cutoff. Estimate the time indexes of its crests and troughs (Fig. 2a).
- 4) Compute the coordinates of the central points for each pair of successive extremes (asterisks in figure 2).
- 5) Estimate more points through cubic interpolation between the central points (Fig. 2a). (Compare the intonation time series with the low pass filtered with cutoff 2 Hz).
- 6) Subtract the intonation signal from the vibrato signal.
- 7) Compute the analytic signal with the equation (4).
- 8) Obtain the rate (instantaneous frequency) and the extent of vibrato from the relation (5) (Fig. 2b-2d).

In step (5), the intonation time series, although it is very close to the low filtered vibrato waveform at four points (Fig. 2a), it deviates at the other points, especially at the

edges. For this reason, the intonation waveform via cubic interpolation was preferred to the other. Specification of the time segment of the vibrato signal to be analyzed was implemented as follows: First, time borders were roughly defined, manually. Then, the time limits were marked precisely so that the left boundary is the midway of the distance between the two first successive extremes, in horizontal and vertical direction. The right border was found in the same manner for the last two successive extremes.

2.2.2 The second algorithm

To assess the accuracy of the above method, the vibrato characteristics rate, extent and intonation were also calculated by the following system of relations (6), (7) and (9). Henceforth we will refer to them as the second method of analysis. All the calculations in this method were based on the extremes in the waveform of vibrato. In Figure (2), the values obtained based on the extremes are represented together with the signals of the rate and extent as estimated by the first algorithm. The equations for estimation of the rate and extent through the second method are the following:

$$rate(m) = 1 / (2(t(k+1) - t(k))) \quad (6)$$

$$extent(m) = |a(k+1) - a(k)| / 2 \quad (7)$$

where $t(k)$ and $a(k)$ are the instant and the value of the extreme k respectively. In this way, the rate and extent is computed for each semi-cycle of vibrato, by, roughly, doubling the number of the time instants.

However, since the variation of the values of the extent also increases, the mean between the two values of successive semi-cycles was calculated. (Fig. 2c, cycles). In a previous study [12], the extent was calculated as the percentage of the mean intonation between two semi-cycles, namely, based on the relation (8):

$$extent(m) = \frac{|a(k-1) - 2a(k) + a(k+1)|}{|a(k-1) + 2a(k) + a(k+1)|} \quad (8)$$

which can be obtained from the equation (7), as the ratio between the semi-sum of pairs of successive extreme values (two semi-cycles) and the mean intonation for these two semi-cycles. This value of the mean intonation is defined as:

$$intonation(m) = a(k) + a(k-1)/2 + (a(k+1) + a(k))/2 \quad (9)$$

where m takes values from 1 up to $k-1$.

The final values of the extent are converted in the music scale of semitones with frequency of reference the note of A3 (220 Hz), for any comparison in the psychoacoustic scale with previous studies, as well as because the order

of the partial selected for the analysis varies among the tones.

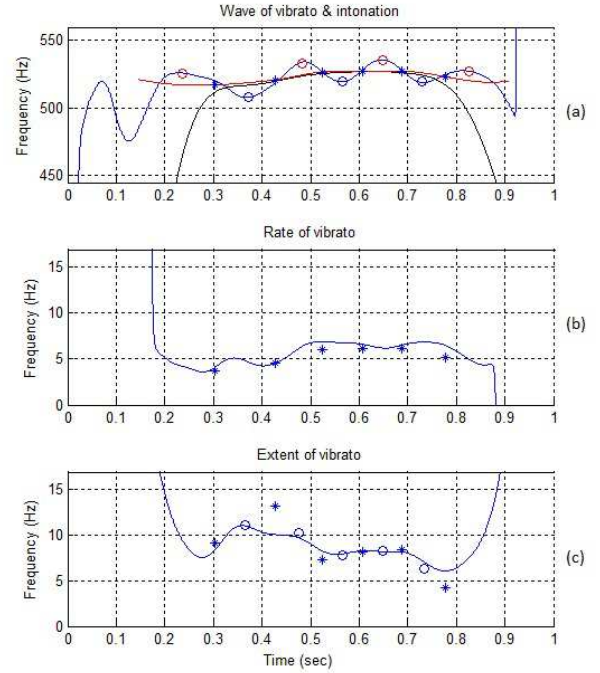


Figure 2. Signals and results of the algorithm of the analysis: (a) Intonation after low pass filtering (black line) and cubic interpolation (red line). Extremes at the smoothed waveform (red and blue cycles) along with the half-way distance points between them. (b) The values of the rate being estimated for each semi cycle of vibrato (asterisks) and the rate's waveform through the analytic signal. (c) Estimated values of the extent through the analytic signal (line), the equation (7) – asterisks and the equation (9) cycles.

The values of the relative intensity level of the sound of the DAMASKINOS corpus [18] refer to the same distance from the microphone (30cm), for all recordings by using the same apparatus and settings. Intensity levels were calculated in relation to the loudness threshold. All the software developed for analysis was implemented in the MATLAB programming environment.

2.3 Material

Choosing the audio sample-tones for the analysis it was a difficult task, because of the rareness of the phenomenon of vibrato in BM. This is probably mainly due to stylistic reasons, an explanation that is supported by the short duration of vibrato [14] (less than 1.5sec, with an average 0.7sec). Assuming an average rate equal to 4.2 Hz (in BM hymns of medium rhythm) then the number of vibrato periods for analysis is approximately three, on average. For the purpose of finding possible longer tones for more accurate analysis, we selected those parts of the

DAMASKINOS corpus [18], which include chanted scales, both ascending and descending ones. The implicit assumption here is that the parameter values of vibrato will be approximately similar to those found in the musical performance of Byzantine Music [14]. Chanters were asked to chant each tone of the scale slowly and trying to keep its F0 invariable. In order to have a small representative set of subjects [19], five chanters were selected so as to be one chanter from each chanting category according to some classification, based on their spectra in a previous study [14]. The ages of these chanters ranged between 40 and 60 years, which were suitable to combine art experience with a cultured voice. The tones for each chanter were selected to belong to six frequency bands of 1 semitone width each and centers being defined by the notes C3, D3 #, F3 #, A3, C4 and D4 # (± 0.5 semitones). Any tone contains only one vowel out of /a/, /o/, /i/, /e/ and /u/. Two tones were selected for each vowel and chanter, one from ascending and one from descending scale. The constant difference in frequency between the notes by three semitones was chosen in order to examine whether there are dependencies on the other

logarithmically varying sizes (intensity level, extent and intonation).

3. RESULTS

Table (1) compares the two methods of vibrato analysis through a paired t-test between the mean values of the vibrato parameters, the rate, extent, intonation and their average standard deviations within each tone. The t-test reveals that the results of the two methods do not differ in the mean standard deviation of the rate, the mean extent and the mean intonation at the level of significance $\alpha = 0.05$; however, they do differ in the mean rate and the mean standard deviations of the extent and intonation. Table (1) also shows the confidence intervals of the average difference between the values of the second and the first method, from which it follows that the maximum and minimum average difference in rate is 0.34 Hz and 0.1 Hz greater in the first method, respectively. In percentage terms, and taking into account the Table (2), these values correspond to 6.6% and 1.9% of the average of the second method, for all the tones.

	Mean of the rate (Hz)	SD of the rate (Hz)	Mean of the extent (sem)	SD of the extent (sem)	Mean of the intonation (Hz)	SD of the intonation (Hz)
h value	1,00	0,00	0,00	1,00	0,00	1,00
p value	0,00	0,78	0,34	0,00	0,99	0,00
CI lower	-0,34	-0,11	-0,04	0,04	-17,30	0,46
CI upper	-0,10	0,08	0,02	0,06	17,47	1,20
SD	0,74	0,57	0,19	0,07	108,41	2,30
t stat	-3,56	-0,28	-0,96	8,72	0,01	4,43
df	598,00	598,00	598,00	598,00	598,00	598,00

Table 1. Paired t-test between the two methods of vibrato analysis.

	Mean of the rate (Hz)	SD of the rate (Hz)	Mean of the extent (sem)	SD of the extent (sem)	Mean of the intonation (Hz)	SD of the intonation (Hz)
Mean (1st)	5,35	0,96	0,51	0,10	486,9	3,24
SD (1st)	0,69	0,46	0,18	0,06	114,3	1,99
Mean (2nd)	5,13	0,95	0,49	0,14	487,1	4,06
SD (2nd)	0,75	0,75	0,18	0,07	114,4	2,41

Table 2. Average values and their standard deviations of vibrato parameters over all subjects and pitches.

The average difference of the standard deviation of the extent is between 0.04 and 0.06 semitones, i.e. 4 and 6 cents, and in percentage values are 40% and 60% greater than the respective values of the first method, for all the tones. The confidence interval of the average difference of the standard deviation of intonation is (0.46 Hz, 1.20 Hz), i.e. (23.1%, 60.3%) greater than the values of the first method, for all the tones.

The average values for the entire set of tones, of the mean rate, mean extent and within the tone their standard deviations, for both methods are presented in Table (2). Moreover, the standard deviations of the within tone means for all tones are provided. The distributions of these values are presented in Figure (3). Although the within tone mean of the rate $\pm 1SD$ is 5.13 ± 0.95 Hz, for the second method, (5.35 ± 0.96 Hz for the first method) the value 4.5 Hz seems to have higher occurrence for the second method (4.8 Hz for the first method - Figure 3a) as the histogram is not symmetric in its maximum. The second method shows a tendency for slightly lower rate values compared to the first and this is not the case in the standard deviations (Fig. 3a, c). The intra-tone variation in the rate is greater than that for all tones and chanters (0.95 Hz versus 0.75 Hz - Table 2). For the rate, the largest percentage of values (mean $\pm 2SD$) of the within tone standard deviation has a value less than 2 Hz (Figure 3c). The corresponding percentage for the rate has a value less than 7 Hz.

The extent shows approximately the same distribution of mean values within the tone (Figure 3b) and its standard deviations tend to be smaller in the first method than the second (Fig. 3d). The majority of values (mean $\pm 2SD$) of the mean vibrato extent over all tones vary within the range of 0.50 ± 0.36 semitones (Table 2). This finding is quite different from the values at the opera, where the extent reaches the maximum value of 123 cents [12].

Examining the dependency of the rate, the extent and their standard deviations on the f_0 , we observe that there is a negligible tendency for the within the tone mean rate to increase by 0.13 Hz per 100 Hz raise of the f_0 (Fig. 4a). The regression analysis was based on the least squares and the assumption that there is a linear relationship between the dependent and independent parameter. In a total change by 300 Hz, the rate can be increased by 0.39 Hz, which corresponds to a change of less than the standard deviation (0.75 Hz - Table 2). For the same reason, the standard deviation of the rate is not affected by the change of the pitch (Figure 4b). Similarly, both the intra-tone mean extent and its mean standard deviation decrease by increasing the pitch by 0.7 cents/semitone and 0.2 cents/semitone, for the mean and standard deviation of the extent, respectively. In a total change of f_0 by 16 semitones, the reduction is 11.2 cents

and 3.2 cents, for the mean and standard deviation of the extent, respectively.

The corresponding standard deviations across the tones are 18 and 7 cents, respectively, which are values greater than of the overall pitch change. The relationship between the intensity of the tone and f_0 in semitones, as expected, is linear, with a slope of about 0.55 dB / semitone (Figure 5a). Also, the within the tone extent of vibrato and its within tone mean standard deviation show a linear relationship with a slope 1.2 semitones of extent / semitone of SD (Figure 5b). In other words, for an increase of 1 semitone in extent, its standard deviation increases by 0.83 semitones. Finally, the extent of the vibrato is not affected by the changes in intensity for all the tones as shown by the regressive analysis in Figure (5c).

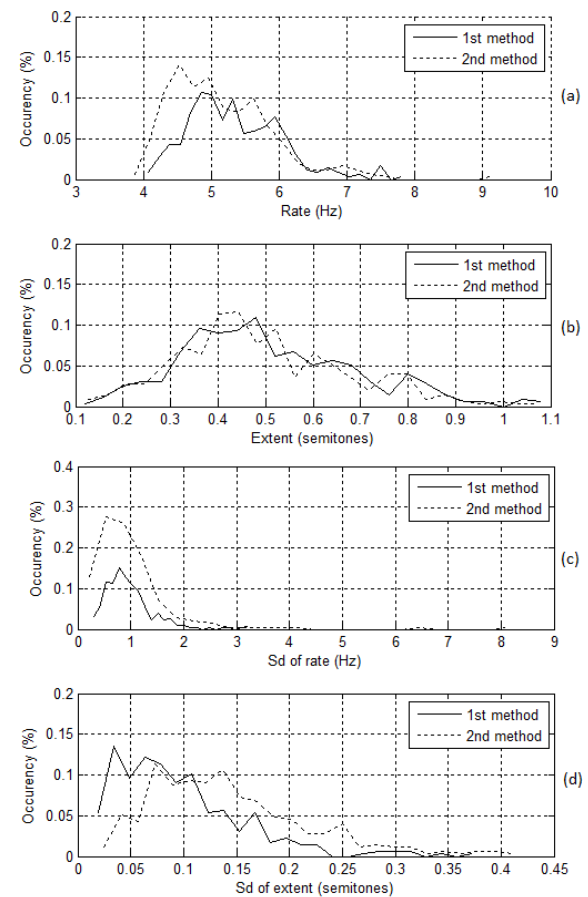


Figure 3. Histograms of the estimated values for the rate (a), the extent (b), the SD of the rate (c), and the SD of the extent (d), by the two methods of analysis.

4. DISCUSSION

The two methods of vibrato analysis give identical results concerning the within tone mean standard

deviation of the vibrato rate, the mean intra-tone vibrato extent and the mean intra-tone intonation. Although it seems that do not agree to each other in the intra-tone mean rate, the maximum difference of 0.34 Hz is rather

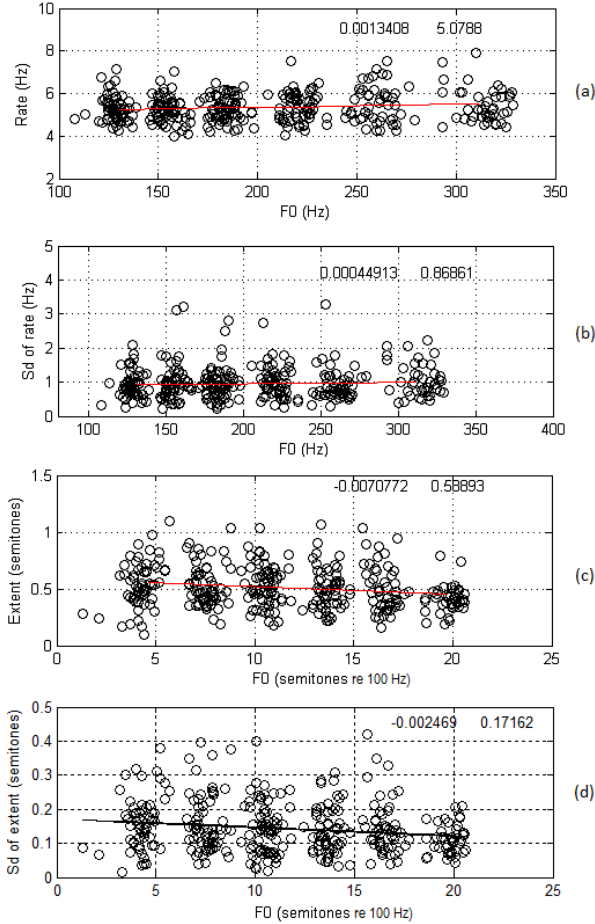


Figure 4. Values of vibrato characteristics in BCM for the (a) rate, (b) SD of the rate, (c) extent, and (d) SD of the extent, in relation to f_0 for all the tones analyzed. Straight lines of Least Squares fitting are shown along with their parameters.

small, provided that the intra-tone standard deviation is 2.5 times greater. The major cause of this difference is due to improper placement of the extremes of vibrato waveform, despite the smoothing of the analysis algorithm. This measurement error of the rate mainly affects the second method. Also, there may be changes in the rate values due to the pseudo-sinusoidal waveform of vibrato, which, although they are taken into account in the first method, do not affect the values estimated by the second one (see Figure 2b, time 0.2-0.4sec). The difference between the two methods in the intra-tone mean standard deviation of the extent by 6 cents (at most), is obviously due to the estimation of the intonation time series. Since the within the tone mean standard

deviation of intonation is a bit bigger in the second method by 1.2 Hz (at most) and the mean difference between the two methods is also about 1.2 Hz, this difference could be ascribed mainly to the intonation time series. Improvements in the analysis algorithm could be made by applying a better smoothing of the vibrato waveform, for the second method and another way of estimation of the intonation time series for the first method.

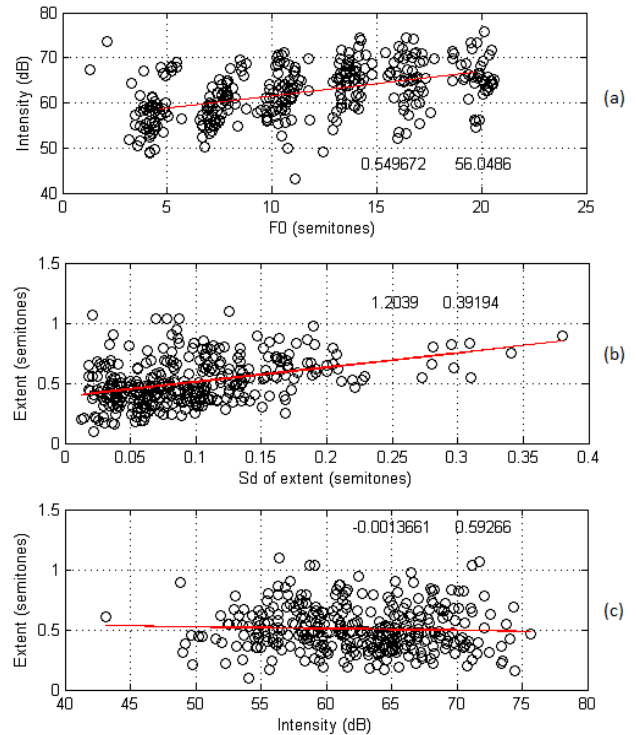


Figure 5. Correlograms for all analyzed tones between: (a) relative intensity and f_0 , (b) extent of vibrato and SD of extent and, (c) extent of vibrato and relative intensity. Straight lines of Least Squares fitting are shown along with their parameters.

It is obvious that the standard deviation across the chanter and vowels is smaller than that within the tone. This may suggest that the vibrato rate depends mainly on the stylistic characteristics of BM rather than the individual characteristics of the chanters. Yet, the fact that this is not affected by the pitch change inside the pitch range of the “average” chanter is in line with the literature, where no systematic differences have been observed [10]. The maximum value of vibrato extent in BM of 86 cents, has been also found in a previous study [14], which was implemented with a different sample of chanters and is likely due to low pitches in BM. This follows from the fact that the intensity varies depending on the pitch (Figure 5a); also it was observed that whenever it increases, the extent of vibrato increases too.

However, when the intensity decreases, the extent tends to remain stable, while exhibiting a maximum in the middle pitches [20]. In this study, the extent does not change with the increase of intensity within the pitch range of the “average” chanter. This may be explained by the fact that in the entire set of tones, half came from descending scales and furthermore these changes are different for each chanter. The marked decrease in the average value over the frequency range (302.3Hz -320.2 Hz: D4 # \pm 0.5 semitones) reveals that a maximum exists in the extent values somewhere in the previous middle pitches. Both the rate and extent of vibrato differ slightly between the tones of musical scales and those of musical performance [14] by 4.5 or 4.8 Hz versus 4.1 Hz and 0.5 semitones versus 0.6 semitones, respectively. In [14], tones extracted from a single Byzantine hymn (of melodic sticheraric type [1]) were analyzed questioning these differences between musical scales and performance. Further investigation is needed for these in order to be considered systematic. Besides this, reverse differences have been reported between sustained tones and real performance of a song. More specifically, lower rate values have been found in sustained notes than inside a song [21]. In addition, these differences in vibrato rate and extent in BM could also be due to the method of the analysis. In the current work other than the autocorrelation method was used and issues such as rate and extent time series, vibrato intonation and various dependencies on f0 and intensity were considered here.

5. CONCLUSIONS

Two methods of vibrato analysis applied in BM showed identical results although with slight differences in the mean intra-tone rate and variability of the standard deviations of the intensity. There was no systematic dependence of the vibrato parameters on the f0 and BM intensity. There is a need to investigate further the variation of the characteristics of BM vibrato in every tone per chanter and by increasing the sample for analysis, by including more singers and tones.

Acknowledgments

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: “THALES - Investing in knowledge society through the European Social Fund”, under the project: “ERASITECHNIS: Processing, Analysis, Synthesis and Context-Based Retrieval for Multimedia Music-Related Data Bases of Traditional Music and Dancing Recordings”, MIS 375435.

6. REFERENCES

- [1] G. Stathis (ed.), *Theory and Praxis of the Psaltic Art, Proceedings of the II International Congress of Byzantine Musicology and Psaltic Art*, Athens, 2003, ISBN 9608835283.
- [2] D. Delviniotis, “A classification of Byzantine singing voices,” in Proc. of EUSIPCO-98, Rhodes, Greece, 1998, pp.129-132.
- [3] I. Zannos, A. Georgaki, D. Delviniotis and G. Kouroupetroglou, “Real-Time Control of Greek Chant Synthesis”, in Proc. of Int. Conf. Sound and Music Computing, Marseille, France, 2006, pp. 47-52.
- [4] G. Chryssochoidis, D. Delviniotis, and G. Kouroupetroglou, “A semi-automated tagging methodology for Orthodox Ecclesiastic Chant Acoustic corpora,” in Proc. of the Intern. Conf. Sound and Music Computing, Lefkada, Greece, 2007, pp. 126-133.
- [5] D. Delviniotis, G. Kouroupetroglou, and S. Theodoridis, “Acoustic Analysis of musical intervals in modern Byzantine Chant Scales,” *J. of the Acoustical Society of America*, vol. 124, No 4, pp. EL262-EL269, 2008.
- [6] E. Ekholm, G.C. Papagiannis, and F.P. Chagnon, “Relating objective measurements to expert evaluation of voice quality in Western classical singing: critical perceptual parameters,” *J. Voice*, vol.12, pp. 182-196, 1998.
- [7] C.E. Seashore (ed.), *The Vibrato*. Iowa: University of Iowa Press, 1932.
- [8] J. Sundberg, *The Science of the Singing Voice*. DesKalb, IL, Northern Illinois University Press, 1988.
- [9] Y. Horii, “Acoustic analysis of vocal vibrato: a theoretical interpretation of data,” *J. Voice*, vol. 3, pp. 36-43, 1989.
- [10] T. Shipp, R. Leanderson, and J. Sundberg, “Some acoustic characteristics of vocal vibrato,” *J. Res. Singing*, vol. 4, pp. 18-25, 1980.
- [11] E. Prame, “Measurement of the vibrato rate of ten singers,” *J. Acoust. Soc. Am.*, vol. 96, pp. 1979-1984, 1994.
- [12] E. Prame, “Vibrato extent and intonation in professional Western lyric singing,” *J. Acoust. Soc. Am.*, vol. 102, pp. 616-621, 1997.
- [13] J.A. Diaz, and H.B. Rothman, “Acoustical comparison between samples of good and poor vibrato in singers,” *J. Voice*, vol. 12, no. 2, pp. 179-184, 2002.
- [14] D. Delviniotis, “Acoustic characteristics of modern Greek Orthodox Church Music,” *J. Voice* (accepted for publication), 2013.
- [15] I. Arroabarren, M. Zivanovic, J. Bretos, A. Ezcurra, and A. Carlosena, “Measurement of Vibrato in

- Lyric Singers,” *IEEE Transactions on Instrum. and Measurement*, vol. 51, no 4, pp. 660-665, 2002.
- [16] H. Suzuki, F. Ma, H. Izumi, O. Yamazaki, S. Okawa, and K. Kido, “Instantaneous frequencies of signals obtained by the analytic signal method,” *Acoustical Science & Technology*, vol. 27, no.3, pp. 163-170, 2006.
- [17] F.N. Fritsch, and R. E. Carlson, “Monotone Piecewise Cubic Interpolation,” *SIAM J. Numerical Analysis*, vol. 17, pp.238-246, 1980.
- [18] D. Delviniotis and G. Kouroupetroglou, “DAMASKINOS: The Prototype Corpus of Greek Orthodox Ecclesiastical Chant Voices”, in *Proc. Int. Conf. Crossroads / Greece as an Intercultural Pole of Musical Thought and Creativity*, Thessaloniki, Greece, pp.1-14, 2011.
<http://speech.di.uoa.gr/sppages/spppdf/DAMASKINOS.pdf>
- [19] D. Delviniotis, *Byzantine Chant analysis with signal processing techniques*, Ph.D. Thesis, University of Athens, 2002 (in Greek).
<http://thesis.ekt.gr/thesisBookReader/id/13220#page/130/mode/2up>
- [20] J.F. Michel, and R.D. Myers, “The effects of crescendo on vocal vibrato,” *J. Voice*, vol. 5, pp. 292-298, 1991.
- [21] J. Sundberg, “Acoustic and psychoacoustic aspects of vocal vibrato,” *Q. Prog. Status Rep.*, vol. 35, pp. 45-68, 1994. KTH Speech Transmission Laboratory.

Parameterization of the Byzantine Chant Ethos through Acoustic Analysis: from theory to praxis

Anastasia Georgaki
Music Department,
University of Athens
georgaki@music.uoa.gr

Achilleas Chaldaeakes
Music Department,
University of Athens
axaldaik@music.uoa.gr

Panagiotis Tzevelekos
Department of Informatics and
Communications,
University of Athens
taktzev@di.uoa.gr

ABSTRACT

In this paper we will analyze the notion of ethos in Byzantine music theory through a systematic parameterization of a vocal performance. In this direction we focus on the analysis of ethos from special pitch contours, melismatic micro-intervallic variations, intensity curves, timbral analysis, spectral variations and energy rates which indicate the emotional state of the performer (respecting the music notation and the meaning of the sacred text) trying to interpret the indicated mood by a concrete melismatic approach. More precisely in our approach we will attempt both structural and performance acoustic analysis as a first step in order to understand the different changes of the ethos (mood) in the macrostructure of the piece and by isolated phrases which indicate the functionality of the sang mode (echos)¹.

INTRODUCTION

Byzantine music holds a very important place as a living culture in Greece and abroad, and comprises two sides of the same coin with regard to modern Greek music: Byzantine music is the music of the Greek Orthodox Church and also the basis of the traditional popular music of the Greek people.² A number of studies have been carried out the last ten years in the field of music informatics and computational musicology in

order to analyze the special character of the notation system and the intervals, which is between the oral and the written tradition, as also to develop new tools for the education, processing and performance of the Byzantine music particularities [Delviniotis and al, 2008], [Zannos and Georgaki, 2006], [Mavromatis, 2005], etc.

¹ Echos” (Ἠχος) in Greek means "sound" in general. It acquired the
²Ethos” (ἦθος) is a Greek word meaning "moral character" that is used to describe the guiding beliefs or ideals that characterize a community, nation, or ideology. The Greeks also used this word to refer to the power of music to influence its hearer's emotions, behaviors, and even morals. The semiotics of 'music-ethos' at the ancient Greeks appears in its totality at the level of main components of music (modes, rhythms, genres, sonorous registers, the instruments utilized) and, therefore by the integrative signification of music. The Ancient Greek doctrine of ethos attributed ethical powers to music and claimed that

Copyright: © 2013 Georgaki et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

rousing or calming, that it produced [Anderson, Warren D. 1966] , [West,1992].

One of the major obstacles to appreciating Byzantine music is that extensive training is required in order to understand the special music notation, as also the way of pronunciation and singing of the sacred text towards comprehending the thoughts behind the musical work.

On the other hand, musical ethos (mood) in Byzantine music has always been left to the chanter's interpretation according to their education and origin. Byzantine music has a strong mathematical and theoretical background which is supposed to be connected to Ancient Greek music. Xenakis in his radical article 'Vers une *metamusique*' [Xenakis, 1970] observes through mathematical intervallic comparison that Byzantine music has best preserved the musical structures of ancient Greece by presenting out the four orders: tone, tetrachord and genera, scale and finally mode (echos), noting both similarities and differences from ancient Greek theory. Although Western music and culture have changed, the doctrine of ethos still holds significance as a piece of Western musical heritage and a model for the examination of music in society.³

1. ON BYZANTINE MUSIC *ETHOS* THROUGH *ECHOS*

Modes (Echos), Ethos and Melopoeia of the Byzantine Music Theory Music are terms being referred at Chrysanthos' *Great Theory of Music* where one can visually understand (philosophically and musicologically) their existing strong (internal and external) connection. It's obvious that we have to explore the sense of *Ethos* (that is the way "*the state of the soul is observed and corrected*") through the structure of *Melopoeia* (that is "*the power to create melos*"; the composer's "*poetic state*"; but in any case we can understand the differences between the existing Ethe (Chrysanthos names three of them, that is the aforementioned *Diastaltic*, *Systaltic* and *Hesychastic*) through *Modes*, i.e. under the specific

³ Byzantium involved mathematics very closely in art. This was in part a result of the rational basis of Byzantine aesthetics and the separation of the world of the mind – *noetos* -- from the material world, *aisthetos*. "The harmonies of number could be translated into geometric terms [...] Though both belonged to the world of *noetos*, both molded subordinate arts in *aisthetos* and the material world." It was this that led Michael Psellos in the 11th century to note the value of mathematics for philosophy since it linked abstract thought and material things [Gervase, 1963]

idea and most of all the unique *sound (echos)* of anyone of the eight in total Byzantine Music Modes. [Chrysanthos of Madyta, 1832/ 2010]

If we compare the notion “echos” (mode) to ethos (mood), we could say that the first one is a pure technical term which has a strict mathematical basis and the second one is based on the aesthetical interpretation of the performer, the understanding of the text and the perceptive attitude. Echos is musical system that relies on a fundamental tone and a set of melodic formulas built around it which are made from a defined set of scale steps that have uniquely tuned intervals, based on well-defined intrinsic information which cannot exist by itself in a composition, but has to be combined and coupled with rhythm, melopoieia, an eligible text, etc. Through the fusion of all the above, it has been stated that during the musical performance some members of a set of characteristics having emotional content emerge. These emotional characteristics, in their potential form, are grouped under the generic term ‘ethos’ of the ‘echos’. On the concrete music vocabulary the application of ‘echos equipped with ethos’ on a special composition gives a new approach which depends on the interpretation of the performer and its respect on the theoretical background of the mathematical basis of the intervals.

In this section of the paper, we propose a first approach on acoustic analysis solution to reveal firstly the semantic structure in Byzantine singing, such that users can gain insights into musical structure and understand better what ethos means by acoustic parameters taking as case study the troparion analyzed above by professor A. Chaldaiakis.⁴ In this analysis we propose to show the correlation between the ethos and the echos (mode) which appears in this original performance and decodify this intervallic architecture in the sung in order to explore the profound roots of on oral tradition which has its roots in the Pythagorean doctrine of ethos through scales. We will investigate also how accurate a performance is realized when observed within the context of two factors: how close are the experimental results to the theoretical scales and how consistent the chanter is within the piece. The experimentally derived scale is close to both theoretical scales considered in this paper, and the chanter is consistent throughout the piece.

We shall indicatively use a musical example based on the following well-known *Doxastikon Idiomelon* of the Vespers service.⁵

⁴ Achilleas Chaldaiakis is Associate Professor of Byzantine musicology at the Musicology Department, University of Athens and at the same time an active performer of Byzantine music.

⁵ Here is an English translation of the text of that poem: “Let us praise today the mystical trumpets of the Spirit, the God bearing Fathers, who sang a harmonious melody of theology in the midst of the Church: one Trinity, unchanging Essence and Godhead; the over throwers of delusion, the champions of the Orthodox, who ever intercede with the Lord that he have mercy”. We will examine the individual meanings of the above poem, searching for their Ethos and Aesthetics, using here a specific -extremely popular and widespread- melody of it, lately (during 20th century) composed by Konstantinos Pringos (+ 1964). The musical text of the said composition is taken here from the following musical edition, an own one of Konstantinos Pringos (Master Protosalt of the Great Church of Christ) himself (recently republished by Georgios N. Konstantinou): Musical Blister, 1st volume, Athens 2007

The intervallic environment of the said composition is a hard *chromatic* one, since its melody is composed in *second plagal mode*. The scale of that mode includes two identical disjunct tetrachords, the lower and the higher one (separated through a major tone, seemed like an ideal “musical bridge”), while each of them is developed by three (both ascending and descending) steps: semitone | augmented tone | semitone.

2. METHODOLOGY

In this case, we propose a methodology in order to parameterize three different states of the musical ethos (*diastaltikon*, *systaltikon* and *hesychastikon*) related to modes (echoi) and genera on this vocal performance of Byzantine music composition (troparion) through the observation of acoustic parameters. The visualized mood (ethos) is a clue with less ambiguity and subjectivity than verbal tags, because figures are drawn with expressive elements obtained from performance data.

- 1) Record Professor’s Chaldaiakis specific interpretation (troparion on hard chromatic genus) according the special notation system
- 2) Create an augmented score where we represent the changes of ethos according the performers’ statements on the function of the modes and genera.
- 3) Explore three different states of ethos through echos by analysing inside time and outside time the phrases:
 - a) Diastaltic ethos: “*tas salpigas tou pneumatos*” (trumpets of spirit) and “*melodisantas*” (sang) (Diatonic genus)
 - b) Systaltic: “*tas mystikas*” (secret) (chromatic genus-soft)
 - c) Hesychastic: “*Melos enarmonion*” (enharmonic genus)
- 4) Implement the prosodic feature extraction in order to get several features that describe performance model of the prosody (register, intensity level, speed, continuity)
- 5) Give a description the pitch contour analysis of this recorded fragments on the content of intervallic frequencies in order to explore the performer’s approach to the mathematical basis of the mode.
- 6) Explore the melismatic⁶ path of the *ethos* trough echos.
- 7) Other implementation details using the Praat Program
 - Sound and Textgrid Tiers of syllabus and words
 - Sound and raw pitch contour (Pitch Tier)
 - Sound and stylised pitch contour (Pitch Tier)

(Apostolic Ministry of the Church of Greece), pp. 185-187 [July’s feast].

⁶Melisma (from Greek – *melody*) is an expressive vocal phrase or passage consisting of several notes sung to one syllable. From the point of view of signal theory, melisma is a non-stationary signal, i.e., characterized by rapidly changing amplitude (intensity) and periodicity (pitch).

3. AUGMENTED SCORE OF *ETHOS* AND *ECHOS*

Our first approach is based on the augmented score where the visualization outline of the overall structure of the piece is described by special annotation of audio files that we have done in the PRAAT platform⁷ (open source software) by coloring the change of the mood of music from the underlying musical elements. We formulate the semantic structure into macro-level layer interactions, micro-level theme variations, and macro-micro relationship between themes and layers.

higher), dynamic level (louder– softer), speed (faster– slower), continuity (gradual– abrupt), and so forth”. [Juslin, Laukka, 2003]

In Byzantine music, the mood is indicated by the several modes and the genera annotated by the Great theory of music by Chrysanthos of Madyta where the melody is sung according the meaning of the sacred text show gratitude, admiration, pride, sadness, vigilance or anticipation, or else.

At the phrase *trumpets of the Spirit*, where an imitation of the sound of trumpets appears in melody, through a change in modality genus,

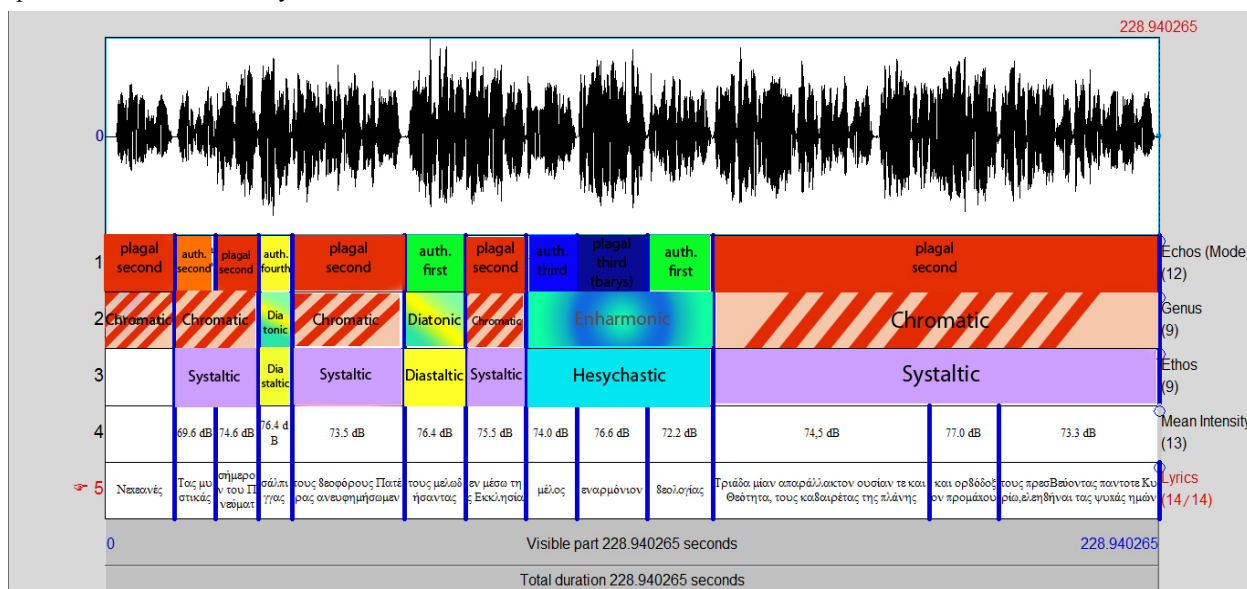


Figure 1. Augmented score of the music ethos (macro-structure).

In this augmented score we can observe 5 different layers which indicate the ethos (mood) variations closely related to the modes (echos), the genera (genus)⁸, the mean intensity of the voice and the lyrics and according to the notion of ethos that has been presented in the first part of this article.

4. ANALYSIS ON *ETHOS*

As we have already mentioned musical mood (ethos) is subjective and is usually left to the performer's and listeners' interpretation. In our approach the notion of 'ethos', expressed via 'statistical parameters', approaches Juslin's acoustic cues: "When listeners or critics describe music as sad, happy, angry and so on, [these states] are delineated by the action of what I have called the 'statistical parameters'. These aspects of sound vary in amount or degree – for example, register (lower–

whilst the chromatic second plagal mode is changing here to the *diatonic four authentic mode* (the so-called *Hagia*). In addition, an interpreter's attention and a researcher's mind cannot avoid focusing on the following five remarkable points of the entire composition on the one hand, two **Diastaltic** Ethos changes, the first one based on **G** (*trumpets of the Spirit*) and the second on **A** (*tous melodisantas*), through a transposition and alteration from the initial second plagal to fourth and first mode respectively; this way, *majesty and virile disposition of the soul*, along with *heroic deeds are expressed*.

Systaltic environment *Ethos* of specific composition includes a low-tone additional *Systaltic* musical passage at the beginning (*the mystical*), balancing with another one also additional *Systaltic* musical mention at its end (*the champions of the Orthodox*), a high-tone one in this case; this way, *the soul is driven to humility, fitted to compassions*⁹. At the phrase *the mystical*, with which compositions starts, precisely by a surprising overthrow: specifically, the melody here, turning from *plagal* to the *authentic second mode*, develops below the base of the initial *plagal* mode, an unusual phenomenon for *plagal* modes structure and identity; music on that phrase (*the mystical*) seems like whispering a secret that should remain an unknown piece of information, even if it is sung through a loud voice.

⁷ <http://www.fon.hum.uva.nl/praat/>

⁸ DIATONIC)Nη-204-Πα-161 - Βου - 133 - Γα - 204 - Δι - 204 - Κε - 161 - Ζω' - 133 - Νη'
CHROMATIC (soft). Nη - 133 - Πα - 232 - Βου - 133 - Γα - 204 - Δι - 133 - Κε - 232 - Ζω' - 133 - Νη'
- CHROMATIC (hard). Πα - 90 - Βου - 337 - Γα - 71 - Δι - 204 - Κε - 90 - Ζω' - 337 - Νη' - 71 - Πα'
- ENHARMONIC. Γα - 204 - Δι - 204 - Κε - 90 - Ζω' - 204 - Νη' - 204 - Πα' - 204 - Βου' - 90 - Γα'.

⁹ Cf. *Great Theory of Music* by Chrysanthos of Madytos, *ibid.*, p. 180.

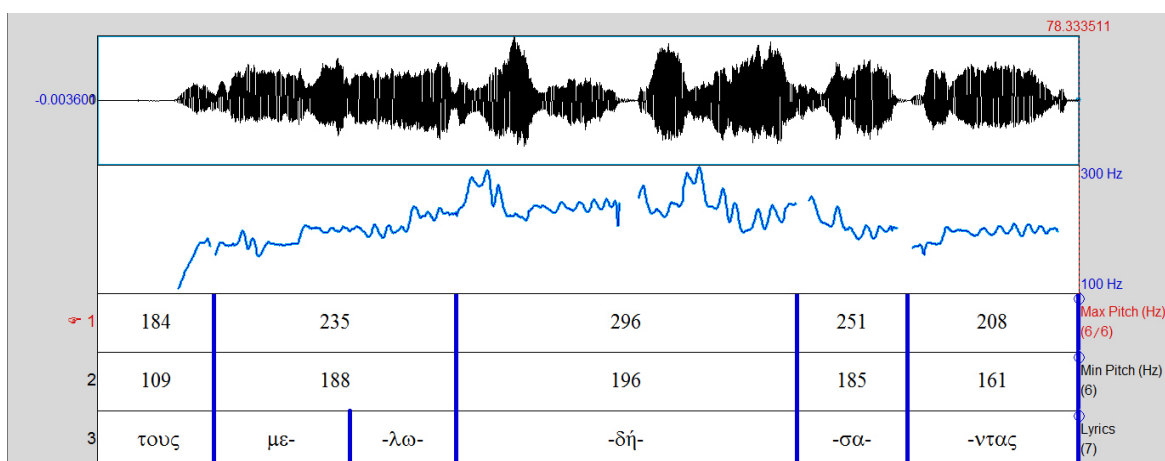


Figure 2. Contour analysis of “melodisantas”.



Figure 3. Fragment of the score in byzantine and western notation.

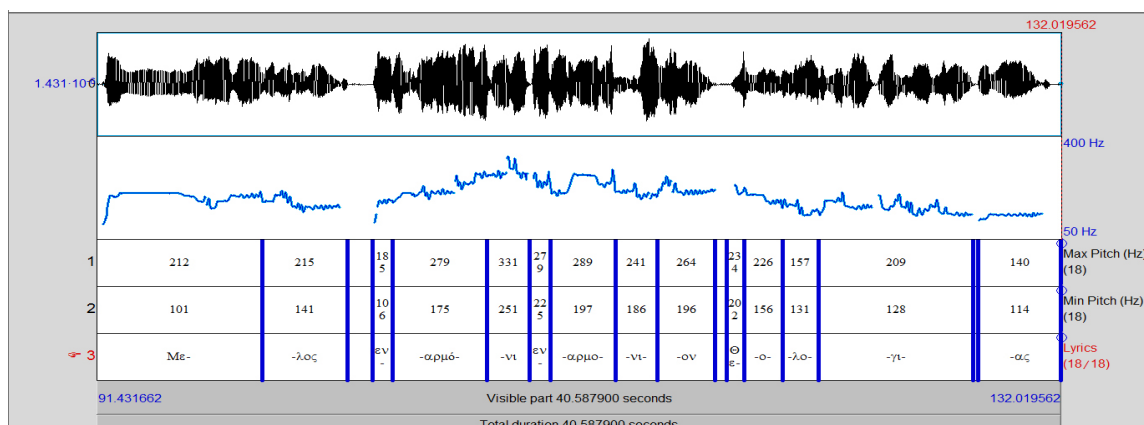


Figure 4. Melodic contour of the Phrase “harmonious melody” (melos enharmonion).

Hesychastic is added (undoubtedly the most indicative of the entire composition, developed on the phrase *a harmonious melody of theology*), consecutively transferring the melody to the third and “barys” modes; this way, *serenity of the soul and a state of freedom and peacefulness are expressed*, feelings totally appropriate for the specific part of this *hymn and song of praise*, a part which is clearly a *paean*¹⁰. At the verb *sang*, where melodic development describes the sense of chanting (included in the meaning of the said verb) using also the

diatonic intervals of the brilliant *first mode*. At the phrase *a harmonious melody of theology*, one can easily observe that at this specific point of the composition the composing idea is developed word by word, using three different modes (a separate one for each word), in order to help any audience to understand immediately what the entire phrase means: it is like *theology* (i.e. the way we speak about religion and God himself), a meaning described here through -plausible- the *first mode*, being a *melody* (and on the word “melody” the *third mode* evolves) and especially a *harmonious* one (composer shows the latter meaning -*harmonious*- using an enharmonic genus mode, the so-called *barys*).

¹⁰ *Ibid.*, pp. 180-181^{§398}.

5. ACOUSTIC ANALYSIS OF VOCAL PERFORMANCE

In the present research, acoustical analysis focuses on four different domains for the parts “*trumpets, who sang, the mystical, the God-bearing Fathers, the champions of the Orthodox, a harmonious melody of theology*” in order to specify similarities and differences regarding the way the singer approaches ethos: pitch domain, intensity domain, temporal domain, and spectral domain. On the pitch domain, minimum pitch, maximum pitch, and median pitch are derived, as well as the entire pitch contour.

On the intensity domain, maximum intensity, mean intensity of the voiced parts weighted on duration, and the entire intensity contour are derived. On the temporal domain, the duration of each part as a percentage of the whole piece, and the velocity factor as the number of note changes per second are derived. On the spectral domain, the spectral centroid and the fourth formant are derived. Results are shown in the following Table 1. Using the variables shown, cluster analysis is performed using SPSS software. Values were normalized and hierarchy cluster analysis option for 3 clusters and was used, resulting to the dendrogram shown in Fig.5.

6. ACOUSTIC ANALYSIS DISCUSSION

In the parts *trumpets* and *sang*, median pitch values are relatively high (>200 Hz), while the mean intensity of the voiced parts is also high (>75 dB). Spectral centroid values are the highest in the entire piece (>2000 Hz), while duration and velocity values denote fast pace. These findings are in accordance with the festive and majestic character of the text, since fast music rate, high sound level, much high-frequency energy and high pitch are connected to emotional states of happiness and joy. Both parts link with the diastaltic ethos.

In the part *harmonious melody*, median pitch lies in the middle (~200 Hz) and melody covers a wide pitch range. Intensity is relatively low (~74 dB), the fourth formant is lowered in comparison to the diastaltic parts, and the spectral centroid is found in the middle frequency area (~1900 Hz). This part has the longest duration, and a slow note-changing rate. One would say that the parameter values lie in their middle-low area, in line with the serenity, balance and tranquility the text expresses. This part links with the hesychastic ethos.

In the *mystical* part, melody moves through its lowest pitch, intensity, spectral centroid and temporal rate, underlining the secretive feel of the systaltic ethos. It should be noted however that the fourth formant lies high (~3900 Hz), marking the sound, though low in volume, rich and audible. The *God-bearing Fathers* part is quite similar overall, with most values slightly transferred toward the middle area. The *champions of the Orthodox* part seems most intriguing. Even though pitch and intensity move in their highest values (255 Hz median pitch, 342 Hz maximum pitch, 77 dB mean intensity),

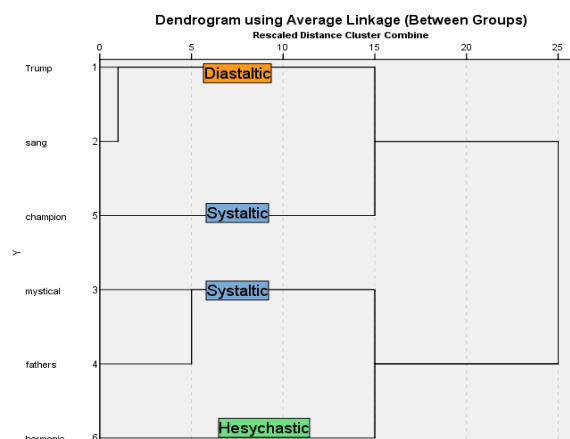


Fig.5.Dendrogram using average linkage

spectral features lie in the middle-low frequency area (spectral centroid 1700 Hz, F4 3550 Hz), offering a different audible result compared to the parts *trumpets* and *sang*.

This difference is also implied in the text. The lyrics speak of a heroic act, but this time through humility and modesty. Therefore, this part is expressed with the systaltic ethos.

The results from the clustering algorithm agree with the former discussion. The parts *mystical* and *sang* are easily grouped together (*diastaltic ethos*), and are later grouped with the *champions* part (high pitch/intensity, systaltic ethos). *Harmonious* part lies alone (*hesychastic ethos*) and is later grouped with the *mystical* and *God-bearing fathers* group (low pitch, systaltic ethos). It seems that diastaltic and hesychastic parts are at the two opposite ends, while systaltic parts lie in the middle, moving towards each end according to the part's particularities.

7. FURTHER RESEARCH

We hope that this methodology on the analysis of complementary feedback with visual data will help young singers and chanters to understand their own performances and interpret the annotated ethos according to the mathematical structure of the echos and the further indications of the score.

8. CONCLUSIONS

The philosophy behind performance acoustic analysis is to display qualitative musical characteristics by acoustic cues in order to get more information than the score itself indicates. In our case, we propose a methodology in order to parameterize musical ethos (mood) related to modes (echoi) and genera on a performance of Byzantine music composition (troparion) through the observation of acoustic parameters in order to spot the difference between the three different states of ethos (mood).

9. REFERENCES

- [1] Anderson, Warren D. (1996): *Ethos and Education in Greek Music*. Cambridge, Mass: Harvard University Press, 1966.
- [2] Chrysanthos of Madytos, *Great Theory of Music* (1832) (translated by Katy Romanou), New York 2010 (The Axion Estin Foundation, New Rochelle), pp. 180-181.
- [3] Delviniotis D., Kouroupetroglou G., and Theodoridis S. (2008): "Acoustic Analysis of musical intervals in modern Byzantine Chant Scales", *Journal of the Acoustical Society of America* (JASA), Vol. 124, No 4, pp. EL262-EL269
- [4] Gervase Mathew (1963): *Byzantine Aesthetics*, London ("The Mathematical Setting" p.23)
- [5] Juslin P., Laukka P. (2003): "Communication of emotions in vocal expression and music performance. Different channels, same code?" in *Psychological 2003*, Vol. 129, No. 5, 770–814.
- [6] Mavromatis, P. (2005): *The Echoi [Modes] of Modern Greek Church Chant in Written and Oral Transmission: A Computational Model and Its Cognitive Implications*, Ph.D. Dissertation. Eastman School of Music, University of Rochester.
- [7] Sundberg, J. (2001): "Expression in music: Comparing vocal and instrumental performance". Karevold, H Jörgensen, IM Hanken, E Nesheim, eds, *Flerstemmige innspill 2000*, NMH-publikasjoner 2000:1, Norges Musikkhøgskole, Oslo.
- [8] West, M.L. (1992): *Ancient Greek Music*. Oxford: Oxford University Press.
- [9] Xenakis, I. (1970): "Towards a metamusic", *Tempo new series*, issue 93, June 1970, pp.2-19.
- [10] Zannos I., Georgaki A., Delviniotis D., Kouroupetroglou G., (2006): "Real time control of Greek chant synthesis", *SMC06 proceedings, Marseille*.

Part	Min Pitch (Hz)	Max Pitch (Hz)	Median Pitch (Hz)	Duration %	Velocity (notes/sec)	Mean Intensity (dB)	Spectral Centroid (Hz)	Fourth Formant (Hz)
<i>trumpets</i>	141	276	214,7	3,4	1,81	76,45	2113	3810
<i>who sang</i>	109	296	204,0	6,2	2,35	76,44	2032	3842
<i>the mystical</i>	85	242	166,7	8,5	1,05	72,49	1416	3932
<i>the God-bearing Fathers</i>	114	266	177,9	11,4	1,19	73,83	1887	3846
<i>the champions of the Orthodox</i>	148	342	255,6	7,1	1,58	77,28	1698	3555
<i>a harmonious melody of theology</i>	102	331	197,0	19	1,11	73,99	1916	3653

Table 1. Acoustic cues of vocal performance.

INTERFERENCE-FREE OBSERVATION OF TEMPORAL AND SPECTRAL FEATURES IN “SHOUT” SINGING VOICES AND THEIR PERCEPTUAL ROLES

Hideki Kawahara

Wakayama University, Japan

kawahara@sys.wakayama-u.ac.jp

Masanori Morise

University of Yamanashi, Japan

mmorise@yamanashi.ac.jp

Ken-Ichi Sakakibara

Health Sciences Univ. of Hokkaido, Japan

kis@hoku-iryo-u.ac.jp

ABSTRACT

A new set of algorithms are introduced to represent and manipulate physical features found in expressive singing performances. The proposed representations are simple in both conceptual as well as computational aspects. This simplicity makes interpretation and control of the representations straightforward. First, a new fundamental frequency (F0) extractor revealed rapid and strong F0 modulation around 70 Hz (in terms of modulation frequency) in a “shout” performance of a Japanese POP song. The F0 extractor is based on higher-order waveform symmetry and has finer temporal resolution than conventional methods. It is followed by a refinement procedure based on an interference-free representation of instantaneous frequency. Second, in interference-free representation of power spectrum revealed a synchronized spectral modulation to the rapid F0 modulation in “shout” performance. Third, one-third octave level differences between plain performances and “shout” performances were observed. A set of manipulated singing voices were synthesized by manipulation of these features using an auto-regressive (AR) model-based notch-filtering in the modulation frequency domain and one-third octave FIR equalizers. The stimuli were synthesized using an extended version of a speech analysis, modification and resynthesis framework, TANDEM-STRAIGHT. Subjective tests illustrated that all these three features are effective contributing factors of “shout” impression. Physiological mechanism and possible applications are also discussed.

1. INTRODUCTION

Expressive singing performance sometimes deviates from beautiful classical voices such as Bel-Canto [1, 2] but still is very impressive and moving. Singing processing systems have to be able to analyze, replicate and control such voices to explore full range of artistic expressions. It also is indispensable for speech science to understand them, because such deviations represent non- and para-linguistic aspects of speech, which make speech a richer media for communications than texts [3].

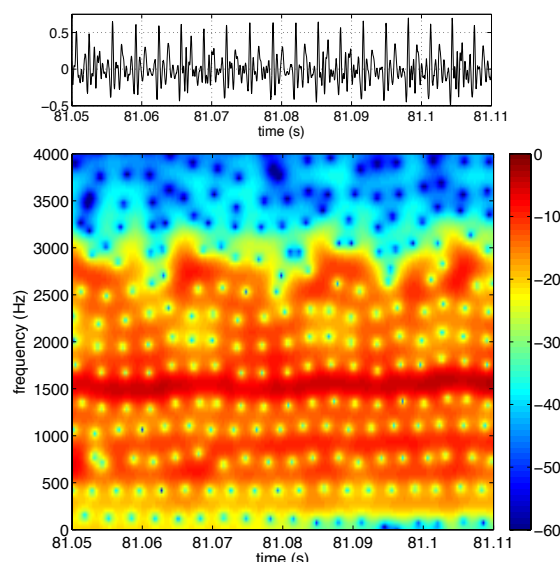


Figure 1. Waveform and spectrogram of an excerpt from an expressive (“shout”) singing performance.

In spite of important roles they play, signal processing aspects of expressive singing voices were not explored extensively [4, 5], possibly because of technical difficulties in analyzing such signals. To analyze such difficult expressive singing voices, a set of interference-free representations of power spectrum [6, 7] and instantaneous frequency [8] are introduced with a temporally-fine F0 extractor based on higher-order waveform symmetry [9].

This article is organized as follows. First, an excerpt of an expressive performance of a Japanese POP song is used to illustrate difficulties in analysis based on conventional analysis procedures. Second, a set of above mentioned interference-free representations are briefly introduced. Third, an F0 extractor based on higher-order waveform symmetry is introduced. The following sections illustrate specific features of “shout” expression and introduce procedures to manipulate those features. Finally, subjective test results using the manipulated test singing stimuli (linked to our Web [10]) are analyzed and discussed.

2. EXPRESSIVE SINGING: AN EXAMPLE

Figure 1 shows an example of “shout” voice waveform and the corresponding spectrogram excerpted from a performance of a Japanese POP song, “Ride,” specially prepared

for research on singing expressions. Singing performances of this song in several different styles, by a Japanese professional male singer, were recorded in 16 bit and 44100 Hz sampling format. The performance in his own style consisted of a strong vocal expression regarded as “shout” and is used in this research. Also the performance without “shout” expression (named “plain”) is used for reference. Length of each performance is 157 s. The excerpt shown in Fig. 1 is 60 ms in length and located at 81.05 s from the beginning. The sung vowel is /a/. The length of the window for spectrogram is 9 ms and the frame update period is 0.25 ms.¹ In this setting, the spectrogram has isometric resolution in terms of the fundamental frequency (about 310 Hz, in this case) and the fundamental period.

The spectrogram indicates that F0 is frequency modulated three times in this 60 ms display. Note that this modulation is not synchronized with the vocal fold vibration corresponding to F0. It also indicates that spectral envelope is modulated synchronously with this F0 frequency modulation. This spectral modulation is salient around 2500 Hz to 3000 Hz region. These features illustrate difficulties for applying conventional analysis procedures. For F0 modulation, conventional F0 extractors [12–14] fail to track this rate of F0 change [5]. For spectral shape modulation, virtually regular dips in conventional spectrogram strongly interfere with spectral envelope estimation. Without physically accurate parameter extraction, analysis and manipulation of such features cannot be implemented. These motivated this proposal of a new set of analysis, modification and resynthesis procedures.

3. INTERFERENCE-FREE REPRESENTATIONS

This periodic interference can be cancelled by using F0 information. This section introduces two representations free from interferences caused by the signal periodicity.

3.1 Power Spectrum

Interference-free power spectrum is calculated by using a two staged procedure. The first stage cancels out temporal variations due to periodicity. The second stage eliminates periodic variations in the frequency domain.

3.1.1 TANDEM: Temporally Stable Power Spectrum

If the effective pass-band of the time window function used in short time Fourier analysis covers two harmonic components of the signal, the temporal variation of the power spectrum is a sinusoid. The frequency of the sinusoidal variation is equal to F0, and can be canceled out by adding the anti-phase signal.

Let T_0 represent the fundamental period. The temporally stable power spectrum $P_T(\omega, t)$ (TANDEM-spectrum) is defined as the average of two short time power spectra cal-

culated at two locations a half fundamental period apart [6].

$$P_T(\omega, t) = \frac{P(\omega, t + \frac{T_0}{4}) + P(\omega, t - \frac{T_0}{4})}{2}, \quad (1)$$

where $P(\omega, t)$ represents a short time power spectrum. The variables ω and t represent angular frequency and time respectively.

3.1.2 STRAIGHT: Spectral Envelope Recovery

The next stage is spectral variation suppression. By introducing moving average on the frequency axis using a rectangular smoother of a width F_0 , eliminates periodic variations on the frequency axis, since the period is equal to F_0 (f_0 represents the value of F_0) and yields a frequency-interference-free power spectrum $P_I(\omega, t)$.

$$P_I(\omega, t) = \frac{1}{\omega_0} \int_{-\frac{\omega_0}{2}}^{\frac{\omega_0}{2}} P_T(\lambda, t) d\lambda, \quad (2)$$

where $\omega_0 = 2\pi f_0$.

However, this simple moving average smears spectral details because time windowing already introduced spectral smoothing. This excessive smoothing problem can be solved by introducing a concept called consistent sampling [15]. For assuring positivity of the power spectrum, this compensation and the moving averaging operations are combined and executed in the cepstrum domain to yield a power spectral envelope, free-from periodic interferences and smearing, $P_{ST}(\omega, t)$ (STRAIGHT-spectrum).

Whole procedures are approximately implemented as liftering in the cepstrum domain [16]. The following equation represents the whole process.

$$P_{ST}(\omega) = \exp(\mathcal{F}[g_A(\tau)g_C(\tau)C_T(\tau)]), \quad (3)$$

where $C_T(\tau)$ represents the cepstrum calculated from the TANDEM spectrum. The symbol \mathcal{F} represents Fourier transform. The lifter $g_A(\tau)$, $A \in \{1, 2\}$ is F0 adaptively designed to eliminate periodic variations due to the harmonic structure. The lifter $g_1(\tau)$ is used in TANDEM-STRAIGHT implementation and equivalent to (2).

$$g_1(\tau) = \frac{\sin(\pi f_0 \tau)}{\pi f_0 \tau} = \mathcal{F}[h_1(\omega)], \quad (4)$$

$$h_1(\omega) = \begin{cases} 0 & |\omega| \geq \frac{\omega_0}{2} \\ \frac{1}{\omega_0} & \text{otherwise} \end{cases}. \quad (5)$$

The lifter $g_2(\tau) = (g_1(\tau))^2$ is the corresponding lifter of the smoothing function used in legacy-STRAIGHT implementation [17].

The lifter $g_C(\tau)$ is the approximate implementation of the compensation digital filter, which was initially designed based on the consistent sampling. In the latest implementation, numerically tuned filter is used. The tuning was conducted to reduce spectral smearing around formant peaks [18], based on a set of simulations using vocal tract area functions of eleven English vowels [19] and a glottal waveform model [20] with random parameter perturbations.

$$g_C(\tau) = \tilde{q}_0 + 2\tilde{q}_1 \cos\left(\frac{2\pi\tau}{T_0}\right). \quad (6)$$

¹ The spectrogram is calculated using a time windowing function with very low (less than -90 dB) side lobe levels and relatively fast (-18 dB/oct) side lobe decay rate. It is the twelfth item of the Table II in Nuttall's paper [11]. This Nuttall window is different from the commonly used Nuttall window, for example, “nuttallwin” in Matlab.

3.2 Instantaneous Frequency

Instantaneous frequency [21] is a relevant concept for representing time varying signals. It is defined by the time derivative of the phase of the signal. Flanagan's equation [22] provides an efficient algorithm which does not require phase unwrapping. The equation also reveals the source of an annoying behavior of instantaneous frequency, periodic discontinuities. These discontinuities, caused by interference between neighboring harmonic components, are canceled out by calculating power spectrum weighted average of instantaneous frequencies obtained at two locations a half fundamental period apart [8].

Let $\omega_i(\omega, t)$ represent instantaneous frequency calculated using a time windowing function positioned at t . Averaged instantaneous frequency $\omega_{Ai}(\omega, t_1, t_2)$ calculated using two windows positioned at t_1 and t_2 is defined by the following equation:

$$\omega_{Ai}(\omega, t_1, t_2) = \frac{P(\omega, t_1)\omega_i(\omega, t_1) + P(\omega, t_2)\omega_i(\omega, t_2)}{P(\omega, t_1) + P(\omega, t_2)}. \quad (7)$$

By substituting $t - \frac{T_0}{4}$ and $t + \frac{T_0}{4}$ with t_1 and t_2 , the denominator of (7) yields $P_T(\omega, t)$ defined in (1), which does not have temporally varying components. For numerator, time varying components are also cancelled out by this setting. Refer to the article [8] for proof and implementation.

4. TEMPORALLY-FINE F0 EXTRACTOR

An efficient and accurate F0 extractor is indispensable for these interference-free representations, because of their F0-adaptive nature. The F0 extractor also has to have finer temporal resolution capable of tracking the rapid frequency modulation illustrated in Fig. 1.

A two staged procedure is also introduced here [9]. First stage selects the best candidate based on waveform symmetry. The next stage refines the initial estimate using the temporally stable instantaneous frequency mentioned above.

4.1 F0 Extraction based on Waveform Symmetry

The first stage of the procedure measures the repetition interval of the fundamental component. It requires isolation of the fundamental component only. LPF-based selection of the fundamental component is employed in the proposed method, because it has two times finer temporal resolution than that of BPF based selection.

To determine the relevant cutoff frequency of the LPF,² an index representing symmetry of the filtered waveform is introduced. This symmetry index is calculated for each of a set of LPFs with log-linearly allocated cutoff frequencies and used to select the best filter for measuring the interval.

Figure 2 shows reference points used to define symmetry index. The point B is the temporal mirror image of C . The temporal deviation from symmetry is represented by the temporal distance between points A and B normalized

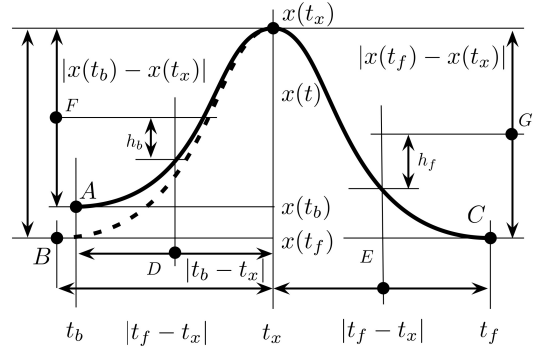


Figure 2. Definition of reference points for measuring deviation from symmetry. Dashed line shows the mirror image of the latter half cycle. Temporal position t_x represents the maximum position, and t_b and t_f represent the temporal positions of the preceding and the following extrema. For details, refer to the article [9].

by the distance between C and A the fundamental period (candidate). The amplitude deviation from symmetry is represented by the level difference between points A and B normalized by the average amplitude of the waveform. The higher-order deviation from symmetry is represented by the distances from the average at the midpoints (h_b and h_f in Fig. 2). These individual distances are combined using Minkowski distance and negative exponentially mapped into the interval $[0, 1]$ to yield the symmetry index (1 represents perfect symmetry). It was found that introduction of higher-order symmetry effectively reduces selection errors of the best LPF when calculating this symmetry index [9].

This first stage can track the rapid frequency modulation of F0. For example, when F0 is 310 Hz, the -3 dB frequency of the modulation frequency transfer function of this initial F0 extraction stage is about 120 Hz. In the current implementation, similar to the example section, a set of Nuttall windows are used as the LPFs' impulse response. They are covering from 32 Hz to 1000 Hz by allocating six LPFs for each octave.

4.2 F0 Refinement based on Interference-free Instantaneous Frequency

The initial estimate of F0 value is refined using the interference-free instantaneous frequencies of the harmonic components. In the current implementation, refinement is executed two times. In the first iteration, the first and the second harmonic components are used. In the final iteration, from first to sixth harmonic components are used to update the estimate. The -3 dB frequency of the modulation transfer function of this second stage is about 80 Hz when F0 is 310 Hz. This cutoff frequency clearly outperforms conventional F0 extractors [9].

This F0 extractor runs faster than realtime using usual PCs. For example, the initial estimation stage requires 0.6 s to process 3.3 s signal (44,100 Hz sampling, 1 ms frame rate), and the refinement stage requires 2.2 s for the same signal (OS X 10.8.2, PowerBook Pro 2.6 GHz Intel Core i7 16GB memory and MATLAB R2012a).

² A special version of the Nuttall's window, which is mentioned in the foot note of section 2, is used as the impulse response of the LPF. Low side-lobe levels and fast decay of side-lobe levels are important.

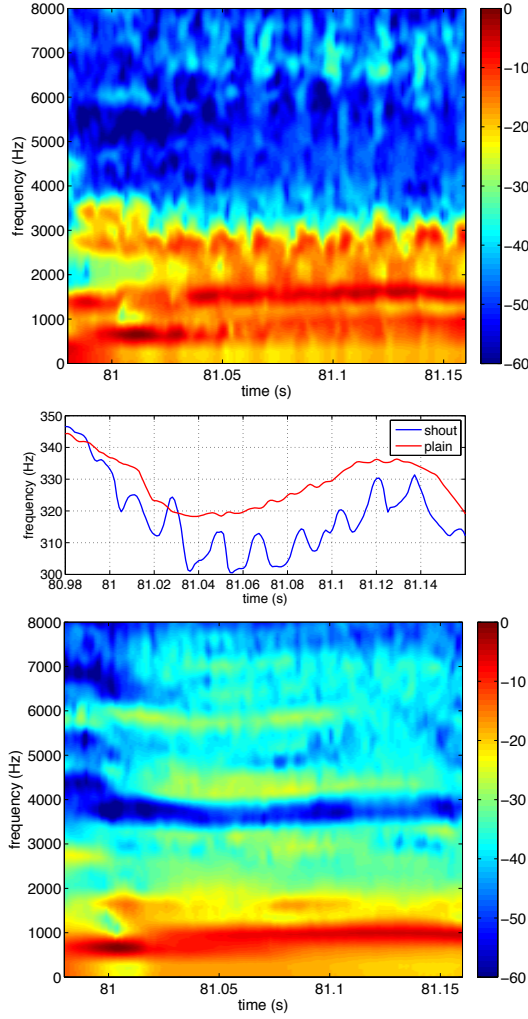


Figure 3. Interference-free spectrograms of “shout” performance (top plot) and “plain” performance (bottom plot). The center plot shows F0 trajectories of two performances. Frame update rate is 1 ms in all plots.

5. FEATURES

Using the procedures introduced in the previous sections, comparative study between “shout” and “plain” performances was conducted. Each performance was subdivided into common thirty two musical “phrases” and analyzed.

Figure 3 shows interference-free spectrograms and F0 trajectories of two different performances. The displayed range is three times wider than Fig. 1 in the time domain and two times wider in the frequency domain. The displayed region is located at the final part of the 17th phrase spanning from 79.457 s to 81.662 s. The lyrics of the phrase is “zaregoto mo tsurai” (“joking does not ease my pain” in English). The abrupt spectral change around 81.00 s corresponds to the consonant /r/.

These interference-free spectrograms clearly show contrasts between “shout” and “plain” performances. The most salient feature of “shout” is the periodic spectral envelope variation around 2500 Hz. The variation is synchronized with the fast F0 frequency modulation found in the center plot. Also, the spectral envelope level from 2000 Hz to

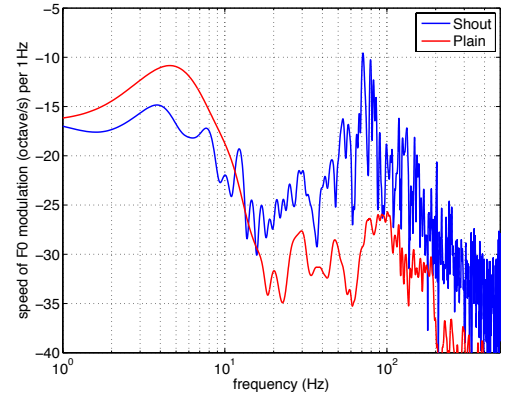


Figure 4. Power spectra of the differentiated logarithmic F0 trajectories with “shout” (blue line) and “plain” (red line) expressions.

3000 Hz of “shout” is higher than that of “plain” performance. Visually, the spectral valley around 4000 Hz found in “plain,” which is caused by the branching of the vocal tract (pyriform fossa) [23], is striking and disappears in “shout” performance.

In the following subsections, these three features (F0 frequency modulation, spectral envelope modulation, and spectral level difference) are quantitatively investigated. The valley feature is not investigated in this article, since it does not have significant contribution on impression of “shout.”

5.1 F0 Frequency Modulation: feature “Q”

5.1.1 Power Spectrum

First, rapid F0 modulation is analyzed in terms of power spectrum of the modulation, $P_S[k]$. Differentiated signals of the logarithmic F0 to base 2, $l_0[n]$, are used in the analyses since our (musical) pitch perception is approximately log-linear. Differentiation is introduced to make the signals zero-mean.

$$l_0[n] = \frac{\log_2(f_0[n+1]) - \log_2(f_0[n])}{t[n+1] - t[n]}, \quad (8)$$

$$P_S[k] = \left| \sum_{m=0}^{N_S-1} w[m] l_0[m + b_S] \exp\left(-j \frac{2\pi k m}{M}\right) \right|^2, \quad (9)$$

where $f_0[n]$ represents the fundamental frequency at n -th frame and $t[n]$ represent the corresponding time of the n -th frame. The number N_S represents the length of the windowing function $w[m]$ of the S -th voiced segment of the F0 trajectory. The initial frame index of the segment is represented by b_S . A discrete frequency index k corresponds to the modulation frequency $f[k] = k f_s / M$, where f_s represents the analysis frame update rate (frequency in Hz) and M represents the buffer length of discrete Fourier transform. A set of contiguous $P_S[k]$ are summed to yield $P_p[k]$, which represents the modulation power spectrum of a larger musical unit “phrase” that consists of several bars. The power spectrum is calibrated to represent speed of logarithmic F0 change in terms of octave/s per 1 Hz.

Figure 4 shows the power spectra of the differentiated logarithmic F0 trajectories. The figure illustrates that the

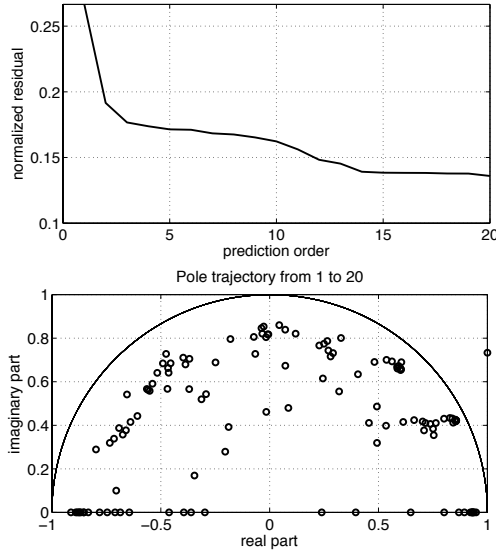


Figure 5. AR model fitting with the order from 1 to 20. Residuals (upper plot) and estimated poles (lower plot).

expressive singing consists of strong modulation power around 70 Hz. In this case, the peak magnitude of the modulation in “shout” performance is over 20 dB higher than that of its “plain” counterpart.

5.1.2 AR Model Analysis

Since resonant-like spectral peaks were found in modulation power spectra of “shout” expression, autoregressive (AR) model [24, 25] parameters $\{a[k]\}_{k=1}^K$ are estimated from the auto-covariance coefficients $r[n]$ calculated by inverse Fourier transform of $P_p[k]$.

$$r[n] = \frac{1}{M} \sum_{k=0}^{M-1} P_p[k] \exp\left(j \frac{2\pi kn}{M}\right), \quad (10)$$

$$\mathbf{a} = \mathbf{R}^{-1} \mathbf{r}, \quad (11)$$

where vector notations of the AR coefficients $\{\mathbf{a}\}_k = a[k]$, auto-covariance coefficients $\{\mathbf{r}\}_k = r[k]$, and a Toeplitz matrix $\{\mathbf{R}\}_{k,l} = r[|k-l|]$ are introduced. Frequencies and bandwidths of the resonant peaks are calculated from the roots $\{z_k\}_{k=1}^K$ of the following polynomial equation.

$$1 - \sum_{k=1}^K a[k] z^{-k} = 0, \quad (12)$$

where the imaginary part and the real part of the complex number z_k provide the frequency and the bandwidth respectively.

Figure 5 shows AR model fitting results. The upper plot shows the normalized residuals as a function of the model order. The lower plot shows extracted pole locations, which are overlaid. The model orders from 1 to 20 are used. The salient spectral peak in “shout” corresponds to the closest poles to the unit circle in the lower plot of Fig. 5. For order 20, an unstable pole is estimated suggesting that the equation is ill-conditional for the order. AIC of this analysis showed local minima at orders 15 and 18. For the sake

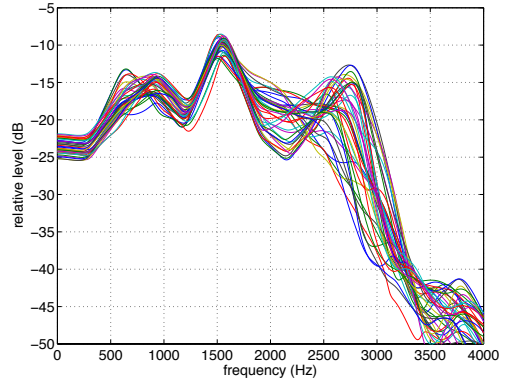


Figure 6. Spectral envelope modulation in “shout.” STRAIGHT-spectra from 81.06 s to 81.10 s are overlaid.

of parsimony, the order 15 is selected here. Substituting this best fitting model parameters in (12) yields poles. The pole frequency and bandwidth, which correspond to the most salient power spectral peak, are found 73.1 Hz and 3.67 Hz respectively.

5.2 Spectral Shape Modulation: Feature “F”

Figure 6 shows STRAIGHT-spectra from 81.06 s to 81.10 s in every 1 ms. It illustrates strong modulation of the third formant frequency and bandwidth. Differentiated logarithmic octave band level modulation in the voiced segments are analyzed in terms of amplitude modulation power spectra. The results revealed similar peak around 70 Hz, reflecting synchronization³ with the F0 frequency modulation. This variation of the third formant may reflect the mechanical vibration of the supra-glottal structures [4]. This is consistent with the AR model mentioned before, since this deformation results in periodic variation of the acoustic impedance and modulates the vocal fold vibration [27].

5.3 Octave-Band Level Difference: Feature “E”

Phrase-wise long-term spectra $L_p[k]$ are calculated from STRAIGHT spectra $P_{ST}(f, t)$ to illustrate the global spectral shape difference between “shout” and “plain” performances.

$$L_p[k] = 10 \log_{10} \left[\frac{1}{N_p} \sum_{m \in S_p} \int_{f_L[k]}^{f_H[k]} P_{ST}(\lambda, t[m]) d\lambda \right], \quad (13)$$

where S_p represents a set of frame indices m of the voiced segments in the p -th musical “phrase” and N_p represents the total number of voiced frames in the p -th “phrase.” The frequencies $f_L[k]$ and $f_H[k]$ represent the lower and the higher frequency boundary of the k -th frequency band, respectively. In the current analysis, one-third octave bands are used.

Figure 7 shows the one-third octave band level difference between “shout” and “plain” performances. In this plot, dB band levels of the “plain” performance is subtracted from

³ This synchronization depends on types of “shout” and disappears in extreme cases [26].

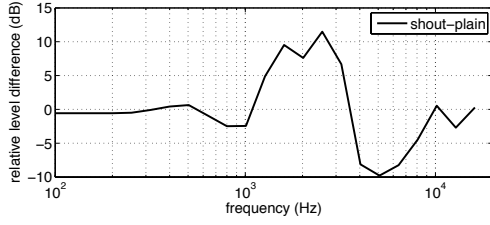


Figure 7. Difference of the one-third octave band levels between “shout” and “plain” performances.

that of the “shout” performance. This difference is normalized to be zero-mean. The positive level region around 2 kHz corresponds to the salient level enhancement shown in the spectrogram of “shout” expression in Fig. 3.

6. FEATURE MANIPULATION TESTS

Simple resynthesis of the “shout” performance using the analysis results by the extended temporally-fine TANDEM-STRAIGHT procedure yielded perceptually equivalent reproduction of the “shout” impression. This makes it possible to evaluate contributions of the above mentioned three features by testing resynthesized performance using all possible manipulation patterns of these features. In this test, only “suppression of feature” is used as the manipulation.

6.1 Procedures of Feature Manipulation

In this section, procedures to suppress three features individually are introduced.

6.1.1 F0 Frequency Modulation: feature “Q”

The inverse digital filter $Q(z)$ is designed to cancel the complex conjugate poles z_k, z_k^* associated with the peak around 70 Hz, that is estimated by the AR model. It is cascaded to the 2nd order critical damping LPF $T(z)$ consisted of complex conjugate poles z_q, z_q^* with the same nominal frequency to yield a notch filter. In addition to this notch filter, the following bilinear equalizer $M(z)$ is cascaded to adjust level differences in the higher modulation frequency region (> 20 Hz) after notch-filtering.

$$|M(z)| = \left| \frac{1 - r_L}{1 - r_H} \right| \left| \frac{1 - r_H z^{-1}}{1 - r_L z^{-1}} \right|, \quad (14)$$

where r_L and r_H are determined to adjust the lower and higher corner frequencies.

The upper plot of Fig. 8 shows the gain of the notch filter $|Q(\omega)T(\omega)|$ using a red line in the upper plot. The component gain of the pole is represented using a blue line and its inverse filter gain $|Q(\omega)|$ is represented using a green line. The bottom plot shows the compensation filter’s gain $|M(\omega)|$. This compensation filter is implemented using a linear phase FIR filter.

6.1.2 Spectral Shape Modulation: Feature “F”

The same modulation suppression filter $Q(z)T(z)M(z)$ is used to suppress the spectral shape modulations of the dB STRAIGHT spectrogram.

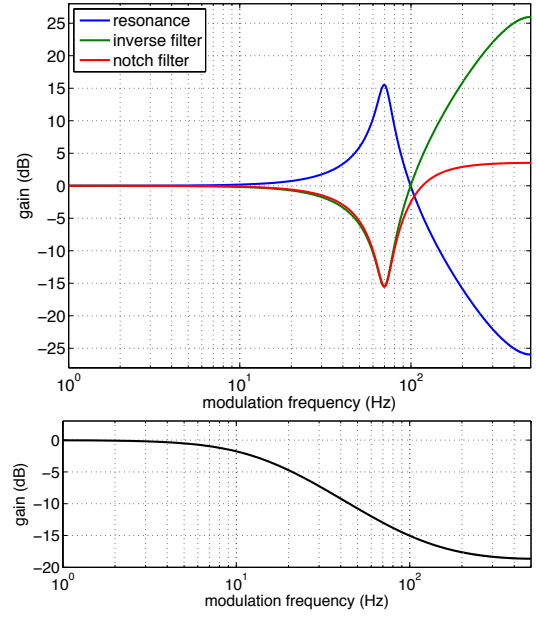


Figure 8. Notch filter gain $|Q(\omega)T(\omega)|$ (red line in the upper plot) and the target pole and its inverse filter gains (blue line and green lines in the upper plot) and compensating filter response $|M(\omega)|$ (lower plot).

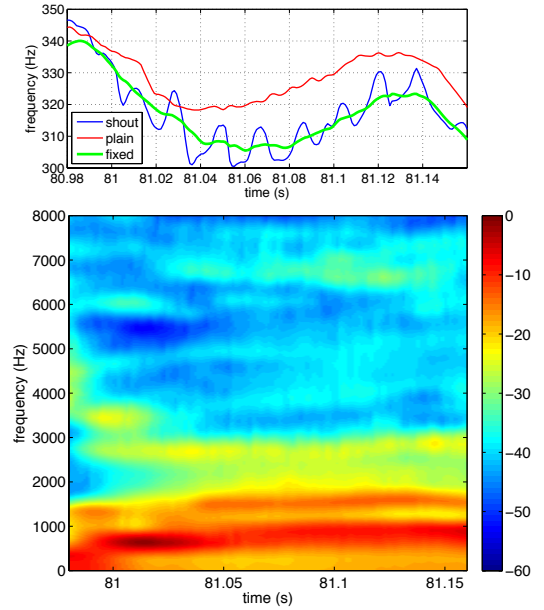


Figure 9. Processed F0 trajectory (green line in the upper plot) and STRAIGHT spectrogram (lower plot).

6.1.3 Octave-Band Level Difference: Feature “E”

The piecewise linearly interpolated negative dB difference, for example, shown in Fig. 7 is used to design a linear phase FIR filter. This target dB difference is converted to the absolute spectrum and the impulse response is calculated by inverse Fourier transform. This response is truncated using a 10 ms Hann window function to yield the response used in suppression.

Figure 9 shows the modified “shout” F0 trajectory and

label	F0 modulation	shape modulation	level difference
Q0F0E0	SUPPRESSED	SUPPRESSED	SUPPRESSED
Q0F0E1	SUPPRESSED	SUPPRESSED	INTACT
Q0F1E0	SUPPRESSED	INTACT	SUPPRESSED
Q0F1E1	SUPPRESSED	INTACT	INTACT
Q1F0E0	INTACT	SUPPRESSED	SUPPRESSED
Q1F0E1	INTACT	SUPPRESSED	INTACT
Q1F1E0	INTACT	INTACT	SUPPRESSED
Q1F1E1	INTACT	INTACT	INTACT

Table 1. Resynthesis conditions of test stimuli

winner	loser (number “xyz” represents the stimulus “QxFyEz”)							
	000	001	010	011	100	101	110	111
Q0F0E0	–	12	0	0	12	0	0	4
Q0F0E1	88	–	4	0	58	15	0	0
Q0F1E0	100	96	–	15	100	88	27	8
Q0F1E1	100	100	85	–	100	100	62	23
Q1F0E0	88	42	0	0	–	15	4	0
Q1F0E1	100	85	12	0	85	–	0	4
Q1F1E0	100	100	73	38	96	100	–	31
Q1F1E1	96	100	92	77	100	96	69	–

Table 2. Summary of forced choice results in %. “winner” represents the selected stimulus pattern and “loser” represents the un-selected pattern, where “0” represents SUPPRESSED and “1” represents INTACT states of features, “Q”, “F” and “E” respectively.

STRAIGHT spectrogram by suppressing all three “shout” specific features. All “shout” specific features are not visible on these modified representations.

6.2 Subjective Evaluation

A subjective evaluation test was conducted using resynthesized stimuli listed in Table 1. Five male and three female subjects were participated in the test. All possible combinations excluding the same stimulus pairs are counterbalanced and randomized. An interactive test program was used in the test. Each subject was instructed to select the stimulus with stronger impression of “shout.”

Table 2 shows the test results. The numbers in the table represents the percentage of the selection. The average preference differences due to feature “Q”, “F” and “E” are 15.6%, 48.3% and 12.7% respectively.

A generalized linear model (GLM) was also applied to analyze the results. The binomial model of GLM [28] was applied using the statistical package R [29]. It revealed that all three features are highly significant (the risk is smaller than 0.1%) and the intercept is virtually zero and insignificant. Therefore, the regression equation yields,

$$y = 1.287Q + 4.968F + 1.558E, \quad (15)$$

where the estimate y represents the logit conversion of the preference probability. Figure 10 shows scatter plot of this predicted scores and the test results.

These suggest that all these features are contributing to “shout” impression and can be controlled. It also suggests that it may be possible to convert a “plain” performance to a “shout” performance by manipulating these features.

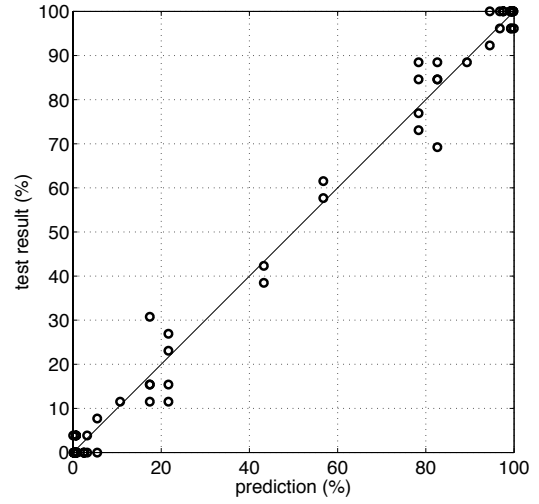


Figure 10. Predicted scores and the test results.

Test stimuli used in this article are stored in this conference proceedings. These stimuli and other materials are linked in our Web page [10].

7. CONCLUSIONS

A set of new analysis, modification and resynthesis procedures are introduced to illustrate physical correlates of “shout” impression in expressive singing voices. An F0 extractor with temporally fine resolution revealed rapid frequency modulation of the F0 trajectory in “shout” singing performance around 70 Hz in terms of modulation frequency. Spectral shape modulation in synchrony with this frequency modulation was also found by an interference-free power spectral representation. In addition to these temporal features, spectral level difference between “shout” and “plain” performance was observed. These newly introduced procedures enabled analysis and manipulation of these new features individually for investigating their perceptual effects. Subjective test results illustrated significant contributions of all three features in “shout” impression. Application of these findings to convert “plain” performance into “shout” is the next target of our study.

Acknowledgments

The authors thank Yuri Nishigaki of our former undergraduate student for preliminary analyses. This work is supported by the Aid for Scientific Research No.24300073, 23700221 and 24650085 from JSPS.

8. REFERENCES

- [1] J. Sundberg, *The science of the singing voice*. Northern Illinois University Press, 1987.
- [2] I. R. Titze, *Principles of voice production*. Prentice Hall, 1994.
- [3] “Cnfcus said: Writing cannot express speech completely. Speech cannot express thoughts completely. :

- from The I Ching or book of changes,” about 1000 BC, [one of the oldest of the Chinese classic text. Commentary maybe dated 475–221 BC].
- [4] K. Sakakibara, H. Fuks, N. Imagawa, and N. Tayama, “Growl voice in ethnic and Pop styles,” in *Proc. Int. Symp. on Musical Acoustics*, 2004.
 - [5] H. Kawahara, M. Morise, and T. Irino, “Analysis and synthesis of strong vocal expressions: Extension and application of audio texture features to singing voice,” *ICASSP2012*, pp. 5389–5392, 2012.
 - [6] M. Morise, T. Takahashi, H. Kawahara, and T. Irino, “Power spectrum estimation method for periodic signals virtually irrespective to time window position,” *Trans. IEICE*, vol. J90-D, no. 12, pp. 3265–3267, 2007, [in Japanese].
 - [7] H. Kawahara, M. Morise, T. Takahashi, R. Nisimura, T. Irino, and H. Banno, “A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0 and aperiodicity estimation,” *ICASSP2008*, pp. 3933–3936, 2008.
 - [8] H. Kawahara, T. Irino, and M. Morise, “An interference-free representation of instantaneous frequency of periodic signals and its application to F0 extraction,” *ICASSP2011*, pp. 5420–5423, May 2011.
 - [9] H. Kawahara, M. Morise, R. Nisimura, and T. Irino, “Higher order waveform symmetry measure and its application to periodicity detectors for speech and singing with fine temporal resolution,” *ICASSP2013*, pp. 6797–6801, 2013.
 - [10] H. Kawahara *et.al.*, “Shout suppression examples.” [Online]. Available: <http://www.wakayama-u.ac.jp/%7ekawahara/SMACSMC2013/>
 - [11] A. H. Nuttall, “Some windows with very good side-lobe behavior,” *IEEE Trans. Audio Speech and Signal Processing*, vol. 29, no. 1, pp. 84–91, 1981.
 - [12] W. Hess, *Pitch Determination of Speech Signals: Algorithms and Devices*. Springer-Verlag, 1983.
 - [13] A. de Cheveigne and H. Kawahara, “YIN, a fundamental frequency estimator for speech and music,” *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1917–1930, 2002.
 - [14] A. Camacho and J. G. Harris, “A sawtooth waveform inspired pitch estimator for speech and music,” *J. Acoust. Soc. Am.*, vol. 124, no. 3, pp. 1638–1652, 2008.
 - [15] M. Unser, “Sampling – 50 years after Shannon,” *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, 2000.
 - [16] H. Kawahara and M. Morise, “Technical foundations of TANDEM-STRAIGHT, a speech analysis, modification and synthesis framework,” *SADHANA*, vol. 36, no. 5, pp. 713–722, 2011.
 - [17] H. Kawahara, I. Masuda-Katsuse, and A. de Cheveigné, “Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction,” *Speech Communication*, vol. 27, no. 3-4, pp. 187–207, 1999.
 - [18] H. Akagiri, M. Morise, T. Irino, and H. Kawahara, “Evaluation and optimization of F0-adaptive spectral envelope estimation based on spectral smoothing with peak emphasis,” *Trans. IEICE*, vol. J94-A, no. 8, pp. 557–567, 2011, [in Japanese].
 - [19] B. H. Story, “Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002,” *J. Acoust. Soc. Am.*, vol. 126, no. 1, pp. 327–335, 2008.
 - [20] G. Fant and J. Liljencrants, “A four-parameter model of glottal flow,” *STL-QPSR*, vol. 26, no. 4, pp. 1–13, 1985.
 - [21] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal. I. fundamentals,” *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520 – 538, 1992.
 - [22] J. L. Flanagan and R. M. Golden, “Phase vocoder,” *Bell System Technical Journal*, pp. 1493–1509, November 1966.
 - [23] J. W. Dang and K. Honda, “Acoustic characteristics of the piriform fossa in models and humans,” *J. Acoust. Soc. Am.*, vol. 101, no. 1, pp. 456–465, JAN 1997.
 - [24] F. Itakura and S. Saito, “A statistical method for estimation of speech spectral density and formant frequency,” *Electro. Comm. Japan*, vol. 53-A, no. 1, pp. 36–43, 1970, [in Japanese].
 - [25] B. S. Atal and S. L. Hanauer, “Speech analysis and synthesis by linear prediction of the speech wave,” *J. Acoust. Soc. Am.*, vol. 50, no. 2B, pp. 637–655, 1971.
 - [26] H. Nishiwaki, H. Banno, and K. Asahi, “Investigation of correlation between temporal fluctuations of F0 and spectrum in scream vocal style,” *IEICE Technical Report SP2012-107*, vol. 112, no. 422, pp. 55–60, 2012, [in Japanese].
 - [27] I. R. Titze, “Nonlinear source–filter coupling in phonation: Theory,” *J. Acoust. Soc. Am.*, vol. 123, no. 5, pp. 2733–2749, May 2008.
 - [28] J. A. Nelder and R. W. M. Wedderburn, “Generalized linear models,” *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, no. 3, pp. 370–384, 1972.
 - [29] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2008.

Formant frequencies of sung vowels intonated by six traditional Japanese *Shigin* singers. Part I: Dataset construction and analysis method

Masashi Nakayama¹, Kosuke Kato², and Masaru Matsunaga³

^{1,3} Kagawa National College of Technology, Japan

²Osaka University, Japan

E-mail: m-nakayama@t.kagawa-nct.ac.jp, kato@uic.osaka-u.ac.jp

ABSTRACT

Shigin, a type of traditional Japanese singing, is a performance of reciting Japanese or Chinese poetry in Japanese. In order to understand the production mechanisms of *Shigin*, and thereby progress toward synthesized singing and singer training systems, the acoustic features of *Shigin* need to be clarified in relation to the interpretation styles of its practitioners. This study investigates each of the four lowest formant frequencies of sung vowels in relation to the singer's gender, the singer's vocal part, intonated vowel (/u/, /o/, /a/, /e/, or /i/), the loudness level, and phonation frequency. Anechoic recordings of the steady-state portions of vowels intonated by six trained *Shigin* singers using different volume levels, pitches, and vowels, were used for the acoustic analysis. This paper (Part I) describes the method to construct the dataset for analysis and the method for analyzing the constructed dataset. Results of the analysis and the statistical investigation will be reported in the following paper (Part II).

1. INTRODUCTION

Shigin, a type of traditional Japanese singing of Japanese or Chinese poetry in Japanese, conforming to a melodic line called *Seicho*. In order to understand the production mechanisms of *Shigin*, and thereby progress toward synthesized singing and singer training systems, we must first investigate the acoustic features of *Shigin* in relation to the interpretation styles of its practitioners.

A number of researchers have studied the acoustic features of classical singing. Notable among these are Sundberg and his coworkers, who conducted systematic investigations of operatic/classical singing [1], and Titze and his colleagues, who provided in-depth cross-disciplinary exploration of the physics and physiology of voice production mechanisms [2]. In the domain of Japanese traditional singing, Nakayama and Yanagida developed a large database of singing voices and discovered instances of formant shift and vibrato characteristics [3], while Tanaka *et al.* demonstrated a relationship between *Shigin* singers' interpretation styles and an acoustic parameter of the singing voice, the effective duration (τ_e) of the autocorrelation function, which showed random fluctuation during song [4]. Among the authors of the present paper, one has proposed a support system for training

Shigin singers [5], and two have presented examples of formant frequencies in *Shigin* singing [6].

Thus far, no systematic attempt has been made to examine the basic acoustic parameters of vowels intonated by *Shigin* singers. In order to extract such parameters from singing voice data objectively and automatically, it is necessary to define all of the relevant procedures for acoustic analysis. In particular, we must define the steady-state and unsteady-state (transient) portions and analyze each separately, as the values for basic acoustic parameters, such as F_0 and the formant frequencies, will differ greatly between the two. The overall goal of this study is to clarify the acoustic features of *Shigin* in relation to the interpretation styles of its practitioners.

As a part of achieving this overall goal, this study attempts to describe each of the four lowest formant frequencies (F_1 , F_2 , F_3 , or F_4) of sung vowels as a function of the singer's gender (*Gender*), the singer's vocal part (*Honsu*), being analogous to the tenor or soprano classifications in opera, intonated vowel (/u/, /o/, /a/, /e/, or /i/) (*Vowel*), the loudness level (L_{Aeq}), and phonation frequency (F_0). Anechoic recordings of the steady-state portions of vowels intonated six trained *Shigin* singers using different volume levels, pitches, and vowels, were used for the acoustic analysis. This paper describes the method to construct the dataset for analysis and the method for analyzing the constructed dataset.

2. TRADITIONAL SHIGIN SINGING

Shigin was established in the 19th century and has been practiced ever since [7]. When *Shigin* was first introduced, there were no melodic rules to follow, and performances remained relatively freeform. In an effort to formalize *Shigin*, Kimura *et al.* invented melodic rules for *Shigin* singers and established schools to train singers according to these rules [8]. Over time, the melodies produced by these formative rules have diverged to a limited degree, as schools and instructors developed their own styles. To achieve a full and rich style of *Shigin*, the student/trainee must undergo special training in breathing and utterance, and they must develop techniques such as *Yuri* (vibrato) and *Fushi* (control of melodic trajectory and singing volume). Correct accents must also be learned, as each word in a poem carries an important meaning, and so must be sung clearly and correctly. To

this end, the *Nippon Ginkenshibu* Foundation, an association dedicated to the traditional activities of *Shigin* and sword dancing, started promoting a standard set of accents for *Shigin* [9]. Musically, *Shigin* uses a relative score, expressed in the form of numbers, such as 2⁺–2–3. In standard musical notation, these are equivalent to C–D–E.

3. DATASET CONSTRUCTION AND ANALYSIS METHOD

3.1 Anechoic recordings of *Shigin* singing

Anechoic recordings of *Shigin* singing, containing songs and individual vowel sounds, were used for analysis and discussion of acoustic characteristics. This section describes the details of this database. Two male singers and four female singers, including a professional and two master-level singers, participated in our recording sessions [4]. Singings and utterances were recorded in an anechoic chamber. We used the B&K 1/2-inch Microphone Type 4819 with 48-kHz sampling and 16-bit resolution. The distance between the singer and the microphone was 50 cm, as is typical for recording of singing. The database consists of five Japanese vowels (/u/, /o/, /a/, /e/, and /i/) in three different keys (low, medium, and high) and three different strengths (weak, medium, and strong) of utterance.

Thus, the database collected a total of 1,620 samples from the six singers, intoning five vowels six times using nine different combinations of keys and strengths. Each utterance was sung by the given singer when a key was provided as a reference scale sound. Hence, the keys of the recorded utterances differed accordingly. In this study, only a half of database was used for subsequent acoustic analysis by extracting initial three trials out of six trials of utterance.

Table 1 shows the key conditions of *Shigin* singers singing at medium strength. We used the following acoustic parameters: F_0 , an equivalent continuous A-weighted sound pressure level L_{Aeq} , and formant frequencies F_1 , F_2 , F_3 , and F_4 . Here, F_0 indicates a factor of the key score, L_{Aeq} indicates the loudness of the vowel and, F_1 , F_2 , F_3 , and F_4 are the resonance frequencies in the vowel. This study analyzed feature parameters including the trajectory of F_0 for Japanese vowels based on a formal definition of the steady-state portion of *Shigin* because the trajectory was not measured correctly in a previous study [6].

Figure 1 shows representations of an utterance of the Japanese vowel /a/ intoned by Singer 2 (7-hon, female) in medium key and medium strength. The upper part of the figure shows the waveform and the lower part shows the corresponding spectrogram. A fundamental frequency of about 300 Hz for the voiced section from spectrogram was found. However, utterance of head and tail sections at voice section has volume at unsteady-state level, and then also finds transient of vibrato. For systematic analysis of recorded singing voice data, the steps to extract parameters of these steady-state portions were formalized.

Singer (<i>Honsu</i> , gender)	Pitch		
	Low	Medium	High
Singer 1 (8-hon, female)	A3	E4	A4
Singer 2 (7-hon, female)	G3	D4	G#4
Singer 3 (6-hon, female)	A3	D4	G4
Singer 4 (6-hon, female)	G3	D4	G4
Singer 5 (3-hon, male)	F3	C4	G4
Singer 6 (2-hon, male)	D#3	A#3	D#4

Table 1. Key conditions of *Shigin* singers.

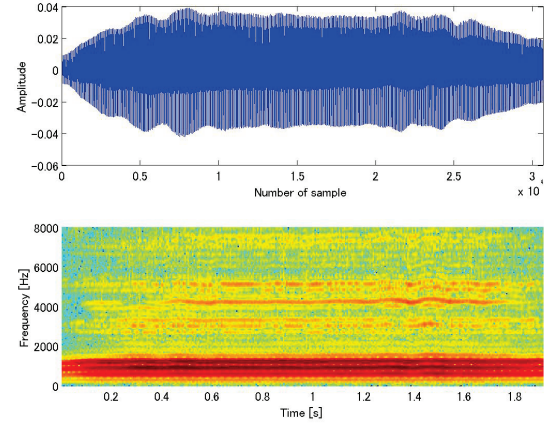


Figure 1. Waveform (upper part) and spectrogram (lower part) of an utterance of the Japanese vowel /a/ intoned by Singer 2 (7-hon, female) in medium key and medium strength.

3.2 Extraction of steady-state portions

Prior to the systematic analysis of recorded singing voice data, we will formalize the following two processing steps:

1. Define the steady-state portions of *Shigin* singing.
2. Extract the F_0 of these portions automatically.

In order to statistically isolate the steady-state portion of a vocal segment, two types of sub-segments were trimmed from the beginning and end of the utterance in advance:

- Type I: Segments whose L_{Aeq} was low.
- Type II: Segments whose F_0 was higher or lower than the median of the F_0 of the sung vowel by 150 cents (approximately the maximum range of the vibrato extent [10]).

The details are described below:

Type I:

1. L_{Aeq} was calculated by the following equation.

$$L_{Aeq} = 10 \log_{10} \left\{ \frac{1}{T} \int_{t_1}^{t_2} \frac{p_A^2(t)}{p_0^2} dt \right\} \quad [\text{dB}] \quad (1)$$

where $p_A(t)$, and the absolute threshold of hearing p_0 are introduced as 20 μPa .

2. Calculate L_{Aeq} for frame n by using the following equation.

$$L_{Aeq}(n) = L_{calib} + 20 \log_{10} \frac{|s_m(n) * p(n)|}{p_0} \quad (2)$$

Here, $s_m(n)$ is the amplitude of each frame, $p(n)$ is the impulse response of A-weighted coefficients, p_0 is the

average amplitude of the calibration signal, and $L_{calib}=92.1$ dBA is the calibration level. For each frame, $s_m(n)$ has been calculated for the following conditions:

- integration interval = 30 ms
 - stepping interval = 30 ms
 - time window = rectangular
3. Calculate the frame bounded by start sample i_{start} and end sample i_{end} by analyzing each of the following equations:

$$i_{start} = \min\{\arg(L_{Aeq}(i) > 55)\} \quad (3)$$

$$i_{end} = \begin{cases} L - \Delta_{tail} & \text{if } \max_{L-\Delta_{tail} < i < L} \{L_{Aeq}(i)\} < 55 \\ \max_{L-\Delta_{tail} < i < L} \{\arg(L_{Aeq}(i) > 55)\} & \text{otherwise} \end{cases} \quad (4)$$

Here, $L = \Delta_{tail} = 20,000$ samples was the tail margin.

4. Calculate a new waveform s' using the following equation:

$$s'(n) = \begin{cases} s_m(n) & \text{if } i_{start} + 1 \leq n \leq i_{end} - 1 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Type II:

1. Track the F_0 contours by making use of the first major peak of the short-term autocorrelation function (s-ACF). Here, the s-ACF has been calculated for the following conditions:
 - integration interval = 20 ms
 - stepping interval = 5 ms
 - time window = blackman
 - interpolation function = spline
2. Calculate the median of F_0 , which we call f_{med} .
3. Calculate the trajectory f'_{cent} . Here, the units were changed from Hz to cents, using the following equation for each frame:

$$f'_{cent}(j) = 1200 \log_2 \left(\frac{f'_0(j)}{f_{med}} \right) \quad (6)$$

4. Calculate the start sample j_{start} and the end sample j_{end} when the trajectory f'_{cent} is within ± 150 cent of f_{med} , as follows:

$$-150 < f'_{cent}(j) < 150 \quad (7)$$

$$j_{start} = \min\{\arg(-150 < f'_{cent}(j) < 150)\} \quad (8)$$

$$j_{end} = \max\{\arg(-150 < f'_{cent}(j) < 150)\} \quad (9)$$

5. Calculate the F_0 of the steady-state portion f''_0 , using the following equation:

$$f''_0(n) = \begin{cases} f'_0(n) & \text{if } j_{start} + 1 \leq n \leq j_{end} - 1 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

6. Search for a local maximum greater than 150 cents at all f''_{cent} . The location of this maximum defines the index k_{max} .
7. Determine the number of samples over which f''_{cent} increases monotonically before reaching $f''_{cent}(k_{max})$, n_{up} . Similarly, determine n_{down} , the number of samples over which f''_{cent} decreases monotonically after reaching $f''_{cent}(k_{max})$.
8. Calculate the start sample k_{start} and the end sample k_{end} according to the following equations (Equations (11) and (12)) in order to extract the longest steady-state portion of the voice sample, is shown Figure 2.

$$k_{start} = \begin{cases} 1 & \text{if } k_{max} > j_{end} - 1 - k_{max} \\ k_{max} + n_{down} & \text{otherwise} \end{cases} \quad (11)$$

$$k_{end} = \begin{cases} k_{max} - n_{up} & \text{if } k_{max} > j_{end} - 1 - k_{max} \\ j_{end} - 1 & \text{otherwise} \end{cases} \quad (12)$$

9. Define the new signal s'' , using the following equation.

$$s''(n) = \begin{cases} s'(n) & \text{if } k_{start} \leq n \leq k_{end} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

3.2.1 Results

Figure 3 shows the SPL contour of a vowel intoned by a *Shigin* singer. Figure 4 shows the SPL contour of a vowel with sub-segments of type I. As can be seen, the low-level segments at the beginning and end of the vowel were eliminated. Figure 5 shows the F_0 contour of a vowel with sub-segments of type I. Figure 6 shows the final result for an F_0 that conformed to both types I and II. Comparison of Figure 6 with Figure 5 shows that non-stationary sections from 0 to 0.18 sec were eliminated. These results demonstrate that the F_0 contour of the steady-state portion of a vowel intoned by a *Shigin* singer can indeed be analyzed objectively and automatically. Judging from these results, our methods worked as expected and provided automatic extraction. The extraction accuracy of F_0 can be evaluated by comparing it to the F_0 from Praat, conventional speech analysis software released by the University of Amsterdam [11], as shown in Figure 7. Note that the trajectory of Figure 6 is similar to but smoother than that of Figure 7, indicating good extraction accuracy.

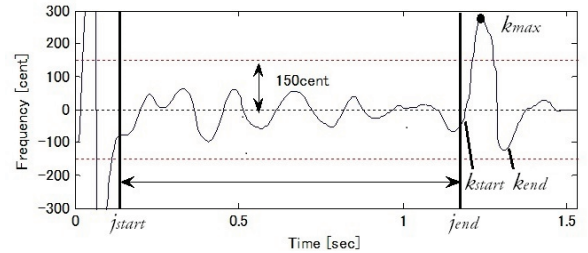


Figure 2. Example of irregular section at vibrato.

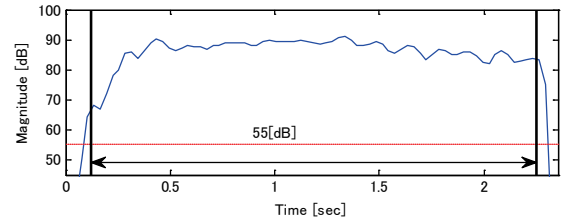


Figure 3. SPL contour of a vowel.

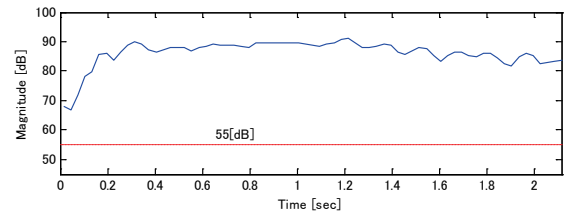


Figure 4. SPL contour of a vowel with type I.

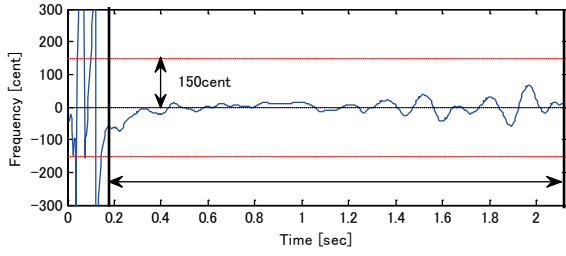


Figure 5. F_0 contour of a vowel with type I.

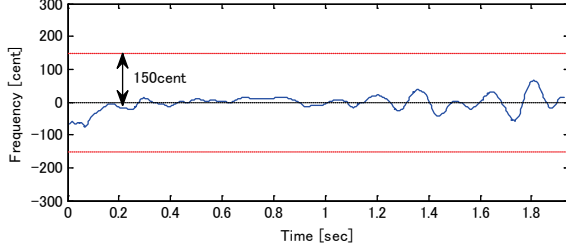


Figure 6. F_0 contour of a vowel with types I and II.

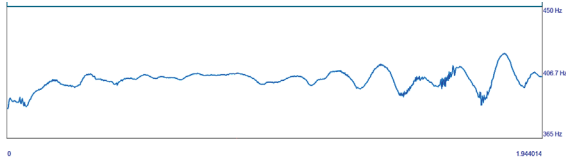


Figure 7. F_0 contour from Praat with types I and II.

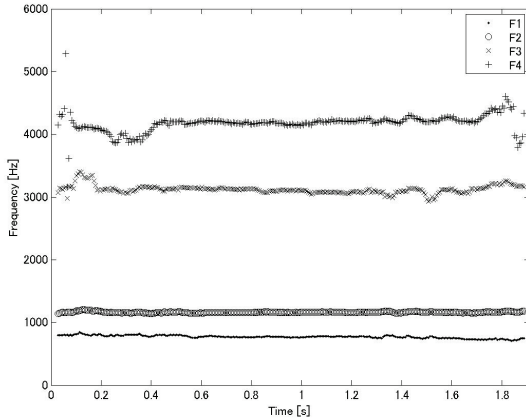


Figure 8. Example of estimated F_1 , F_2 , F_3 , and F_4 as a function of time.

3.3 Extractions of F_1 , F_2 , F_3 , and F_4 from steady-state portions

3.3.1 Materials and methods

In order to extract each of F_1 , F_2 , F_3 , and F_4 at steady-state portions of the singing voice dataset, the robust formant tracking function of Praat software [11] with the following default parameters were used.

- Time step: auto
- Maximum number of formants: 5
- Maximum formant: 5,500 [Hz]
- Window length: 0.025 [s]
- Pre-emphasis from: 50 [Hz]
- Number of standard deviations: 1.5
- Maximum number of iterations: 5
- Tolerance: 0.000001

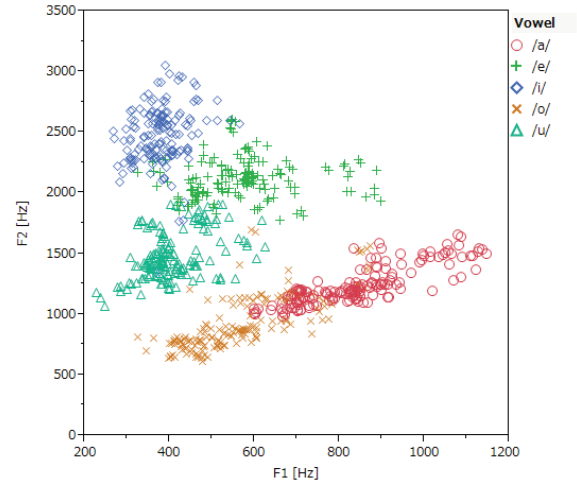


Figure 9. Distribution of extracted F_1 and F_2 .

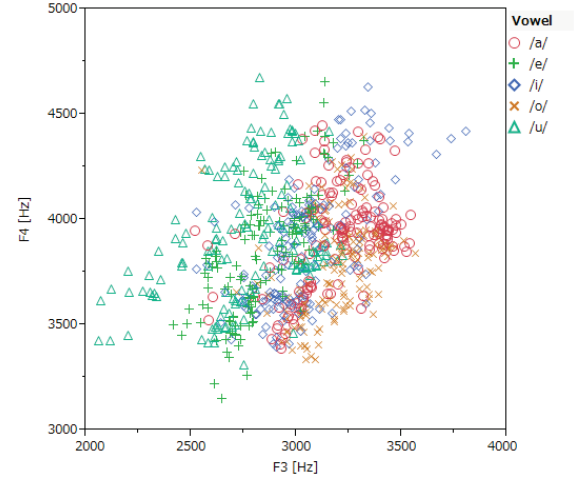


Figure 10. Distribution of extracted F_3 and F_4 .

3.3.2 Results

Figure 8 shows an example of extracted F_1 , F_2 , F_3 , and F_4 as a function of time, for an utterance of the Japanese vowel /a/ intonated by Singer 2 (7-hon, female) in medium key and medium strength.

Figure 9 shows distribution of extracted F_1 and F_2 , while Figure 10 shows distribution of extracted F_3 and F_4 .

3.4 Dataset construction

3.4.1 Feature vector for Shigin singing analysis

In order to describe F_1 , F_2 , F_3 , or F_4 in relation to both qualitative attributes (*Singer*, *Gender*, *Honsu*, *Vowel*, *Pitch*, *Volume*, and *Trial*) and acoustic features (L_{Aeq} , F_0 , F_1 , F_2 , F_3 , and F_4), the feature vector for each utterance was defined as shown in Equation (14).

$$\mathbf{F}_{Singer, Gender, Honsu, Vowel, Pitch, Volume, Trial} = (L_{Aeq}, F_0, F_1, F_2, F_3, F_4) \quad (14)$$

3.4.2 Dimensional compressions of each feature

In order to compress the dimensions of each of F_0 , F_1 , F_2 , F_3 , and F_4 , the geometric mean of each of F_0 , F_1 , F_2 , F_3 , and F_4 in the steady-state portion was calculated. This was done because these frequency features were time-series vectors, consisting of many values.

3.4.3 Dataset structure

In advance of analysis, the seven samples whose L_{Aeq} lower than 55 dBA had been omitted from the dataset.

Consequently, the final dataset was composed of 13 variables (see Equation (14)) \times 803 samples (6 singers \times 5 vowels \times 3 volume levels \times 3 pitches \times 3 trials – 7 data whose L_{Aeq} were lower than 55 dBA).

3.5 Statistical analysis

Using JMP 10 software [12], the following statistical analysis were conducted.

3.5.1 Four-way ANOVA and linear prediction model on each of F_1 , F_2 , and F_3

On the basis of previous studies on formant frequencies of Western operatic singing [1], this study attempted to describe each of the three lowest formant frequencies (F_1 , F_2 , and F_3) of sung vowels as a function of *Vowel*, *Gender*, L_{Aeq} , and F_0 . Hence, each of the four variables: *Vowel*, *Gender*, L_{Aeq} , and F_0 was dealt as an explanatory variable, while each of the four variables: F_1 , F_2 , and F_3 was regarded as an objective variable. Consequently, a four-way ANOVA was performed.

The linear prediction model for each of F_1 , F_2 , and F_3 that employs the four variables: *Vowel*, *Gender*, L_{Aeq} , and F_0 was also analyzed because each model is useful for e.g. singing voice synthesis, development of training support system, and understanding of the voice production mechanism of singers.

3.5.2 Four-way ANOVA on F_4

Referring to the previous studies of Western operatic singing [1], this study attempted to describe F_4 of sung vowels as a function of *Vowel*, *Singer*, L_{Aeq} , and F_0 . Hence, each of the four variables: *Vowel*, *Singer*, L_{Aeq} , and F_0 was dealt as an explanatory variable, while the variable F_4 was regarded as an objective variable. Consequently, a four-way ANOVA was performed.

4. CONCLUSIONS

In order to describe each of the four lowest formant frequencies (F_1 , F_2 , F_3 , or F_4) of vowels intonated by *Shigin* singers as a function of the singer's gender (*Gender*), the singer's vocal part (*Honsu*), intonated vowel (/u/, /o/, /a/, /e/, or /i/) (*Vowel*), the loudness level (L_{Aeq}), and phonation frequency (F_0), this paper (Part I) showed the method to construct the dataset for analysis and the method for analyzing the constructed dataset.

Results of the analysis and the statistical investigations will be reported in the following paper (Part II) [13].

ACKNOWLEDGEMENTS

The authors thank Prof. Takashi Yano and Dr. Keiji Kawai of Kumamoto University for providing access to a database of anechoic recordings *Shigin* utterances. The authors thank Mr. Syouhei Mizobuchi of Wakayama University for building a basic program for acoustic analysis.

REFERENCES

- [1] J. Sundberg, *The Science of The Singing Voice*, Northern Illinois University Press, 1987.
- [2] I. Titze, *Principles of Voice Production Second Printing*, National Center for Voice and Speech, 2000.
- [3] I. Nakayama and M. Yanagida, "Introduction to database of traditional Japanese singing with examples of comparative studies on formant shifts and vibrato among genres," *Acoustical Science and Technology*, Volume 29, Issue 1, pp. 58–65, 2008.
- [4] S. Tanaka, K. Kato, K. Kawai, T. Yano, and Y. Ando, "Traditional Japanese singing Shigin described by the effective duration of the running autocorrelation function of variably singing volume voice signals," *Proceedings of Student Meeting, Kyushu Chapter of the Acoustical Society of Japan*, 2005 (in Japanese).
- [5] M. Nakayama, "Fundamental research on a singing training support system for Shigin: Japanese traditional singing," *IEEE SoutheastCon2012*, S4-A-1, Florida, U.S.A., 2012.
- [6] M. Nakayama, S. Mizobuchi, and K. Kato, "A study of formant frequencies in Japanese traditional Shigin singers," *Proceedings of 2012 Annual Meeting Record IEE Japan*, 3-015, 2012 (in Japanese).
- [7] K. Sakakibara, *Gindo Taikan*, Kinensya, 1972 (in Japanese).
- [8] K. Hiroshige, *Renshuukyohon Dai Isyuu*, Koufuuryu Sanyogineikai, 1972 (in Japanese).
- [9] *Ginkenshibudo Akusentotsuki Knshisyuu (Zetsukuhen)*, NIPPON GINKENSHIBU Foundation, 1977 (in Japanese).
- [10] E. Prame, "Vibrato extent and intonation in professional Western lyric singing," *J. Acoust. Soc. Am.*, Volume 102, pp. 616–621, 1997.
- [11] Praat, <http://www.fon.hum.uva.nl/praat/>
- [12] JMP, <http://www.jmp.com/>
- [13] K. Kato, M. Nakayama, and M. Matsunaga, "Formant frequencies of sung vowels intonated by six traditional Japanese *Shigin* singers. Part II: Results of analysis and statistical investigations," *Proceedings of Stockholm Music Acoustics Conference 2013*, 2013.

Formant frequencies of sung vowels intonated by six traditional Japanese *Shigin* singers. Part II: Results of analysis and statistical investigations

Kosuke Kato¹, Masashi Nakayama², and Masaru Matsunaga²

¹Osaka University, Japan

²Kagawa National College of Technology, Japan

E-mail: kato@uic.osaka-u.ac.jp, m-nakayama@t.kagawa-nct.ac.jp

ABSTRACT

This study aims to clarify the acoustic characteristics of *Shigin*, a type of traditional Japanese singing. On the basis of the materials and method as described in Part I, this study investigates the four lowest formant frequencies of sung vowels in relation to the singer's gender, intonated vowel, the loudness level, and the phonation frequency. An analysis showed that: (1) The third formant frequency was about 400–500 Hz higher and the fourth formant frequency was about 1,000 Hz higher than those for operatic singing. (2) The mean of the second formant frequency of the /u/ vowel was similar to that for normal speech mode in Japanese but about 400 Hz higher than that for *Noh* singing and about 800 Hz higher than that for operatic singing. (3) 92.1% of the variance in the second formant frequency and 83.0% of the variance in the first formant frequency were described by a linear model that employs the four variables: intonated vowel, the singer's gender, the loudness level, and the phonation frequency. (4) Each of the first formant frequency and the second formant frequency was significantly increased with the phonation frequency. These findings can be considered as the major acoustic characteristics of *Shigin*.

1. INTRODUCTION

The overall goal of our study is to clarify the acoustic characteristics of a type of traditional Japanese singing, *Shigin*. As a part of achieving this overall goal, on the basis of the materials and method as described in Part I [1], this study investigates the four lowest formant frequencies (F_1 , F_2 , F_3 , and F_4) of sung vowels in relation to the singer's gender (*Gender*), the singer's vocal part (*Honsu*), intonated vowel (/u/, /o/, /a/, /e/, or /i/) (*Vowel*), the loudness level (L_{Aeq}), and phonation frequency (F_0). Anechoic recordings of the steady-state portions of vowels intonated six trained *Shigin* singers using different volume levels, pitches, and vowels, were used for the acoustic analysis. The characteristics of *Shigin* are discussed in comparison with the cases of normal speech mode in Japanese, Western operatic singing, choir singing and choir speech, and other traditional Japanese singing styles.

2. RESULTS OF ANALYSIS AND STATISTICAL INVESTIGATIONS

The first question is whether F_1 , F_2 , F_3 , and F_4 for *Shigin* singing are different from those for other singing style and/or those for normal speech mode.

Figure 1 and Figure 2 shows the distributions of the condition means of extracted F_1 , F_2 , F_3 , and F_4 values for each vowel. As can be seen in this figure, the extracted F_1 , F_2 , F_3 , and F_4 for *Shigin* singing were different from those for the cases of Western operatic singing [2, 3], choir singing and choir speech [4], one of the other Japanese traditional singing styles, *Noh* [5], and/or the normal speech mode in Japanese [6].

Figure 1(a) demonstrates that the values of F_1 of each vowel for *Shigin* singing were between those for the normal speech mode in Japanese [6] and those for Western operatic singing [2, 3] or *Noh* singing [5].

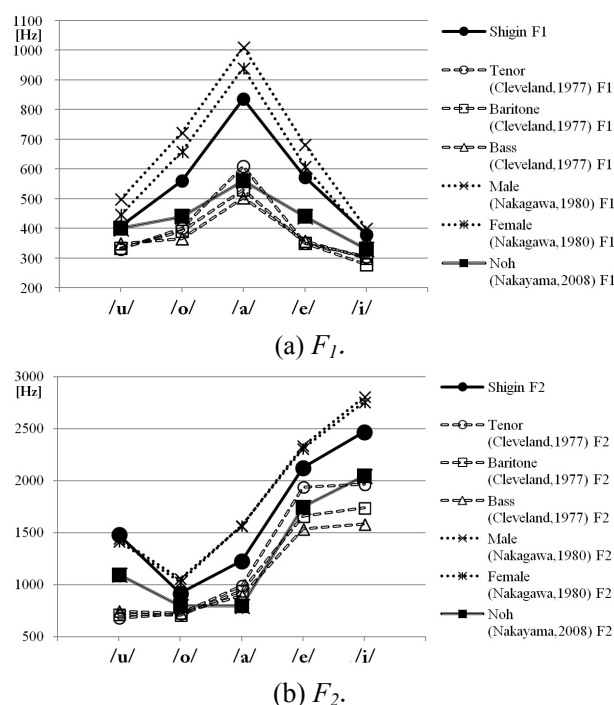


Figure 1. Distributions of extracted F_1 and F_2 of five vowels for *Shigin*: Comparison with the cases of normal speech mode in Japanese, Western operatic singing, and *Noh* singing.

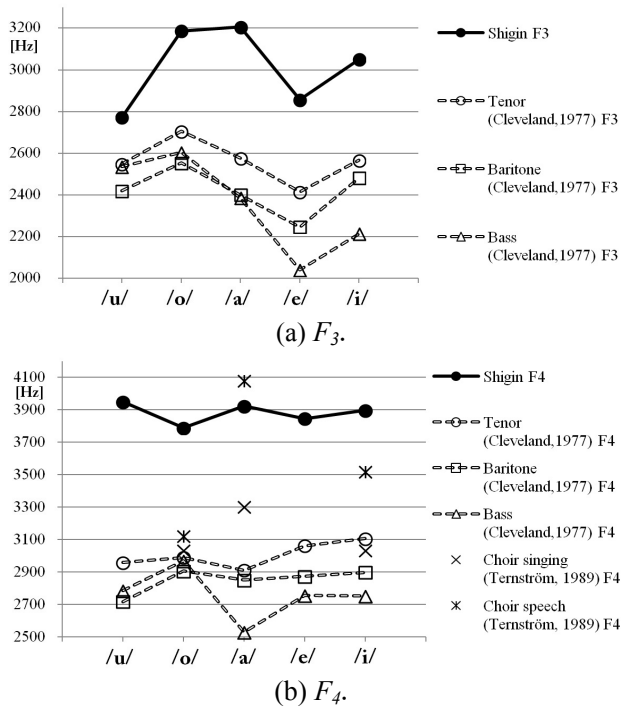


Figure 2. Distributions of extracted F_3 and F_4 of five vowels for *Shigin*: Comparison with the cases of Western operatic singing and choir singing.

Figure 1 (b) demonstrates that the values of F_2 of /u/ vowel for *Shigin* singing were similar to those for the normal speech mode in Japanese [6] but about 400 Hz higher than those for *Noh* singing [5] and about 800 Hz higher than those for Western operatic singing [2, 3]. Figure 1 (b) further demonstrates that the values of F_2 of the other vowels (/o/, /a/, /e/, and /i/) for *Shigin* singing were between the values for the normal speech mode in Japanese [6] and the values for Western operatic singing [2, 3] or *Noh* singing [5].

Figure 2 (a) demonstrates that the values of F_3 of the five Japanese vowels for *Shigin* singing were about 400 - 500 Hz higher than those for Western operatic singing [2, 3], and Figure 2 (b) demonstrates the values of F_4 of the five Japanese vowels for *Shigin* singing were about 1,000 Hz higher than those for Western operatic singing [2, 3] and about 600-900 Hz higher than those for choir singing [4].

The second question is: how is the distribution of each of F_1 , F_2 , and F_3 for *Shigin* singing?

Figure 3 shows the condition means of F_1 for each vowel, with box plots showing the mean ± 1 standard deviations and the maximum and the minimum values. The values of F_1 of individual utterance ranged between 228 Hz (/u/) to 1,147 Hz (/a/). The vowel arithmetic means varied between 379 Hz (/i/) to 836 Hz (/a/). The results of one-way ANOVA, setting *Vowel* as a factor, inputting 803 data points of F_1 , showed that 71.0% ($p < 0.001$) of the variance in F_1 was described by the Japanese five vowels.

Figure 4 shows the condition means of F_2 for each vowel, with box plots showing the mean ± 1 standard deviations and the maximum and the minimum values. The values of F_2 of individual utterance ranged between

610 Hz (/o/) to 3,050 Hz (/i/). The vowel arithmetic means varied between 922 Hz (/o/) to 2,473 Hz (/i/). The results of one-way ANOVA, setting *Vowel* as a factor, inputting 803 data points of F_2 , showed that 89.2% ($p < 0.001$) of the variance in F_2 was described by the Japanese five vowels.

Figure 5 shows the condition means of F_3 for each vowel, with box plots showing the mean ± 1 standard deviations and the maximum and the minimum values. The values of F_3 of individual utterance ranged between 2,059 Hz (/u/) to 3,803 Hz (/i/). The vowel arithmetic means varied between 2,768 Hz (/u/) to 3,206 Hz (/a/). The results of one-way ANOVA, setting *Vowel* as a factor, inputting 803 data points of F_3 , showed that 41.4% ($p < 0.001$) of the variance in F_3 was described by the five Japanese vowels.

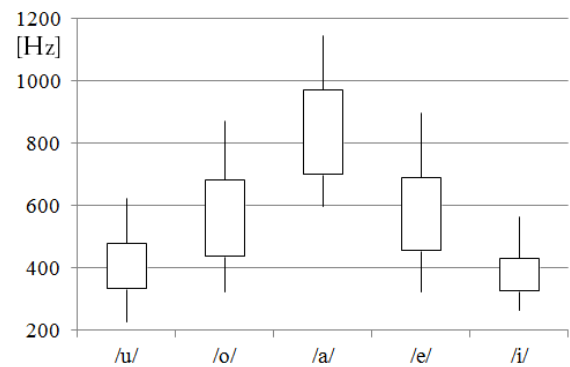


Figure 3. Distributions of extracted F_1 of sung vowels for *Shigin*. $R^2 = 0.710$ ($p < 0.001$).

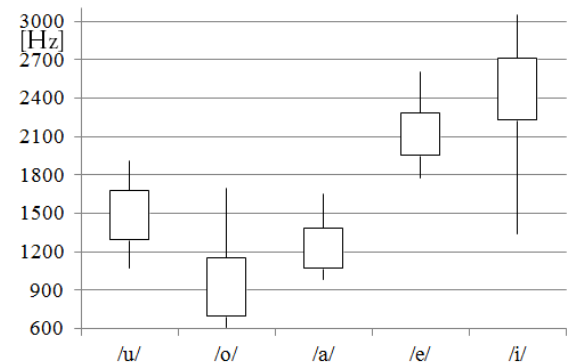


Figure 4. Distributions of extracted F_2 of sung vowels for *Shigin*. $R^2 = 0.892$ ($p < 0.001$).

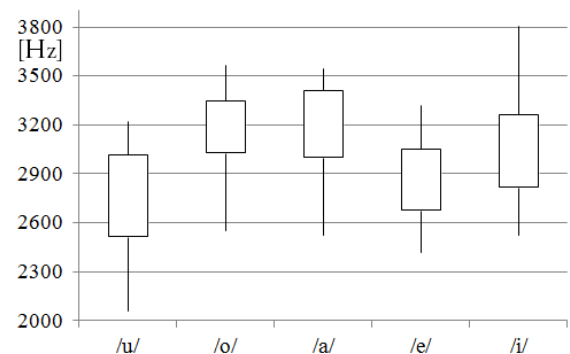


Figure 5. Distributions of extracted F_3 of sung vowels for *Shigin*. $R^2 = 0.414$ ($p < 0.001$).

Factor	DF	Sum of Squares	Mean square	F-ratio
Model	22	26,584,082	1,208,367	276.0734
Error	780	3,414,044	4,377	
Total	802	29,998,125		(Prob>F) <.0001

(a) Results of ANOVA.

Factor	DF	Sum of Squares	Prob > F
Vowel	4	7,669,540	<.0001
Gender	1	370,199	<.0001
L_{Aeq}	1	225,304	<.0001
F_0	1	538,093	<.0001
Vowel*Gender	4	292,091	<.0001
Vowel * L_{Aeq}	4	128,238	<.0001
Vowel * F_0	4	79,181	0.0013
Gender * L_{Aeq}	1	270,322	<.0001
Gender * F_0	1	130,592	<.0001
L_{Aeq} * F_0	1	549,022	<.0001

(b) Results of effect tests.

Table 1. Results of four-way ANOVA for F_1 .
Four most significant factors are typed in Bold font.

Factor	DF	Sum of Squares	Mean square	F-ratio
Model	22	38,141,602	1,733,709	67.1817
Error	780	20,128,893	25,806	
Total	802	58,270,494		(Prob>F) <.0001

(a) Results of ANOVA.

Factor	DF	Sum of Squares	Prob > F
Vowel	4	16,309,348	<.0001
Gender	1	1,278,541	<.0001
L_{Aeq}	1	1,439,901	<.0001
F_0	1	1,309,296	<.0001
Vowel*Gender	4	438,727	0.0021
Vowel * L_{Aeq}	4	687,023	<.0001
Vowel * F_0	4	784,349	<.0001
Gender * L_{Aeq}	1	314	0.9122
Gender * F_0	1	252,988	0.0018
L_{Aeq} * F_0	1	333,673	0.0003

(b) Results of effect tests.

Table 3. Results of four-way ANOVA for F_3 .
Four most significant factors are typed in Bold font.

Factor	DF	Sum of Squares	Mean square	F-ratio
Model	22	274,839,947	12,492,725	516.1370
Error	780	18,879,338	24,204	
Total	802	293,719,284		(Prob>F) <.0001

(a) Results of ANOVA.

Factor	DF	Sum of Squares	Prob > F
Vowel	4	162,587,452	<.0001
Gender	1	1,191,813	<.0001
L_{Aeq}	1	365,640	0.0001
F_0	1	1,441,123	<.0001
Vowel*Gender	4	530,506	0.0002
Vowel * L_{Aeq}	4	585,346	<.0001
Vowel * F_0	4	1,256,313	<.0001
Gender * L_{Aeq}	1	255,547	0.0012
Gender * F_0	1	124,687	0.0235
L_{Aeq} * F_0	1	247,228	0.0014

(b) Results of effect tests.

Table 2. Results of four-way ANOVA for F_2 .
Four most significant factors are typed in Bold font.

Factor	DF	Sum of Squares	Mean square	F-ratio
Model	50	41,815,179	836,304	33.1123
Error	752	18,992,972	25,257	
Total	802	60,808,150		(Prob>F) <.0001

(a) Results of ANOVA.

Factor	DF	Sum of Squares	Prob > F
Singer	4	5,626,861	<.0001
Vowel	1	568,045	0.0002
L_{Aeq}	1	275,761	0.0010
F_0	20	464,196	<.0001
Singer*Vowel	5	4,537,345	<.0001
Singer*L_{Aeq}	5	948,181	<.0001
Singer* F_0	4	659,795	0.0001
Vowel*L_{Aeq}	4	1,020,305	<.0001
Vowel * F_0	1	833,034	<.0001
L_{Aeq} * F_0	4	317,063	0.0004

(b) Results of effect tests.

Table 4. Results of four-way ANOVA for F_4 .
Four most significant factors are typed in Bold font.

The third question is: what impacts each of F_1 , F_2 , F_3 , and F_4 for *Shigin* singing?

Tables 1, 2, and 3 list each of the results of four-way ANOVA, setting *Vowel*, *Gender*, L_{Aeq} , and F_0 as factors, inputting 803 data points of F_1 , F_2 , and F_3 as an objective variable. The full factorial ANOVA was performed in which the main effects of each variable and interactions

among the variables were included in the model.

Being common among the results of F_1 , F_2 , and F_3 , the *Vowel* factor was the most significant factor while the *Gender* factor was the fourth significant factor, showing the largest and the fourth largest sum of squares, respectively (Tables 1(b), 2 (b), and 3(b)).

However, the second and the third largest sum of squares were different among those formant frequencies. As listed in Tables 1(b), 2(b), and 3(b), the cross-effect $L_{Aeq} * F_0$ was the second significant factor and the F_0 factor was the third significant factor for F_1 , the F_0 factor was the second significant factor and the cross-effect $Vowel * F_0$ was the third significant factor for F_2 , whereas the L_{Aeq} factor was the second significant factor and the F_0 factor was the third significant factor for F_3 .

Table 4 lists the results of four-way ANOVA, setting *Singer*, *Vowel*, L_{Aeq} , and F_0 as factors, inputting 803 data points of F_4 as an objective variable. The full factorial ANOVA was performed in which the main effects of each variable and interactions among the variables were included in the model. As can be seen in Table 4 (b), the four most significant factors were, *Singer*, *Singer*Vowel*, *Vowel*L_{Aeq}*, and *Singer*L_{Aeq}*, respectively.

The fourth question is whether we can describe each of F_1 , F_2 , and F_3 using a simple linear prediction model that employs the four variables: *Vowel*, *Gender*, L_{Aeq} , and F_0 . This model can be formulated as

$$F_{\{1,2,3\}} = a_0 + a_1(Vowel) + a_2(Gender) + a_3(L_{Aeq}) + a_4(F_0) \quad (1)$$

where a_0 , $a_1(Vowel)$, $a_2(Gender)$, $a_3(L_{Aeq})$, and $a_4(F_0)$ are values calculated from a multiple regression analysis with dummy variables; these were fitted to the data.

Table 5 lists the values of a_0 , $a_1(Vowel)$, $a_2(Gender)$, $a_3(L_{Aeq})$, and $a_4(F_0)$ in Equation (1). R^2 values varied between 0.565 ($p < 0.001$) and 0.921 ($p < 0.001$). The resulted coefficient a_2 shows that F_1 , F_2 , and F_3 for female singer are 20 – 70 Hz higher than those for male singer. The obtained coefficient a_3 indicates that louder voice cause the increase in F_1 but cause the decrease in F_2 and F_3 . The resulted coefficient a_4 demonstrates that higher pitched voice cause the increase in F_1 , F_2 , and F_3 , indicating the existence of formant tuning.

Figure 6 shows how each of F_1 and F_2 changed with their phonation frequencies (F_0): both F_1 and F_2 increased with increasing phonation frequency (F_0). In order to quantify these observations, a linear regression analysis was conducted. Table 6 shows the results of a linear regression analysis, fitting equation (2) to the data points of each of F_1 and F_2 for each vowel.

$$F_{\{1,2\},\{/u/,/o/,/a/,/e/,/i/\}} = b_0 + b_1(F_0) \quad (2)$$

Results show that a significant regression line was fitted to each of F_1 and F_2 for each vowel, except for F_2 of /e/ and /i/. In particular, the value of b_1 was close to 1 for the case of F_1 of /o/, /a/, and /e/ and F_2 of /u/ and /a/, which means that a 100 Hz increase in F_0 would cause about 100 Hz increase in F_1 or F_2 .

Factor (Item)	Category	Coefficient		
		F_1	F_2	F_3
a_0		+277	+1,924	+3,504
$a_1(Vowel)$	/u/	0	0	0
	/o/	+8	-713	+188
	/a/	+281	-367	+246
	/e/	+21	+482	-139
	/i/	-170	+797	-5
$a_2(Gender)$	Male	0	0	0
	Female	+20	+44	+70
$a_3(L_{Aeq})$		+0.942	-8.06	-9.22
$a_4(F_0)$		+0.693	+1.05	+0.581

Table 5. Values of a_0 , $a_1(Vowel)$, $a_2(Gender)$, $a_3(L_{Aeq})$, and $a_4(F_0)$ in Equation (1). $R^2 = 0.830$ ($p < 0.001$) for F_1 ; $R^2 = 0.921$ ($p < 0.001$) for F_2 ; and $R^2 = 0.565$ ($p < 0.001$) for F_3 .

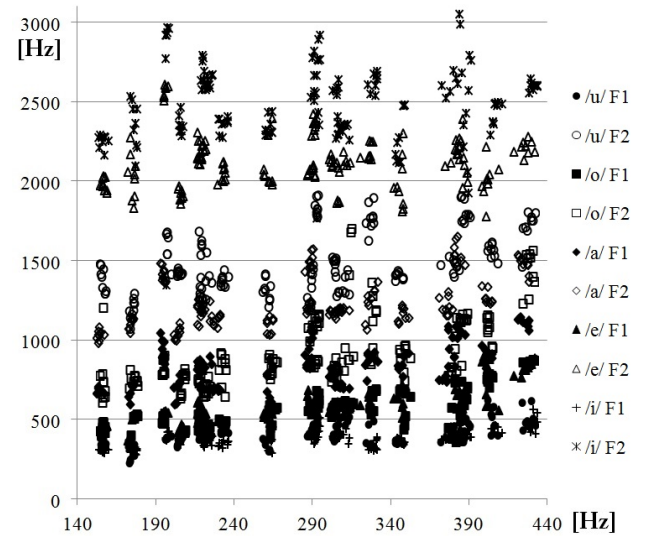


Figure 6. Changes in F_1 and F_2 with phonation frequency.

	b_0	b_1	R^2	p
/u/	269	0.482	0.282	<0.001
/o/	224	1.18	0.597	<0.001
/a/	544	1.03	0.374	<0.001
/e/	287	1.00	0.474	<0.001
/i/	270	0.379	0.350	<0.001

(a) F_1

	b_0	b_1	R^2	p
/u/	1120	1.28	0.276	<0.001
/o/	377	1.90	0.450	<0.001
/a/	946	1.00	0.264	<0.001
/e/	2094	0.093	0.002	0.571
/i/	2380	0.325	0.012	0.170

(b) F_2

Table 6. Values of b_0 and b_1 in Equation (2), coefficient of determination R^2 , and significant level p .

3. DISCUSSION

In the following, the major findings of this study on F_1 , F_2 , F_3 , and F_4 of *Shigin* singing are discussed, in comparison with the cases of operatic singing [2, 3, 7, 8], choir singing and choir speech [4], other traditional Japanese singing styles [5], and normal speech mode in Japanese [6]. In particular, perceptual aspects and production styles of *Shigin* singing are speculated on the basis of the observations in this study.

3.1 Values of F_3 and F_4 were higher than those for operatic singing or choir singing

Why were the values of F_3 and F_4 of the five Japanese vowels for *Shigin* singing were much higher than those for Western operatic singing or choir singing (Figure 2)? It can be speculated that the *Shigin* singer intonated vowels with his/her tongue tip in front and/or his/her larynx in higher position than that for operatic singing or choir singing. Note that (1) closing the tongue tip to the teeth and/or (2) lifting the larynx will generally cause the increase in F_3 and F_4 [8]. Since the value of F_4 is closely related to the perception of voice quality [8], such large difference in F_3 and F_4 would cause the difference in voice timbre.

3.2 The values of F_2 of /u/ vowel for *Shigin* singing were similar to those for normal speech mode in Japanese but higher than those for Western operatic singing or *Noh* singing

Remarkably, the values of F_2 of /u/ vowel for *Shigin* singing were similar to those for the normal speech mode in Japanese [6] but about 400 Hz higher than those for *Noh* singing [5] and about 800 Hz higher than those for Western operatic singing [2, 3] (Figure 1 (b)). A *Shigin* singer may intonate the /u/ vowel with his/her tongue position being similar to that for normal speech mode in Japanese rather than being similar to that for *Noh* singing or Western operatic singing, mindful that F_2 normally increases when moving the tongue from a front to a back position [7].

3.3 The variance of F_1 , F_2 , and F_3 can be significantly described by which vowel the *Shigin* singer intonated

Note that the variance in each of the F_1 , F_2 , and F_3 values was significantly described by the five Japanese vowels (Figures 3, 4, and 5). In particular, the variance of F_2 was 89.2% described by the five Japanese vowels (Figure 4). This ratio may be higher than those for operatic singing, normal speech mode [6] and other traditional Japanese singing styles [5]. *Shigin* singers might accurately adjust the opening of their mouths and jaws and the positioning of their tongues according to the targeted vowel. These findings show that it may be possible for the listeners of *Shigin* singing to easily recognize which vowel the *Shigin* singer intonated.

3.4 F_1 , F_2 , and F_3 of *Shigin* singing can be predicted by the four factors: Vowel, Gender, L_{Aeq} , and F_0 .

It is also interesting to note that the variance in each of the F_1 , F_2 , and F_3 values was significantly described by a linear model that employs the four variables: Vowel, Gender, L_{Aeq} , and F_0 (Table 5). According to this model, the louder voice generally cause the increase in F_1 , while the louder voice generally cause the decrease in F_2 and F_3 . The singer may open their mouths and jaws wider and move the tongue from a back to a front position when producing louder voice [8].

3.5 Formant tunings in *Shigin*: both F_1 and F_2 increased significantly with increasing phonation frequency (F_0)

Similar to the cases of operatic singing [3], significant formant tunings of F_1 and F_2 with phonation frequency (F_0) were observed (Tables 1, 2, 6, and Figure 6). Again, it can be speculated that *Shigin* singers adjust their utterance, in a complex manner, e.g., opening of their mouths and jaws, and varying the position of their tongues between forward and rear positions in the mouth, in order to attain such kinds of formant tunings of F_1 or F_2 [7]. Note that such formant tunings normally increase the audibility of the voice [8].

4. CONCLUSIONS

This study investigated the four lowest formant frequencies, F_1 , F_2 , F_3 , and F_4 , of sung vowels in the steady-state portions of vowels intonated by six *Shigin* singers in anechoic chamber. In this study, the singers' phonation frequency: F_0 was ranged between 152-433 Hz. Five of the major findings of this study, in comparison with the cases of operatic singing, choir singing, other traditional Japanese singing styles, and normal speech mode in Japanese, are summarized as follows:

1. F_3 was about 400–500 Hz higher and F_4 was about 1,000 Hz higher than those for Western operatic singing (Figure 2). F_4 was about 600-900 Hz higher than those for choir singing (Figure 2 (b)).
2. The mean F_2 of the /u/ vowel as intonated by *Shigin* singers was similar to those for the normal speech mode in Japanese but about 400 Hz higher than those for *Noh* singing and about 800 Hz higher than those for Western operatic singing (Figure 1(b)).
3. 89.2% of the variance in F_2 , 71.0% of the variance in F_1 , and 41.4% of the variance in F_3 , were described by five Japanese vowels (Figures 3, 4, and 5).
4. 92.1% of the variance in F_2 , 83.0% of the variance in F_1 , 56.5% of the variance in F_3 were described by a linear model that employs the four variables: Vowel, Gender, L_{Aeq} , and F_0 (Table 5).
5. Similar to the case of operatic singing, formant tunings in both F_1 and F_2 were observed: both F_1 and F_2 increased with increasing F_0 (Table 1, 2, 6, and Figure 6).

These findings are useful for e.g. singing voice synthesis, development of training support system, and understanding of the voice production mechanisms of singers.

ACKNOWLEDGEMENTS

The authors thank Prof. Takashi Yano and Dr. Keiji Kawai of Kumamoto University for providing access to a database of anechoic recordings *Shigin* utterances. The authors thank Mr. Syouhei Mizobuchi of Wakayama University for building a basic program for acoustic analysis, and also thank Prof. Seiichi Nakagawa of Toyohashi University of Technology for providing reference data on formant frequencies in Japanese vowels with the corresponding age and gender information [6].

REFERENCES

- [1] M. Nakayama, K. Kato, and M. Matsunaga, "Formant frequencies of sung vowels intonated by six traditional Japanese *Shigin* singers. Part I: Dataset construction and analysis method," *Proceedings of Stockholm Music Acoustics Conference 2013*, 2013.
- [2] J. Sundberg, "Articulatory interpretation of the 'singing formant'," *J. Acoust. Soc. Am.*, Volume 55, Issue 4, pp. 838–844, 1974.
- [3] T. F. Cleveland, "Acoustic properties of voice timbre types and their influence on voice classification," *J. Acoust. Soc. Am.*, Volume 61, Issue 6, pp. 1622–1629, 1977.
- [4] S. Ternström and J. Sundberg, "Formant frequencies of choir singers," *J. Acoust. Soc. Am.*, Volume 86, Issue 2, pp. 517–522, 1989.
- [5] I. Nakayama and M. Yanagida, "Introduction to database of traditional Japanese singing with examples of comparative studies on formant shifts and vibrato among genres," *Acoustical Science and Technology*, Volume 29, Issue 1, pp. 58–65, 2008.
- [6] S. Nakagawa, H. Shirakata, M. Yamao, and T. Sakai, "Differences in Feature Parameters of Japanese Vowels with Sex and Age," *Studia Phonologica*, XIV, pp. 113–132, 1980.
- [7] B. Lindblom and J. Sundberg, "Acoustical consequences of lip, tongue, jaw, and larynx movements," *J. Acoust. Soc. Am.*, Volume 50, pp. 1166–1179, 1971.
- [8] J. Sundberg, *The Science of The Singing Voice*, Northern Illinois University Press, 1987.

A PILOT STUDY OF VIBRATION PATTERN MEASUREMENT FOR FACIAL SURFACE DURING SINGING BY USING SCANNING VIBROMETER

Tatsuya Kitamura

Konan University

t-kitamu@konan-u.ac.jp

Hiroaki Hatano

Konan University

hatano.hiroaki@gmail.com

Takeshi Saitou

Kanazawa University

t-saitou@ec.t.kanazawa-u.ac.jp

Yui Shimokura

Showa University of Music

Eri Haneishi

Showa University of Music

haneishi@tosei-showa-music.ac.jp

Hiroko Kishimoto

Showa University of Music

kisimoto@tosei-showa-music.ac.jp

ABSTRACT

We attempted to measure the vibration velocity patterns of the facial surface during sustained singing using a scanning laser Doppler vibrometer. The measurement system used allowed laser-based, noncontact, and multipoint measurements of the vibration of objects. Three Japanese female professional singers (A, B, and C) participated in three experiments. In the first experiment, the vibration velocity patterns for the Japanese vowels /a/ and /i/ were obtained for the singers. The patterns for the vowels demonstrated a clear contrast, especially around the nose and cheek. In the second experiment, the differences in the vibration velocity patterns due to the pitch frequency were compared for Singer A. The results show that the amplitude of the vibration velocity of the forehead for a higher pitch frequency (F5=698.5 Hz) is larger than that for a lower pitch frequency (A4=440 Hz). In the third experiment, the effects of methods of vocalization on the vibration velocity pattern were measured while Singer A sang in falsetto and modal voice. A significant difference in the amplitude of the volume velocity of the cheek was observed.

1. INTRODUCTION

Expert singers can express their somesthesia during singing in their unique expressions. For example, one of the authors once heard of a singer who sings by expanding her skull – of course, we cannot expand our own skull by ourselves. If the causal relationship between these expressions and various physical phenomena on the body surface and inside the body is clarified, it could lead to the understanding of the singing mechanism and the improvement of methods of singing exercise. To this end, in the present study, we conducted a pilot study measuring one of such physical phenomena, the vibration velocity pattern of the facial surface during singing for expert singers.

The vibration patterns of skin surfaces during phonation

and singing have been measured by contact and noncontact methods [1]. In their pioneering work, Kirikae *et al.* [2] used vibrometers to measure phonatory body vibrations at more than forty points. Suzuki *et al.* [3] employed acceleration pickups and measured the acceleration of the nasal wall, the neck wall, and the cheek during speech production. Toyoda and Fujinami [4] utilized an optical fiber sensor to measure the vibration at nine different locations on the body during singing. It is, however, undeniable that the sensors used in contact with the measurement objects affect their vibration, although the acceleration pickup used by Suzuki *et al.* [3] is very light and small.

Noncontact methods, on the other hand, have a clear advantage in that they allow measurements without affecting the vibration of objects. Pawluczyk and Kraska [5] developed a laser-based method that can show equal-vibration-amplitude contours of objects and measured the nodal patterns of the facial and neck surfaces during singing.

Recently, laser Doppler vibrometers (LDVs) have been utilized to measure the skin surface vibration during phonation. Avargel and Cohen [6] attempted to obtain speech information from the vibration of the surface of the larynx using a single-point LDV. Kitamura [7] reported the vibration velocity pattern of the facial surface during phonation using a scanning LDV. The scanning vibrometer permits multipoint measurements of the vibration of targets.

In this research, we conducted a preliminary study to obtain the vibration velocity pattern of the facial surface during singing using the scanning LDV.

2. METHODS

2.1 Scanning laser Doppler vibrometer

The vibration velocity of the facial surface was obtained with a scanning LDV system, Polytec PSV-400-M4. The LDV is an optical transducer that senses the frequency shift of the light reflected from a vibrating surface on the basis of the Doppler effect and can determine the vibration velocity and displacement at a certain point [8]. The scanning LDV can scan and probe multiple points of a vibrating surface automatically.

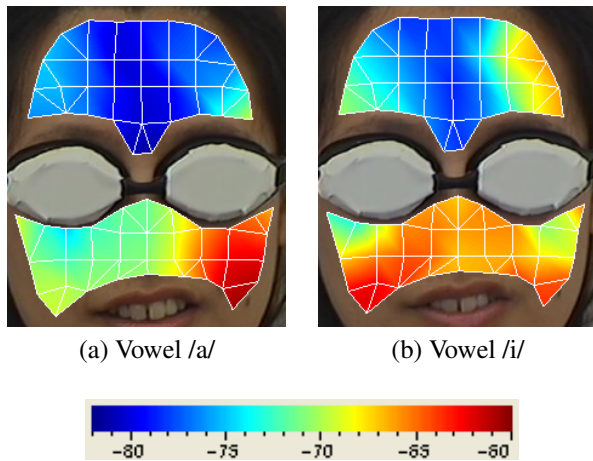


Figure 1. Vibration velocity of facial surface of Singer A during singing of Japanese vowels (a) /a/ and (b) /i/ in falsetto at her most comfortable pitch frequency ($A_4=440$ Hz). The unit is m/s [dB] and 0 dB is equal to 1 m/s.

2.2 Participants

Three Japanese female expert singers, who are faculty members of Showa College of Music, participated in the measurements. Two of them (Singers A and B) are classical singers and the other one (Singer C) is a musical singer.

2.3 Data acquisition

We measured the vibration velocity pattern of the facial surface from the frontal direction, which is perpendicular to the forehead. The scanning head of the vibrometer was mounted on a tripod approximately 2 m away from the head of the participant. The participant sat on a high-back chair. Her head was held from either side by two steel poles covered with polyurethane foam that were clamped to the back of the chair. Her neck was also stabilized by a vacuum cushion. She wore protective goggles to protect her eyes from the laser emission.

In the experiment, scanning points on the facial surface were first determined using system control software. During the measurement, the vibrometer scanned each point and obtained the vibration velocity of the point. One measurement point was probed within 1 s. The vibration velocity and singing voices were measured from 100 Hz to 10 kHz.

Data were acquired during sustained singing of a Japanese vowel. The participants were asked to sing repeatedly while keeping their head immobile during the measurements. Singing voices were recorded through a microphone (PCB PIEZOTRONICS 130E20).

Written informed consent was obtained from the participants prior to the measurement. The experimental protocol was approved by the ethical and safety committees of Konan University.

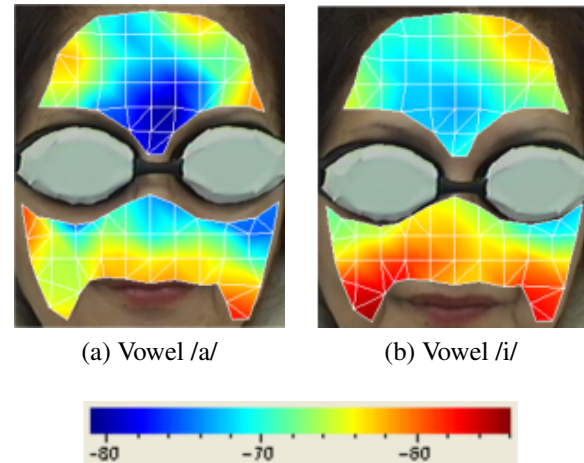


Figure 2. Vibration velocity of facial surface of Singer B during singing of Japanese vowels (a) /a/ and (b) /i/ in falsetto at her most comfortable pitch frequency ($?=??$ Hz). The unit is m/s [dB] and 0 dB is equal to 1 m/s.

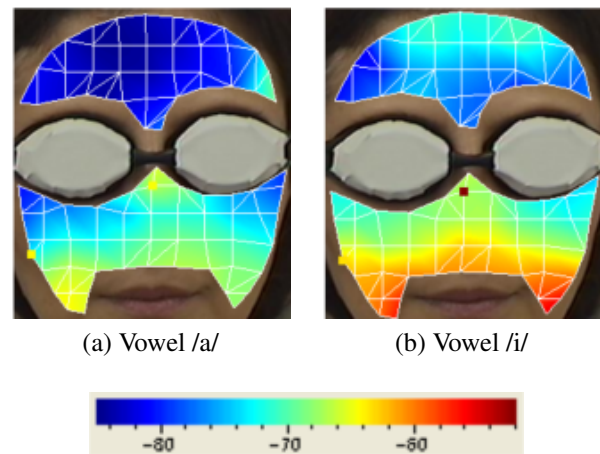


Figure 3. Vibration velocity of facial surface of Singer C during singing of Japanese vowels (a) /a/ and (b) /i/ in falsetto at her most comfortable pitch frequency ($?=??$ Hz). The unit is m/s [dB] and 0 dB is equal to 1 m/s.

2.4 Experiments

Three experiments were carried out in this study. In the first experiment, we compared the vibration velocity patterns of the facial surface during sustained falsetto singing of the Japanese vowels /a/ and /i/, back and front vowels. The three singers sang in their most comfortable pitch frequencies ($A_4=440$ Hz for Singer A, $G_4=392$ Hz for Singer B, and $A_4=440$ Hz for Singer C).

The second experiment was conducted to investigate the differences in the vibration velocity pattern due to the pitch frequency. Singer A sang in her higher comfortable pitch frequency ($F_5=698.5$ Hz) during the measurement.

The third experiment aims to reveal the differences between falsetto singing and modal singing. The vibration velocity pattern was measured while Singer A sang the vowel /i/ sustainedly in her modal voice and compared it

with the patterns measured in the first experiment. The pitch frequency of the modal voice of Singer A was the same as her comfortable pitch frequency in experiment 1.

3. RESULTS AND DISCUSSION

3.1 Experiment 1

Figures 1, 2, and 3 show the root mean square (RMS) values of the vibration velocities of the facial surfaces of the three singers. The grid points in the figures are measurement points and the RMS values of the other colored points were interpolated from the RMS values of adjacent measurement points. The cold and warm colors of the patterns indicate low and high vibration velocities, respectively.

The results show that the amplitude of the vibration velocity for the vowel /i/ is larger than that for the vowel /a/ around the nose and cheek for all the participants. This is probably because the variation in intraoral pressure during the production of the vowel /i/ is higher than that during the production of the vowel /a/.

The vibrometer utilized in this study can visualize the vibration velocity pattern for a specific vibration frequency band of interest. Figure 4 shows the vibration velocity pattern during sustained singing of the vowel /a/ for the harmonic frequencies (438 Hz and 875 Hz) and the frequencies intermediate between the harmonic frequencies (291 Hz, 657 Hz, and 1,094 Hz). This figure illustrates that the amplitude of the vibration velocity of the facial surface at the harmonic frequencies is significantly larger (approximately 20 dB) than that at the other frequencies.

3.2 Experiment 2

The vibration velocity pattern of the facial surface of Singer A for her higher comfortable pitch frequency is shown in Fig. 5. The vibration velocity of the forehead increased compared with that for the pitch frequency of A4 (Fig. 1). For the vowel /i/, the vibration velocity of the bilateral cheeks also increased, whereas that around the nose decreased.

Singer A reported that she vocalized so that the voice went through the top of her head when she sang in a high pitch frequency. Her introspection might correspond to the results for F5, that is, the vibration of the forehead increased probably because the singing voice radiated from there.

3.3 Experiment 3

The vibration velocity patterns of the facial surface of Singer A for her falsetto singing and modal singing are shown in Fig. 6, indicating that the vibration velocity patterns of the cheeks notably decreased for the modal singing even though the vowel and the pitch frequency were the same. Note that the result for the falsetto singing (Fig. 6(a)) is the same as that shown in Fig. 1(b).

Singer A has an image that the modal singing voice propagates inward of the body. Her introspection is possibly linked to the differences between Figs. 6(a) and 6(b).

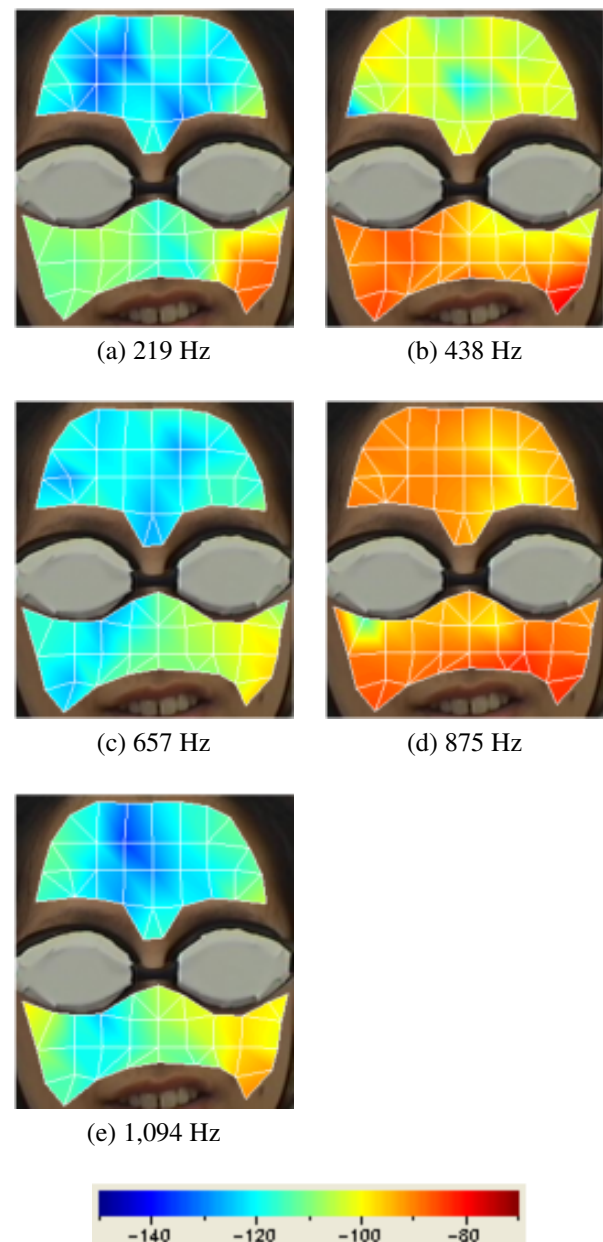


Figure 4. Vibration velocity of facial surface of Singer A during singing of Japanese vowel /a/ in falsetto at her most comfortable pitch (A4=440 Hz). The figures show the vibration velocity patterns at frequencies of (a) 219 Hz, (b) 438 Hz, (c) 657 Hz, (d) 875 Hz, and (e) 1,094 Hz. The unit is m/s [dB] and 0 dB is equal to 1 m/s.

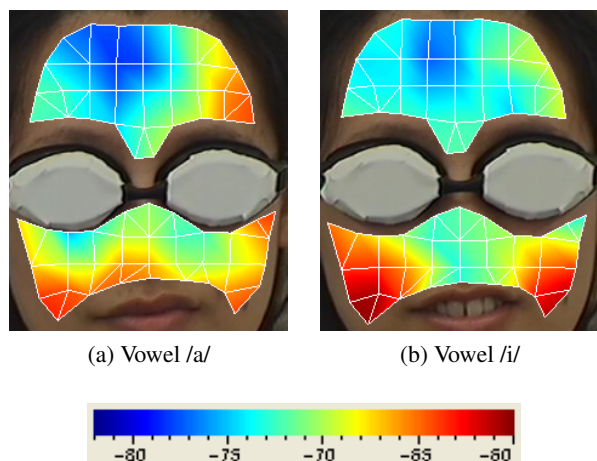


Figure 5. Vibration velocity patterns of facial surface of Singer A during singing of Japanese vowels (a) /a/ and (b) /i/ in falsetto at her higher comfortable pitch frequency ($F5=698.5$ Hz). The unit is m/s [dB] and 0 dB is equal to 1 m/s.

4. CONCLUSIONS

In this study, we reported the results of a pilot study of measuring the vibration velocity patterns of the facial surface during singing. The obtained results show clear contrasts between the patterns for the vowels, pitch frequencies, and vocalization methods. A scanning LDV proved to be extremely useful for measuring the vibration velocity patterns during singing, even though we need to confirm the reproducibility of the measurement.

The proposed method enables us to evaluate singing voices. The vibration patterns may be helpful as a visual feedback of a singing exercise. The vibration patterns may be easier to relate to the somesthesia than the spectra of the speech sounds or trainer's comments.

Acknowledgments

This study was supported by JSPS KAKENHI (nos. 21300071 and 24650088) and Hyogo Foundation for Science and Technology. The author wishes to thank Dr. Kazuhito Ito (National Institute of Advanced Industrial Science and Technology) for his generous assistance.

5. REFERENCES

- [1] J. Sundberg, "Phonatory vibrations in singers: A critical review," in *STL-QPSR*, 1991, pp. 37–51.
- [2] I. Kirikae, T. Sato, H. Oshima and K. Nomoto, "Vibration of the body during phonation of vowels," in *Revue de Laryngologie, Otologie, Rhinologie*, 85, 1964, pp. 317–345.
- [3] H. Suzuki, J. Dang and T. Nakai, "Measurement of sound and vibration at the lips, nostrils and pharynx wall in speech utterance and simulation of sound leakage from the oral cavity to the nasal cavity in non-

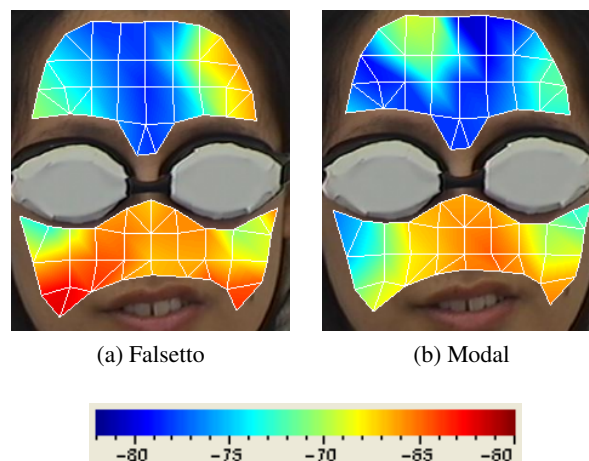


Figure 6. Vibration velocity patterns of facial surface of Singer A during singing of Japanese vowel /i/ in (a) falsetto and (b) modal voice at her most comfortable pitch frequency ($A4=440$ Hz). The unit is m/s [dB] and 0 dB is equal to 1 m/s.

nasal sounds," in *IEICE Transactions*, J74-A, 1991, pp. 1705–1714.

- [4] K. Toyoda and T. Fujinami, "An effective singing for musical expressions," in *Proc. of 10th International Conference on Music Perception and Cognition*, 2008, pp. 372–380.
- [5] R. Pawluczyk and Z. Kraska, "Diffuse illumination in holographic double-aperture interferometry," in *Applied Optics*, 1985, pp. 3072–3078.
- [6] Y. Avargel and I. Cohen, "Speech measurements using a laser doppler vibrometer sensor: Application to speech enhancement," in *Proc. of Joint Workshop on Hands-free Speech Communication and Microphone Arrays*, 2011, pp. 109–114.
- [7] T. Kitamura, "Measurement of vibration velocity pattern of facial surface during phonation using scanning vibrometer," in *Acoustical Science and Technology*, 2012, pp. 126–128.
- [8] Polytec web page, <http://www.polytec.com/>

POSTURAL SWAY IN VOCAL DUETS

Erik Koopmans, Caroline Palmer, Frances Spidle

Department of Psychology

McGill University

1205 Dr. Penfield Ave

Montreal, QC, Canada, H3A 1B1

{erik.koopmans, caroline.palmer, frances.spidle}@mcgill.ca

ABSTRACT

Postural control is important for singing, and is constrained by the demands of breath support. We examined postural sway in vocal duets and the influence of visual and auditory cues on posture during vocal performance. Experienced vocalists performed simple melodies in Solo, Unison (same musical part sung by two vocalists simultaneously), and Round conditions (same musical part sung at a temporal delay) while standing on force plates. Visual feedback about the partners' movements was manipulated by facing the singers Inward (with full view of each other) and Outward (with no view of each other), while normal acoustic feedback was maintained across all conditions. Measures of ground reaction forces and center of pressure both indicated increased variability in sway in Rounds compared with Unison performances, and when the vocalists had full visual cues in Inward-facing conditions. Correlations of vertical ground reaction force measures indicated more agreement between singers during production of Unison singing than Round singing. Although visual cues about one's partner had some influence on the variability of vocalists' posture during duet performance, only the type of performance (Unison/Round) influenced the degree of correspondence in posture between vocal duet members.

Experimental study of the frequency leap interval produced by the change of laryngeal vibratory mechanism during sustained notes

Sylvain Lamesch

IJLRA.

Equipe Lutheries,
Acoustique et Musique,
lamesch@lam.jussieu.fr

Boris Doval

IJLRA.

Equipe Lutheries,
Acoustique et Musique,
boris.doval@upmc.fr

Michèle Castellengo

IJLRA.

Equipe Lutheries,
Acoustique et Musique,
michele.castellengo@upmc.fr

ABSTRACT

The transitions between two different laryngeal vibratory mechanisms are often characterised by frequency jumps. The leap interval of these frequency jumps is studied for the transitions from M1 to M2 and conversely from M2 to M1. Its correlation with the starting fundamental frequency, the vocal intensity and the vowel is investigated.

Seven singers have produced sustained notes with laryngeal mechanisms transitions occurring during the production. The sound and the electroglottographic signals were recorded.

The leap intervals values depend on the subject. However global tendencies can be observed for most of them: the leap interval rises with the musical dynamics for the M1→M2 transition, and decreases with the frequency for some subjects. Concerning the M2→M1 transition, no tendency was observed. The frequency leap interval does not depend on the vowel; however the results show individual strategies.

The subglottal pressure at the beginning of the jump could play a role in the leap interval variation. Results show that the relation between the fundamental frequency and the subglottal pressure could be different in M1 and in M2.

1. INTRODUCTION

Classical singers use only their vocal apparatus in a "stable" configuration. However it is now known that instabilities can occur during phonation. These instabilities can be frequency jumps, period doubling or tripling, or phases of "chaos" (which can be produced at the physiological level by an irregular oscillation of the vocal folds). They are used in non-classical techniques (yodel, tahrir) [1], or specific singing styles like contemporary music [2].

These instabilities may result from different phenomena, especially acoustic [3] and biomechanic ones. This paper is about frequency jumps which are observed at the transition between the two main laryngeal vibration configurations of the vocal folds, which are commonly used in classical singing, contemporary commercial music as well as

in traditional music: the laryngeal vibratory mechanisms M1 and M2.

These two mechanisms are already well-documented and correspond to modal and falsetto registers when these terms refer exclusively to the laryngeal oscillation of the vocal folds. The vocalis, which corresponds to the internal layer of the vocal folds [4], participates with the vibration in M1, whereas it does not in M2. As a result, the surface of contact is greater in M1 than in M2 for a given fundamental frequency and vocal intensity [5]. The subglottal pressure is usually greater in M1 than in M2, the open quotient takes different values, etc. For more details, see [5].

One particularity of the frequency jumps is their very short duration (from a few milliseconds to 120 ms, depending on the observations and experimental conditions [6, 7]). It allows one to compare glottal elements that don't change much before and after the transition, such as regulation of subglottal pressure and articulatory adjustments.

Instabilities have been reproduced from excised canine [8] or human [9–11] larynx by changing continuously the longitudinal tension of the vocal folds. Biomechanical modeling (two, three mass model) [12, 13] has confirmed these observations, allowing authors to consider the laryngeal function as a non-linear dynamic system. A significant result is that a range of longitudinal tension can be observed both in M1 and in M2.

Studies with living subjects are based on different protocols: either yodel-like productions (with a change of note at the change of laryngeal mechanism) [9, 14, 15], or sustained notes or glissandos [6, 7]. All these studies show a great variability among subjects, and smaller frequency leap intervals for females than for males. However the relation between source or resonant parameters with the leap intervals remain poorly documented.

This paper aims at studying the M1→M2 and M2→M1 transitions for different vocal intensities, starting fundamental frequencies and vowels. In order to study these transitions, a well adapted experimental protocol has to be chosen. Svec et al [9] noticed for the yodel-like productions that the frequency jumps are composed of two parts: a transient part and a smoother variation (gliding part). According to them, only the transient part corresponds to the biomechanical phenomenon of laryngeal mechanism transition. The authors of this present paper hypothesize that a protocol based on sustained notes will allow them to distinguish more easily these two parts, and then to study pre-

cisely the transient part of the frequency jumps. This hypothesis will be discussed in part 4.

2. METHOD

2.1 Protocol

The instruction given to the subjects is to produce changes of laryngeal mechanisms while producing a sustained note. This instruction is realistic because the laryngeal mechanisms M1 and M2 have a great overlapping area in terms of fundamental frequency and vocal intensity [16, 17]. Subjects had to switch from one laryngeal mechanism to the other in the following order: M1-M2-M1-M2-M1.

Productions were asked for different notes and different levels. For each note, 3 or 4 levels were asked (depending on the subject). The notes were spaced by one tone. The task was first done for the vowel /a/, and then repeated for /i/ and for the close vowel /o/.

2.2 Subjects

Seven subjects were recruited for the experiment: 3 sopranos (S1, S3, S4), 1 mezzo-soprano (MS2), 1 tenor (T3), one baritone (Bar1) and one bass (B3). All the subjects were amateur or professional singers, so that they could be trained at smoothing the transition between the laryngeal mechanisms M1 and M2. All were untrained for the protocol. No one sang yodel-music or any similar vocal style.

The seven subjects had previously taken part in an experiment about Voice Range Profiles for the laryngeal mechanisms M1 and M2 separately and for the 3 vowels. [17]. Consequently the overlapping area of M1 and M2 was known by the experimentators for each subject.

2.3 Protocol

The sound signal was recorded with a 1/2" microphone (Brüel & Kjaer 4191), placed at 30cm from the singer's mouth. This microphone was connected to a preamplifier Brüel & Kjaer 2669 then to an amplifier Brüel & Kjaer Nexus 2690. The EGG signal was measured with an electroglottograph EG-2-PC with two pairs of electrodes, which allowed us to take into account the vertical movements of the larynx. The sound and EGG signals were recorded with a Metric Halo Mobile 2882 soundcard connected to a Macintosh Power Mac G5. A digital oscilloscope has also been used to visualize the EGG signal during recording.

The sound level was calibrated on a stable vocal production, without vibrato and with a speech voice quality. At the beginning of each set, the sound pressure level was measured at the location of the microphone to compute the gain of the data acquisition chain.

The recording was carried out in a room that was large enough (15 m²) for both the singer and the experimenter, which is more comfortable for singers and allows them to interact during the recording. It was a quiet room, isolated from the outside and with little reverberation (reverberation time at 60 dB: 0.3 s at 1kHz).

2.4 Data processing

2.4.1 Fundamental frequency

As the frequency jumps are especially fast (of the order of a few glottal cycles), the usual pitch determination algorithms, which are commonly based on an average frequency over a large window, do not allow to describe the dramatic pitch variation with enough precision. A solution could be to estimate the frequency cycle by cycle as the inverse of the fundamental period. However with this method, the frequency estimation is characterized by small instabilities just before and after the jumps which prevent one to estimate properly the frequency and therefore the leap interval. To solve this problem, the following procedure was used:

- the detection of the glottal closure instants (GCI) from the DEGG signal around the laryngeal mechanism transitions was manually validated;
- The fundamental frequency was computed as the inverse of the duration between two successive GCIs, and its curve was then passed through a 5 point median filter. Using a median filter allows one to cope with the instabilities rounding the jumps while keeping the leap interval.

2.4.2 Jump labelling

The starting time (t_{st}) and the ending time (t_{end}) of the frequency jumps were manually determined from the visualisation of several parameters: the radiated sound spectrogram, the fundamental frequency curve, the open quotient and the amplitude of the EGG signal. It allows to establish the onset frequency of the jump (f_{0st}), the ending frequency after the readjustment phase (f_{0end}), together with the frequency leap interval Δf_0 . All these parameters are illustrated on figure 1. The sound pressure level before leap number k has been computed. It corresponds to the mean SPL between the end instant of the $(k-1)^{th}$ leap and the begin instant of the k^{th} leap. The intensity after each leap is computed similarly. Therefore, for leap number k :

$$I_{db_bf}(k) = \text{mean}(I_{db}(t_{end}(k-1) : t_{st}(k)))$$

$$I_{db_af}(k) = \text{mean}(I_{db}(t_{end}(k) : t_{st}(k+1)))$$

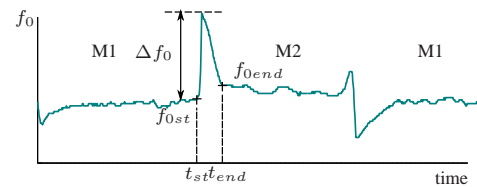


Figure 1. Adopted parameters to study the frequency jumps.

2.4.3 Fundamental frequency and voice sound level: a numerical estimation of the dynamics

Studying the correlation between a parameter and either the fundamental frequency or the sound level can be rather

complicated because these two parameters are linked: the average sound level increases with the frequency [18–21]. In order to study the correlation between a particular intensity scale (independent of the fundamental frequency) and the leap interval, a new intensity scale, SPL_{cor} , is proposed. It is defined as follows:

$$SPL_{cor} = SPL - k_v \cdot \log(f_0)$$

where k_v is the slope of the linear regression line linking $\log(f_0)$ and SPL computed on all the transitions of the same dynamics on a given vowel. It is estimated in mechanism M1 for each vowel separately, as the result of the average of the coefficient values computed for each dynamics produced by the singer. SPL_{cor} is a numeric scale expressed in dB, where the low values correspond to the pianissimi and high values to the fortissimi.

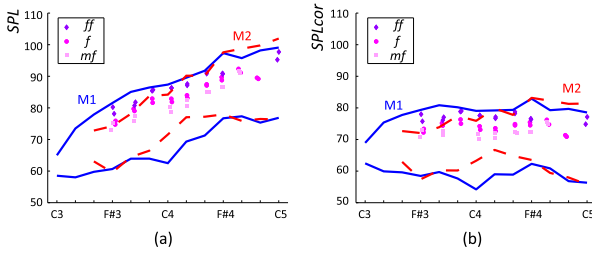


Figure 2. Voice Range Profiles (VRP) from MS2, for each mechanism, with indications of the starting positions of the jumps for the shift M1→M2. (a): Usual VRP (SPL vs f_0). (b): SPL_{cor} vs f_0 . On this example, 78 dB (SPL_{cor}) correspond to a fortissimo, whatever the fundamental frequency. Vowel: /i/.

3. RESULTS

3.1 Qualitative observation of transitions M1→M2 and M2→M1 using the EGG signal

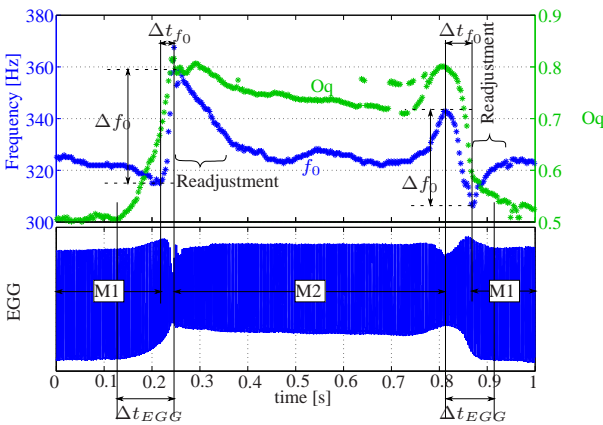


Figure 3. Example of laryngeal mechanisms transitions, produced by B3. Tone: E4, vowel /i/.

Figure 3 shows an example of M1→M2 and M2→M1 transitions, with the instruction to stay on the given note (here an E4, 330 Hz). The frequency leap that is noticeable

after 0.2s characterizes the M1→M2 transition. It goes along with a decrease of the EGG signal and an increase of the open quotient. At 0.81s, the downward frequency leap characterizes the M2→M1 transition.

As the protocol requires that the subject keep constant the fundamental frequency as much as possible, he/she has to readjust it after the transition. It is what is done by the subject between 0.24s and 0.35s for the M1→M2 transition and between 0.86s and 0.92s for the M2→M1 transition. Moreover, before the jump that characterizes the M1→M2 transition, the fundamental frequency decreases slightly, and increases before the M2→M1 transition.

Besides, for the M1→M2 transition, one can observe that the open quotient variation is much slower and more continuous than the frequency jump. For the given example, the open phase duration increases as soon as 0.12s while the frequency jump starts around 10ms later. Similarly, the variation of the EGG signal amplitude starts before the frequency jump.

Then, two phenomena are superimposed: the fast frequency jump, and the slower variations of the EGG signal amplitude and of the open quotient.

3.2 Study of the leap interval

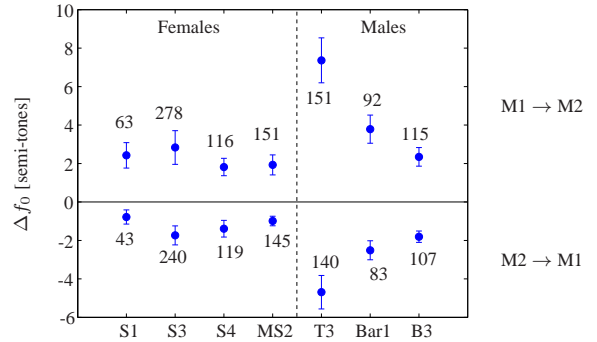


Figure 4. Frequency leap intervals (mean and standard deviation values), in semi-tones, for the 7 subjects, and for the 3 vowels. The positive values correspond to the M1→M2 transitions and the negative ones to the M2→M1 transitions. The indicated numbers correspond to the number of jumps.

Figure 4 shows the mean frequency leap intervals measured on the whole database. Some remarks can be made: while T3 shows a mean value of more than 7 semitones for the M1 to M2 transitions, S4 only shows a 2 semitone mean value. So there is an important variability between the different subjects with respect to the frequency leap interval.

The leap interval is greater for males than for females, for the M1→M2 direction as well as for the M2→M1 one (in absolute value). These observations (inter-subject variability and male-female differences) are in accordance with the results of Svec et al. [9] and Miller et al. [15] obtained respectively on two males and one female, and five males and six females (but, in both cases, with a very different recording protocol). They also support the results of Roubeau et

al. [7] obtained this time with a similar protocol of sustained notes.

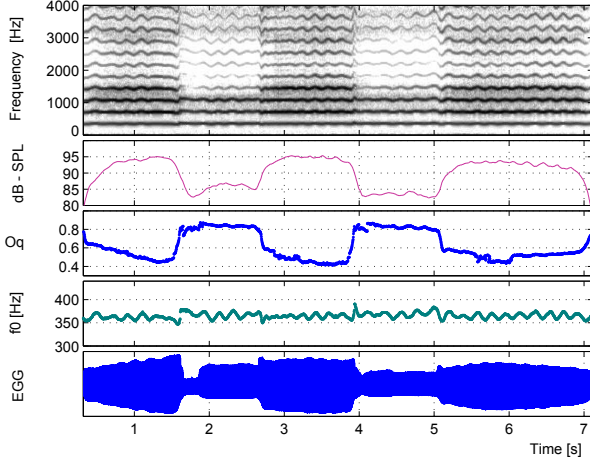


Figure 5. An example of the required production. The alternance of laryngeal vibratory mechanisms is characterized by the variation in terms of EGG amplitude and open quotient values. But the leap intervals are small. Vowel: /a/, tone: F#4.

Figure 5 shows an example of a voice production with alternating laryngeal mechanisms that shows very small frequency jumps, or sometimes even no jump at all. The open quotient changes (between around 0.5 and 0.8) and the EGG signal amplitude changes are related to the alternating laryngeal mechanisms imposed by the protocol. However, the corresponding frequency leap intervals are very small (between one and two semitones), or cannot even be seen as it is the case for the last M2→M1 transition (around 5s) where the frequency control loss (1.3 semitones) is close to the vibrato amplitude (0.74 semitones).

Finally, for each singer, **the M2→M1 leap intervals are smaller than the M1→M2 ones**. This asymmetry has been scarcely described in the literature because the usual protocols cannot show it clearly [9, 15]. Roubeau et al. [22] pointed to this before in a protocol similar to ours using sustained notes. When the singer do not try to readjust the fundamental frequency after the transition, his/her production is of a “yodel” type, moving from a given frequency in M1 to another one in M2. This type of production cannot highlight the asymmetry of the leap intervals along with the direction of the transition.

3.3 Vocal intensity and fundamental frequency

3.3.1 Direction M1→M2

Table 1 shows the partial correlation coefficients for the frequency leap intervals Δf_0 and the starting frequency f_{0st} or the vocal intensity before the jump SPL_{bf} , for the data which were obtained for the M1→M2 transition. Results show that the partial correlation coefficients between Δf_0 and f_{0st} are mostly negative, and the ones between Δf_0 and SPL_{bf} are mostly positive. It means that **the frequency leap interval increases with the vocal dynamics**

singer	/a/		/o/		/i/	
	f_{0st}	SPL_{bf}	f_{0st}	SPL_{bf}	f_{0st}	SPL_{bf}
S1	-0.37	0.46	-0.65	-0.09	-0.14	-0.11
S3	0.51	-0.16	-0.16	0.10	-0.02	0.09
S4	-0.46	0.82	-0.31	0.47	-0.53	0.65
MS2	-0.30	0.62	-0.60	0.67	-0.47	0.42
T3	-0.79	0.78	-0.81	0.85	-0.89	0.82
Bar1	-0.43	0.78	-0.53	0.75	-0.78	0.89
B3	-0.85	0.86	-0.30	0.47	-0.92	0.92

Table 1. Partial correlation coefficients for Δf_0 (in semitones) and f_{0st} or SPL_{bf} , for the M1→M2 transitions. Results are presented separately for the 3 vowels and the 7 singers.

and decreases with the starting fundamental frequency. Besides, **strong correlations were mostly observed for male subjects**.

3.3.2 Direction M2→M1

singer	/a/		/o/		/i/	
	f_{0st}	SPL_{bf}	f_{0st}	SPL_{bf}	f_{0st}	SPL_{bf}
S1	-0.43	0.04	-0.22	-0.28	0.32	-0.31
S3	-0.43	0.05	-0.18	0.16	-0.10	0.01
S4	-0.04	-0.21	0.10	-0.25	0.00	-0.02
MS2	-0.00	-0.48	0.22	-0.22	0.11	-0.19
T3	0.00	-0.20	-0.15	-0.19	0.31	-0.54
Bar1	-0.18	0.06	0.10	-0.37	0.32	-0.43
B3	0.31	-0.36	-0.02	0.05	0.55	-0.52

Table 2. Partial correlation coefficients for Δf_0 (in semitones) and f_{0st} or SPL_{bf} , for the M2→M1 transitions. Results are presented separately for the 3 vowels and the 7 singers.

Table 2 shows the partial correlation coefficients for the leap intervals measured at **M2→M1 transitions**. Since leap intervals are negative in this case, one must take the opposite of the correlation coefficients to obtain a description of the correlations between Δf_0 (absolute value) and f_{0st} or SPL_{bf} .

The most important factor is the **absence of clear tendency** contrary to what happens in the M1→M2 transitions: no strong correlation was obtained and only few coefficients greater than 0.5 (absolute value) were found (no one for females and 3 for males). **The M2→M1 transition is not the reverse phenomenon of the M1→M2 one.**

3.4 The influence of vowels on the leap intervals

Figure 6 presents the leap interval for the transition M1→M2, for the 3 different vowels and for different levels of SPL_{cor} . For the 7 subjects, the evolution of the Δf_0 values with SPL_{cor} is very similar for the 3 vowels. Consequently, one can say that **the influence of the vowel on the frequency leap interval is at most a second order effect**, whereas the influence of the dynamics corresponds rather to a first order effect.

Results also show that **the influence of the vowel depends on the singer**. Indeed for T3, the frequency leap interval is smaller for /i/ than for /o/ (especially for high levels). For S3, the leap intervals are also smaller on /i/ than on /a/ and /o/, but on the contrary, S4 obtained her larger leap intervals on the vowel /i/.

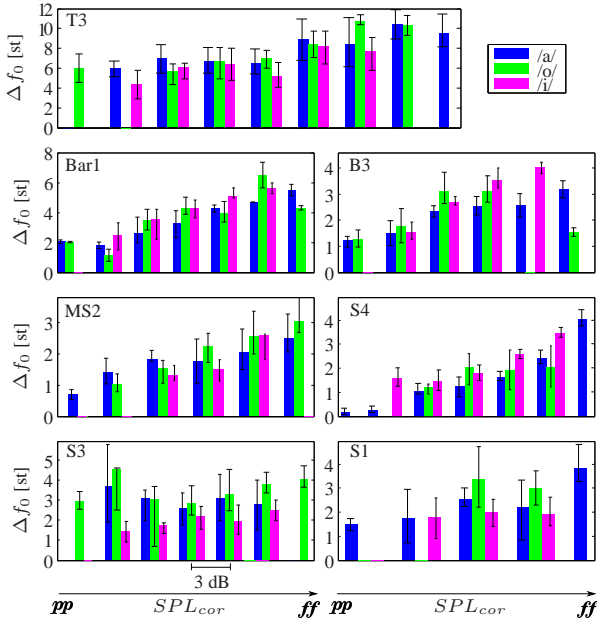


Figure 6. Variations of the frequency leap interval according to SPL_{cor} , for M1→M2 transitions, for the different vowels and for the 7 subjects. Each group of bars is spaced by 3 dB (SPL_{cor}).

3.5 Explored phonetographic area

Figure 7 shows the Voice Range Profile (VRP) of S3, T3 and B3 for the vowels /a/ and /i/. In these representations, the starting and ending positions of the frequency jumps for the M1→M2 transitions are added. It shows that the starting positions cover quite largely the overlapping area for S3 and T3 (G#3 - C5 for S3, E3 - G4 for T3). However the jumps for B3 cover only the high frequency range of the overlapping area (Bb3 - G#4). It confirms the ability of singers to produce laryngeal mechanisms transitions on a large frequency range (larger than one octave for S3 and T3), even if these subjects are non-expert for these vocal productions.

Figure 7 shows also interesting aspects regarding the vocal intensities: T3 covered a large level range in M1, but the ending levels in M2 are all located at a smaller level range. B3 essentially explored his fortissimo productions in M1 as well as in M2. These observations may provide a partial explanation of the inter-subject variability.

4. DISCUSSION AND CONCLUSION

The electroglottographic study shows that the uncontrolled frequency jumps are going along with a fundamental frequency readjustment produced by the singer to return to the targeted value. This readjustment is very similar to the slow frequency variation observed by Svec et al [9] on a yodel-like protocol. The difference between these two protocols relies upon the direction of frequency variation. In the sustained notes protocol, the frequency variation of the readjustment is in the opposite direction of the frequency variation of the transient part, while in the yodel-like protocol the direction of the frequency variation is the same for

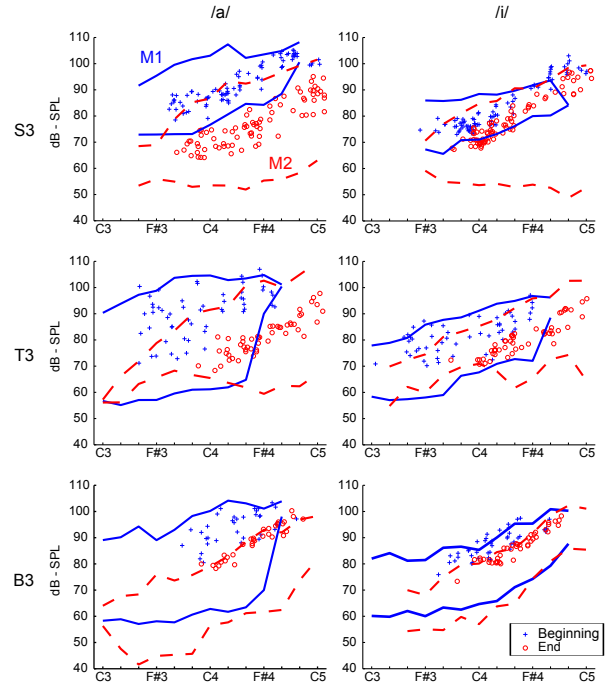


Figure 7. Starting (f_{0st} , SPL_{bf}) and ending (f_{0end} , SPL_{aft}) positions of the frequency jumps for the M1→M2 transitions, 3 singers, and the vowels /a/ and /i/.

both, which makes it difficult to identify the two phases. The sustained notes protocol is therefore more efficient to isolate the transient part which corresponds to the phenomenon of voice instability. However, it is possibly more difficult to produce for the subjects.

One can link the observed results to the physiology and the acoustics of voice production. The vocal fold oscillation frequency is regulated mainly by three parameters: the vibrating mass, the longitudinal tension and the vibrating length. For the vocal folds to oscillate, an air flow must be provided which is adapted to the laryngeal system. The subglottal pressure and the glottal flow result from the equilibrium of the laryngeal resistance (which depends amongst others on the parameters that regulate the fundamental frequency) and from the forces that generate the air flow (resulting from the breathing muscles).

At the muscular level, the vibrating mass is mainly determined by the thyroarytenoid (TA) muscle, while the tension and the length of the vocal folds are due to the equilibrium between the thyroarytenoid and the cricothyroid (CT) muscle. The laryngeal resistance depends more on the TA and other muscles as the inter-arytenoids and the lateral cricothyroids [23].

The main difference between M1 and M2 mechanisms is the involvement of the TA muscle in the vocal fold oscillation itself. During the M1→M2 transition, the decoupling of the different layers which constitute the vocal folds triggers a dramatic decrease of the vibrating mass, which in turn is probably responsible for a large part of the frequency jump. Inter-subject differences may then be linked to vocal fold morphological and internal constitution differences. The fact that on average women have a smaller and a lighter vocal muscle than men could in particular ex-

plain why their frequency jumps are smaller.

In their works, Tokuda and Horacek et al [10–13] show that it is possible to reproduce laryngeal mechanism transitions by continuously varying the longitudinal tension, or more precisely the CT activity. For experimental reasons, these experiments are carried out at a constant flow rate. Large et al. [24] experiments show that the M1→M2 transition is marked by a dramatic variation of the flow (the flow being larger in M2 than in M1). Besides, Miller et al. [15] set up a protocol that requires the subject to produce their transitions at a constant subglottal pressure. These observations allow us to put forward the hypothesis that **during the transition, the CT activity and the subglottal pressure are not varying much**. In the presented protocol, those two parameters can be subject to little or no variation during the transitions, but can be modified during the fundamental frequency readjustment phase. **The leap interval could provide us an indication of the pitch difference between the vocal productions emitted in M1 and in M2 with the same subglottal pressure and CT activity**.

The leap interval is smaller in the M2→M1 transition than in the M1→M2 one. An explanation could be that the subglottal pressure is lower in M2 than in M1 [25] so that the subglottal pressure measured at the jump is lower in the M2→M1 direction than in the M1→M2 one for given initial fundamental frequency and sound level.

In mechanism M1, the vocal sound level is linked to the subglottal pressure [26, 27]. Besides, the results show that the frequency leap interval increases with the dynamics in the M1→M2 direction. These two observations lead to the following hypothesis: **for a given CT activity, the greater the pressure, the larger the frequency difference between the two mechanisms**. Therefore, **the role played by the subglottal pressure in the fundamental frequency control could be different in M1 and in M2**.

In the M2→M1 direction, the frequency leap interval is not correlated with the dynamics. This result clearly shows that when considering only such parameters as the fundamental frequency and the sound level, **the M2→M1 transition does not correspond to the inverse phenomenon of the M1→M2 transition**. As reported by Svec et al [9], the M2→M1 transition is produced at a lower tension than the M1→M2 one because of the hysteresis which is characteristic of the phenomenon. Aerodynamic and biomechanic supplementary data are needed to interpret the obtained results.

The results show that despite the rapidity of the fundamental frequency jump, some glottal parameters like the EGG amplitude or the open quotient (which takes different values in M1 and M2), vary much slower than the fundamental frequency. Roubeau et al [6] already made a similar observation about the EGG while stressing the fact that the duration of variation does not depend on the transition direction nor on the vocal training. As a matter of fact, in the M1→M2 direction, the results show that the contact surface decreases but that the open phase duration begins to increase before the beginning of the jump. These observations suggest that a preparation of the mechanism tran-

sition takes place before the loss of control itself. Further investigation is needed to describe in more detail the observed phenomenon.

Acknowledgments

The authors warmly thank the seven singers that generously accepted to pass this quite constraining protocol.

This work is a part of Sylvain Lamesch's Phd thesis, UPMC - univ. Paris 6 [28], financed by the Fondation de France.

5. REFERENCES

- [1] M. Castellengo, "Continuité, rupture, ornementation, ou les bons usages de la transition entre deux modes d'émission vocale," *Cahiers de musiques traditionnelles*, vol. 4, pp. 155–165, 1991. [Online]. Available: <http://ethnomusicologie.revues.org/1583>
- [2] J. Neubauer, M. Edgerton, and H. Herzel, "Nonlinear phenomena in contemporary vocal music," *J. Voice*, vol. 18, no. 1, pp. 1 – 12, 2004.
- [3] I. R. Titze, "Nonlinear source-filter coupling in phonation: Theory," *J. Acous. Soc. Am.*, vol. 123, no. 5, pp. 2733–2749, May 2008.
- [4] M. Hirano, *Vox Humana*. Institute of the Finnish Language and Communication, 1982, ch. The role of the layer structure of the vocal folds in register control, pp. 50–62.
- [5] B. Roubeau, N. Henrich, and M. Castellengo, "Laryngeal vibratory mechanisms: The notion of vocal register revisited," *J. Voice*, vol. 23, no. 4, pp. 425 – 438, 2009.
- [6] B. Roubeau, C. Chevrier-Muller, and C. Arabia, "Electroglottographic study of the changes of voice register," *Folia Phoniat.*, vol. 39, pp. 280–289, 1987.
- [7] —, "Control of laryngeal vibration in register change," in *Vocal Fold Physiology Conference, Stockholm*. Gauffin J, Hammaberg B, eds., 1989, pp. 279–286.
- [8] A. Shiotani, H. Fukuda, M. Kawaida, and J. Kanzaki, "Vocal fold vibration in simulated head voice phonation in excised canine larynges," *European Archives of Oto-Rhino-Laryngology*, vol. 253, pp. 356–363, 1996.
- [9] J. G. Svec, H. K. Schutte, and D. G. Miller, "On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges," *J. Acous. Soc. Am.*, vol. 106, no. 3, pp. 1523–1531, September 1999.
- [10] J. Horacek, J. Svec, J. Vesely, and E. Vilkman, "Experimental study of the vocal-fold vibration in excised larynx: Measurement set-up and techniques," in *Interaction and Feedbacks '2000*, I. Zolotarev, Ed., 2000, pp. 27–34.

- [11] J. Horacek, J. G. Svec, J. Vesely, E. Vilkman, I. Klepacek, and A. Vetesnik, "Measurement of the vocal-fold vibration behaviour in excised human larynges," in *2nd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications*, 2001.
- [12] I. T. Tokuda, J. Horacek, J. G. Svec, and H. Herzel, "Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments," *J. Acous. Soc. Am.*, vol. 122, no. 1, pp. 519–531, 2007.
- [13] I. T. Tokuda, M. Zemke, M. Kob, and H. Herzel, "Biomechanical modeling of register transitions and the role of vocal tract resonators," *J. Acous. Soc. Am.*, vol. 127, no. 3, pp. 1528–1536, 2010.
- [14] J. Svec and J. Pesak, "Vocal breaks from the modal to falsetto register," *Folia Phoniatr. Logop.*, vol. 46, pp. 97–103, 1994.
- [15] D. G. Miller, J. G. Svec, and H. K. Schutte, "Measurement of characteristic leap interval between chest and falsetto registers," *J. Voice*, vol. 16, no. 1, pp. 8–19, 2002.
- [16] B. Roubeau, M. Castellengo, P. Bodin, and M. Ragot, "Phonétogramme par registre laryngé," *Folia Phoniatr. Logop.*, vol. 56, pp. 321–333, 2004.
- [17] S. Lamesch, B. Doval, and M. Castellengo, "Toward a more informative voice range profile: The role of laryngeal vibratory mechanisms on vowels dynamic range," *J. Voice*, vol. 26, no. 5, pp. 672.e9 – 672.e18, 2012.
- [18] I. R. Titze, "Acoustic interpretation of the Voice Range Profile (Phonetogram)," *J. Speech Hear. Res.*, vol. 35, no. 1, pp. 21–34, 1992.
- [19] A. M. Sulter, H. K. Schutte, and D. G. Miller, "Differences in phonetogram features between male and female subjects with and without vocal training," *J. Voice*, vol. 9, no. 4, pp. 363–377, 1995.
- [20] N. Henrich, C. d'Alessandro, B. Doval, and M. Castellengo, "Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency," *J. Acous. Soc. Am.*, vol. 117, no. 3, pp. 1417–1430, March 2005.
- [21] J. P. H. Pabon, "Objective acoustic voice-quality parameters in the computer phonetogram," *J. Voice*, vol. 5, no. 3, pp. 203–216, 1991.
- [22] B. Roubeau, "Mécanismes vibratoires laryngés et contrôle neuromusculaire de la fréquence fondamentale," Ph.D. dissertation, Université Paris XI, Orsay, 1993.
- [23] J. Lacau St Guily and B. Roubeau, "Voies nerveuses et physiologie de la phonation," *EMC Oto-rhino-laryngologie*, 1994.
- [24] J. Large, S. Iwata, and H. von Leden, "The male operatic head register versus falsetto," *Folia Phoniatr.*, vol. 24, pp. 19–29, 1972.
- [25] J. Sundberg and C. Högset, "Voice source differences between falsetto and modal registers in counter tenors, tenors and baritones," *Log. Phon. Vocol.*, vol. 26, pp. 26–36, 2001.
- [26] I. R. Titze and J. Sundberg, "Vocal intensity in speakers and singers," *J. Acous. Soc. Am.*, vol. 91, no. 5, pp. 2936–2946, May 1992.
- [27] J. Sundberg, M. Andersson, and C. Hultqvist, "Effects of subglottal pressure variation on professional baritone singers' voice sources," *J. Acous. Soc. Am.*, vol. 105, no. 3, pp. 1965–1971, 1999.
- [28] S. Lamesch, "Mécanismes laryngés et voyelles en voix chantée. Dynamique vocale, phonétogrammes de paramètres acoustiques et spectraux, transitions de mécanismes," Ph.D. dissertation, UPMC univ. Paris 6, 2010.

AN ATTEMPT TO DEVELOP A SINGING SYNTHESIZER BY COLLABORATIVE CREATION

Masanori Morise

Faculty of Engineering, University of Yamanashi, Japan
mmorise@yamanashi.ac.jp

ABSTRACT

This paper presents singing synthesizers collaboratively designed by several developers. In the video-sharing Web site Nico Nico Douga, many creators jointly create songs with a singing synthesis system called Vocaloid. To synthesize various styles of singer, another singing system UTAU which is a free software, is being developed and used by many creators. However, the sound quality of this system has not been yet as good as Vocaloid. The purpose of this study is to develop a singing synthesizer for UTAU by collaborative creation. Developers were encouraged to design a singing synthesizer by using a high-quality speech synthesis system named WORLD that can synthesize a singing voice that sounds as natural as a human voice. We released WORLD and a singing synthesizer for UTAU as free software with C language source code and attempted to encourage collaborative creation. As a result of our attempt, six singing synthesizers for UTAU and two original singing synthesis systems were developed and released. These were used to create many songs that were evaluated as high-quality singing by audiences on a video-sharing Web site Nico Nico Douga.

1. INTRODUCTION

Singing synthesis is a major research target in the field of sound synthesis, and several commercial applications such as Melodyne and Auto-Tune have already been used to tune singing voices. Text-To-Speech synthesis systems for singing have been released as computers are sufficiently developed. However, the sales of these applications have been poor.

After the releasing of the Vocaloid 2 Hatsune Miku [1], singing synthesis systems have played an important role in entertainment culture on the video-sharing Web site Nico Nico Douga, and many amateur creators have been uploading songs to the site. Several studies for Vocaloid have been carried out to synthesize natural singing voices [2, 3]. As a result, “Vocaloid music” is now a category of Japanese pop culture, in what has been termed the “Hatsune Miku Phenomenon” [1].

“Social Creativity” [4] which is a collaborative creation [5] by multiple creators, has been gaining popularity as a

new style of creation to improve the quality of contents on video-sharing Web sites. Today, many creators jointly create many contents including songs, promotional videos, and comments¹.

The purpose of this study is to develop singing synthesizers by collaborative creation. We implemented a base technology named WORLD and a singing synthesizer for UTAU and released them on the Web site to encourage collaboration. This attempt functions as indirect support for creating songs by the developed synthesizers.

The rest of this article is organized as follows. Section 2 describes conventional singing synthesis systems and outlines the requirements for developing one. In Section 3, we explain the principle behind and effectiveness of the base technology, WORLD. Section 4 reveals whether the synthesizers could be developed by other developers and discusses the result of our attempt. We conclude in Section 5 with a brief summary of our research and a mention of future work.

2. SINGING SYNTHESIS SYSTEMS: VOCALOID AND UTAU

Vocaloid is a Text-To-Speech synthesis system for singing in which creators can synthesize the singing using lyrics and scores. The demands to synthesize various kinds of singing have been growing due to the rapid growth of Vocaloid, but it has been virtually impossible to support this demand. To solve this problem, UTAU² was developed as a singing synthesis system.

2.1 UTAU

UTAU is a Japanese singing synthesis system similar to Vocaloid. As shown in Fig. 1, the framework consists of an editor to manipulate parameters, a synthesizer, and a voice library associated with the singer. UTAU can switch the voice library to synthesize various styles of singing and synthesizers to improve the sound quality. Although Vocaloid has a few voice libraries (around 20), UTAU has far more (over 5,000) because creating a voice library of amateur singers is easy.

The following sections describe the voice library for UTAU and the requirements for the singing synthesizer. To develop a synthesizer for UTAU, it is necessary to adjust the format for the voice and labeling data.

¹ On Nico Nico Douga, audiences can effectively overlay text comments onto video content. This feature provides the audience with the sense of sharing the viewing experience [6].

² <http://en.wikipedia.org/wiki/Utau>

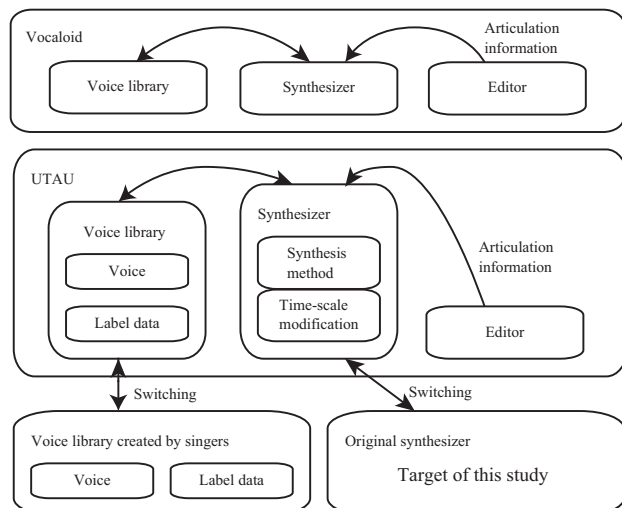


Figure 1. Overview of Vocaloid and UTAU. They consist of an editor, voice library, and synthesis method. Switching voice libraries and synthesizers is possible in UTAU.

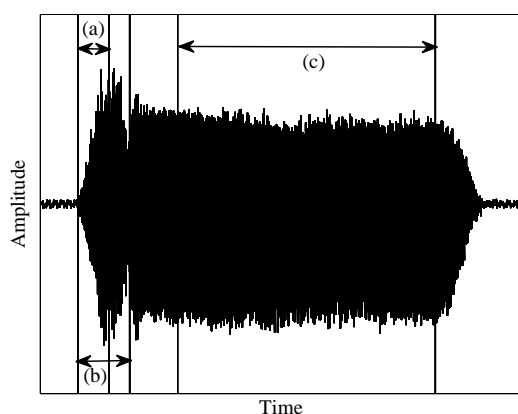


Figure 2. Labeling data of a singing voice /shi/. UTAU requires determining these positions.

2.1.1 Voice library

UTAU supports both CV and VCV synthesis for Japanese singing³, and recording all phonemes is required for the voice library. The recorded voices are then labeled to fulfill the requirement for UTAU. This is done automatically by using free software developed by another developer.

2.1.2 Labeling data

Figure 2 shows an example of labeling data in a CV voice /shi/. Three intervals are used for synthesis: Interval (a) represents the interval for smoothly mixing two voices, Interval (b) represents the consonant interval, with the endpoint used as the origin of a CV voice, and Interval (c) represents the voiced speech and is used to stretch the duration of the voice. A phoneme boundary between V and C is added for VCV voices.

³ Singing in other languages is available, but it is difficult because the editor cannot support other languages.

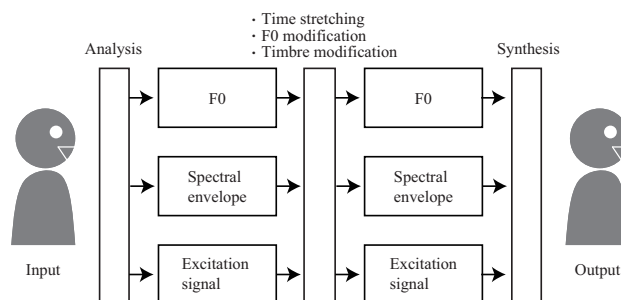


Figure 3. Overview of WORLD.

To develop the singing synthesizer for UTAU, developers must implement at least the following functions

- Time-stretching function
- F0-modification function
- Timbre-modification (at least Formant shift) functions

These functions are used to adjust the voices to the desired musical note, and developers can implement these functions by their original algorithm. However, F0-modification function must be implemented based on two processes that normalize F0 contour to musical scale and add the detailed contour given by the editor.

2.2 Problem of UTAU

Three synthesizers officially released by the UTAU developer to manipulate sensitive atmospheres. Other three synthesizers have been released by other developers because differences between synthesizers affect the quality of the content. Since the sound quality often deteriorates by the compatibility between the voice library and synthesizer, various types of synthesizers should be developed.

In this study, we developed a singing synthesizer for UTAU based on a high-quality speech synthesis and attempted to induce collaborative creation by releasing its C language source code.

3. DEVELOPMENT OF A SINGING SYNTHESIZER BASED ON WORLD

The base technology WORLD is a vocoder-based system [7] that decomposes the input speech into an F0, a spectral envelope, and an excitation signal and is able to synthesize a speech that sounds as natural as a human speech. Three parameters can be modified to fulfill the requirements for UTAU: time stretching, F0 modification, and timbre modification (as shown in Fig. 3).

WORLD estimates the F0 and then uses the F0 information to estimate the spectral envelope. The excitation signal is then extracted by the F0 and spectral envelope information.

3.1 DIO: F0 estimation method

F0 of a voiced sound is defined as the inverse value of the shortest period of glottal vibrations. It is one of the

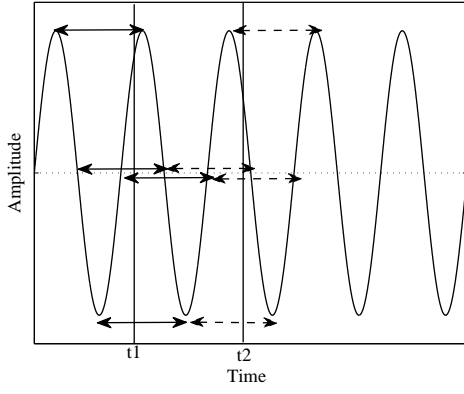


Figure 4. Four intervals used for determining the F0. Inverse value of the average is an F0 candidate, and that of the standard deviation is used as the index to determine the best of the candidates.

most important parameters for speech modification. Many F0 estimation methods, (such as Cepstrum [8] and auto-correlation-based method [9]) have therefore been proposed for accurate estimation. Although these methods can accurately estimate F0, they require extensive calculation such as FFT.

DIO [10] is a rapid F0 estimation method for high-SNR speech that is based on fundamental component extraction. The fundamental component is extracted by low-pass filters and the F0 is calculated as its frequency. Since the cut-off frequency to extract only the fundamental component is unknown, DIO uses many low-pass filters with different cut-off frequencies and the periodicity score to determine the final F0 of all candidates.

DIO consists of three steps to calculate F0 candidates and periodicity scores:

- Step1: Filtering by many low-pass filters with the different cutoff frequencies from low frequency to high frequency.
- Step2: Calculation of F0 candidates and periodicity scores.
- Step3: Determination of final F0 based on periodicity scores.

In the first step, the input waveform is filtered by many low-pass filters. DIO uses a Nuttall window [11] as a low-pass filter with a sidelobe of around -90 dB. The filtered signal is the sine wave with F0 Hz, provided that the filter is designed so that only the fundamental component is extracted.

In the second step, four intervals (shown in Fig. 4) are calculated at all temporal positions. The four intervals are defined as the negative and positive going zero-crossing intervals and the intervals between the peaks and dips. In the case that the filtered signal is a sine wave, four intervals indicate the same value, and the inverse value of the average of four intervals indicates the F0. That of the standard deviation can be used as its periodicity score.

In the final step, the F0 with the highest periodicity score is selected as the final F0. DIO calculates the F0 much faster than the conventional method because it does not use frame-by-frame processing to calculate FFT. The F0 estimation performance of DIO is the same as the conventional techniques [12–14], while the elapsed time is at least 28 times faster than the other methods [15].

3.2 STAR: Spectral envelope estimation method

Since voiced speech has an F0, the speech waveform includes not only the spectral envelope but also the F0 information. Many methods based on linear predictive coding (LPC) [16] and Cepstrum [17] have been proposed. Among them, STRAIGHT [18] can accurately estimate the spectral envelope and can synthesize high-quality speech. TANDEM-STRAIGHT [19] produces the same results as STRAIGHT but at a lower computational cost, and STAR can reduce the computational cost even more than TANDEM-STRAIGHT [20]. To calculate the spectral envelope, TANDEM-STRAIGHT uses two power spectra windowed by two window functions, whereas STAR produces the same result using only one power spectrum.

In STAR, spectral envelope $|\hat{H}(\omega, \tau)|$ is given by

$$|\hat{H}(\omega, \tau)|^2 = \exp \left(\frac{2}{\omega_0} \int_{-\frac{\omega_0}{2}}^{\frac{\omega_0}{2}} \log(|S(\omega + \lambda, \tau)|) d\lambda \right), \quad (1)$$

where $S(\omega, \tau)$ represents the spectrum of the windowed waveform and τ represents the temporal position for windowing. A Hanning window, which is used as the window function, has a length of $3T_0$ and is based on pitch synchronous analysis [21]. ω_0 represents fundamental angular frequency ($2\pi f_0$). By windowing this window function and by smoothing with Eq. (1), $|\hat{H}(\omega, \tau)|^2$ is temporally stable.

Figure 5 shows an example of the estimation result. The target spectral envelope consisted of a pole and a dip, but LPC could not accurately estimate the envelope from the spectral envelope including dip. In contrast, TANDEM-STRAIGHT and STAR could, with STAR completing the estimation in half the time of TANDEM-STRAIGHT.

3.3 PLATINUM: Excitation signal extraction method

PLATINUM is the excitation signal extraction method from the windowed waveform, spectral envelope and F0 information [22]. In typical vocoder-based systems, the pulse is used as the excitation signal and the signal calculated based on the spectral envelope with minimum phase is used as the impulse response of voiced speech. White noise is used as the excitation signal to synthesize consonants. PLATINUM can calculate the phase information of the windowed waveform and use it when synthesizing.

The observed spectrum $Y(\omega)$ is defined as the product of the spectral envelope $H(\omega)$ and target spectrum $X(\omega)$ for reconstructing the waveform. In the case that the phase of spectral envelope $H(\omega)$ is minimum, an inverse filter is given by simply calculating the inverse of $H(\omega)$. Since minimum phase is used as the phase information of $H(\omega)$

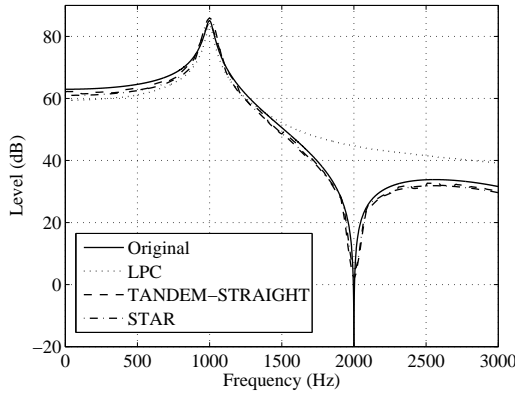


Figure 5. Spectral envelope estimated by STAR. The target spectrum consists of a pole and a dip. Linear predictive coding (LPC) could not estimate the spectral envelope, whereas TANDEM-STRAIGHT and STAR could.

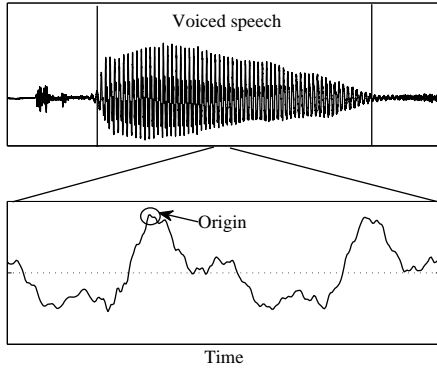


Figure 6. Determination of origin in a voiced speech. The index with maximum amplitude round the center of speech is determined. Other positions are automatically calculated by origin and F0 contour.

in vocoder-based systems, the target spectrum $X(\omega)$ is given by

$$X(\omega) = \frac{Y(\omega)}{H(\omega)}. \quad (2)$$

As shown in Eq. (1), the spectral envelope $H(\omega)$ estimated by STAR is smoothed by a rectangular window. The inverse value of $H(\omega)$ can be calculated without an extremely high amplitude.

The pitch marking required for TD-PSOLA [23] is crucial because PLATINUM uses windowed waveform as the glottal vibration for synthesis. To calculate the temporal positions for calculating the spectrum $Y(\omega)$, PLATINUM uses an origin from voiced speech and an F0 contour. The origin of each voiced speech is determined in the manner shown in Fig. 6. The center interval of the voiced speech is selected, and the time with the maximum amplitude is extracted as the origin for windowing. Other positions are automatically calculated by the F0 contour.

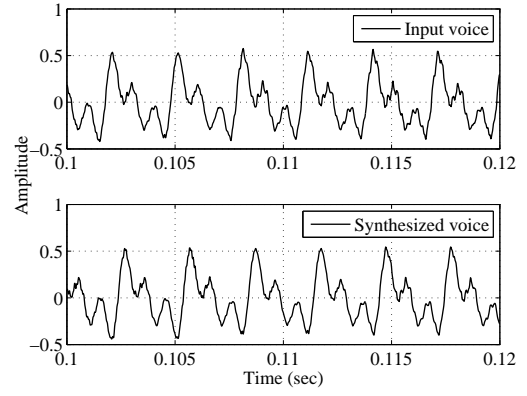


Figure 7. Waveforms of input speech (upper) and synthesized speech (bottom). Since PLATINUM can synthesize a windowed waveform, the output speech is almost all the same except for the temporal position of each excitation

3.4 Sound quality of the synthesized speech

Figure 7 shows the waveforms of both input and synthesized speeches. The waveform synthesized with WORLD is almost completely the same as the input waveform because PLATINUM can compensate for the windowed waveform by the minimum and maximum phase. The temporal positions of each glottal vibration are shifted because the F0 contour does not include the origin of the glottal vibrations.

In reference [22], a MUSHRA-based evaluation [24] was carried out. WORLD was compared with STRAIGHT [18] and TANDEM-STRAIGHT [19] as modern techniques, and Cepstrum [17] as a conventional one. Not only a synthesized speech but also F0-scaled speeches ($F0 \pm 25\%$) and Formant-shifted speeches ($\pm 15\%$) were tested to determine the robustness of the modification. The speeches used for the evaluation were of three males and three females. The sampling was 44,100 Hz/16 bit, and a 32-dB (A weighted) room was used. Five subjects with normal hearing ability participated. This article showed only the result of WORLD, STRAIGHT and TANDEM-STRAIGHT because the sound quality of Cepstrum is clearly low compared with these three. The results are shown in Table 1. Under almost all conditions, WORLD can synthesize the best speech.

4. EVALUATION

WORLD and a singing synthesizer that fulfills the requirement for UTAU were developed and released via a Web site⁴. Both the execute file and the C language source code were released to encourage collaborative creation by developers. Developers could use WORLD and release their synthesizer without any permission from us (they were released under the modified BSD license). An evaluation was performed to determine whether other singing synthesizers were developed and released. The number of contents uploaded on the video-sharing Web site was also

⁴ <http://ml.cs.yamanashi.ac.jp/world/>

	STRAIGHT	TANDEM-STRAIGHT	WORLD
Synthesized speech	88.2	83.2	97.3
F0-scaled speech (+25%)	77.4	72.1	88.4
F0-scaled speech (−25%)	70.1	67.9	79.3
Formant-shifted speech (+15%)	71.4	71.4	73.2
Formant-shifted speech (−15%)	70.1	67.9	68.1

Table 1. The sound quality of speech synthesized with each method.

counted to collect comments on the sound quality of the synthesizers.

4.1 Created singing synthesizers

As of now (April 2013), six synthesizers created by four developers have been released, and two original singing synthesis applications have been created by one developer. More than 70 contents were uploaded by several creators in Nico Nico Douga.

The comments collected from Nico Nico Douga were analyzed to determine the effectiveness of the synthesizers. Almost all comments on the sound quality of the developed synthesizers were positive. It was also suggested that the sensitive atmospheres depended on the synthesizer even if they were synthesized by WORLD-based synthesizers. On the other hand, there are some indications about the compatibility between the synthesizer and the voice library.

4.2 Discussion

Six synthesizers based on WORLD were developed by four developers, and many contents were created and uploaded on the video-sharing Web site. In this section, we discuss our evaluation of the synthesizers.

4.2.1 Synthesizers as content generation software

Vocaloid and UTAU are singing synthesis systems used for supporting creation activities. Although the simplest evaluation of a singing synthesizer is a MOS evaluation of the synthesized singing voice, the content consists of not only the singing but also the music. Post-processing such as adding reverb affects the quality of the music, and compatibility between the synthesizer and the library (including labeling data) also affects the quality. As shown in Fig. 8, there are various factors to evaluate the performance of the synthesizer as content generation software.

4.2.2 Effectiveness of the collaborative creation

The purpose of this study was to support collaborative creation by developers. We consider our attempt a success because six synthesizers were developed and subsequently used to create music. In the past, three synthesizers were released by the developer of UTAU, and three synthesizers that do not use WORLD have been released by other developers. In our case, six synthesizers with WORLD were released. The performance of these synthesizers was verified by other people, which is the collaborative element of the verification.

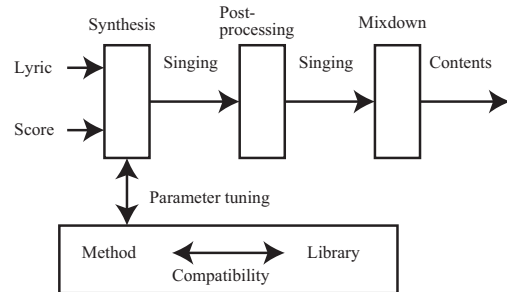


Figure 8. Music creation process. It is difficult to evaluate the synthesized singing voice because the quality of the music does not depend solely on only the singing voice. Not only the synthesizer but also the post-processing and mixdown can change the sensitive atmosphere of the singing voice.

The next step of this attempt is to develop other synthesizers that do not depend on UTAU. Although two such systems have already been developed, they rely on the labeling data and functions of UTAU. Since UTAU requires adjusting the format to synthesize the singing voice, WORLD does not achieve its full potential. Singing voice morphing [25] has potential for use in the field of singing synthesis. More flexible modification will be the primary focus of our future work.

5. CONCLUSIONS

In this article, we described the development of singing synthesizers for UTAU by collaborative creation among many developers. The synthesizers were based on WORLD, which is a high-quality speech synthesis system, and released via a Web site with C language source code. In total, six synthesizers were developed, released, and used to create music.

We also discussed our evaluation of the singing synthesizer. Although WORLD can synthesize a speech that sounds as natural as the input speech, it is difficult to evaluate each synthesizer because there are so many factors in the music creation process.

We consider the proposed attempt to be a success because six synthesizers—a half of all the synthesizers of UTAU—were developed, many creators used them, and their contents were evaluated as good. A discussion of how to evaluate the effectiveness of the singing synthesizer will be the key focus of our future work. We will attempt to develop another singing synthesis system that does not depend on UTAU by collaborative creation.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers 23700221, 24300073, and 24650085.

6. REFERENCES

- [1] H. Kenmochi, "Vocaloid and hatsune miku phenomenon in japan," in *Proc. INTERSINGING2010*, 2010, pp. 1–4.
- [2] T. Nakano and M. Goto, "Vocalistener: A singing-to-singing synthesis system based on iterative parameter estimation," in *Proc. SMC2009*, 2009, pp. 343–348.
- [3] —, "Vocalistener2: A singing synthesis system able to mimic a user's singing in terms of voice timbre changes as well as pitch and dynamics," in *Proc. ICASSP2011*, 2011, pp. 453–456.
- [4] G. Fischer, "Symmetry of ignorance, social creativity, and meta-design," *Knowledge-Based Systems Journal*, vol. 13, no. 7-8, pp. 527–537, 2000.
- [5] M. Hamasaki, H. Takeda, and T. Nishimura, "Network analysis of massively collaborative creation of multimedia contents — case study of hatsune miku videos on nico nico douga—," in *Proc. uxTV2008*, 2008, pp. 165–168.
- [6] K. Yoshii and M. Goto, "Musiccommentator: generating comments synchronized with musical audio signals by a joint probabilistic model of acoustic and textual features," *Lecture Notes in Computer Science, LNCS 5709*, pp. 85–97, 2009.
- [7] H. Dudley, "Remaking speech," *J. Acoust. Soc. Am.*, vol. 11, no. 2, pp. 169–177, 1939.
- [8] A. M. Noll, "Cepstrum pitch determination," *J. Acoust. Soc. Am.*, vol. 41, no. 2, pp. 293–309, 1967.
- [9] L. R. Rabiner, "On the use of autocorrelation analysis for pitch detection," in *IEEE Trans. Acoust. Speech, and Signal Process.*, vol. 25, no. 1, 1977, pp. 24–33.
- [10] M. Morise, H. Kawahara, and H. Katayose, "Fast and reliable f0 estimation method based on the period extraction of vocal fold vibration of singing voice and speech," in *Proc. AES 35th International Conference*, 2009, pp. CD-ROM.
- [11] A. H. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Trans. on Acoust. Speech, and Signal process.*, vol. 29, no. 1, pp. 84–91, 1981.
- [12] H. Kawahara, A. Cheveigne, H. Banno, T. Takahashi, and T. Irino, "Nearly defect-free f0 trajectory extraction for expressive speech modifications based on straight," in *Proc. ICSLP2005*, 2005, pp. 537–540.
- [13] A. Cheveigne and H. Kawahara, "Yin: a fundamental frequency estimator for speech and music," *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1917–1930, 2002.
- [14] A. Camacho and J. Harris, "A sawtooth waveform inspired pitch estimator for speech and music," *J. Acoust. Soc. Am.*, vol. 124, no. 3, pp. 1638–1652, 2008.
- [15] M. Morise, H. Kawahara, and T. Nishiura, "Rapid f0 estimation for high-snr speech based on fundamental component extraction," *IEICE Trans. on Information and Systems*, vol. J93-D, no. 2, pp. 109–117, 2010. (in Japanese).
- [16] B. S. Atal and S. L. Hanauer, "Speech analysis and synthesis by linear prediction of the speech wave," *J. Acoust. Soc. Am.*, vol. 50, no. 2B, pp. 637–655, 1971.
- [17] A. M. Noll, "Short-time spectrum and "cepstrum" techniques for vocal pitch detection," *J. Acoust. Soc. Am.*, vol. 36, no. 2, pp. 269–302, 1964.
- [18] H. Kawahara, I. Masuda-Katsuse, and A. Cheveigne, "Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based f0 extraction," *Speech Communication*, vol. 27, no. 3–4, pp. 187–207, 1999.
- [19] H. Kawahara, M. Morise, T. Takahashi, R. Nisimura, T. Irino, and H. Banno, "Tandem-straight: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, f0, and aperiodicity estimation," in *Proc. ICASSP2008*, 2008, pp. 3933–3936.
- [20] M. Morise, T. Matsubara, K. Nakano, and T. Nishiura, "A rapid spectrum envelope estimatino technique of vowel for high-quality speech synthesis," *IEICE Trans. on Information and Systems*, vol. J94-D, no. 7, pp. 1079–1087, 2011. (in Japanese).
- [21] M. V. Mathews, J. E. Miller, and E. E. David, "Pitch synchronous analysis of voiced sounds," *J. Acoust. Soc. Am.*, vol. 33, no. 2, pp. 179–185, 1961.
- [22] M. Morise, "Platinum: A method to extract excitation signals for voice synthesis system," *Acoust. Soc. & Tech.*, vol. 33, no. 2, pp. 123–125, 2012.
- [23] E. M. C. Hanon and F. Charpentier, "A diphone synthesis system based on time-domain prosodic modifications of speech," in *Proc. ICASSP89*, 1989, pp. 238–241.
- [24] *Method for the subjective assessment of intermediate quality level of coding systems.* ITU-R Recommendation BS.1534-1, 2003.
- [25] M. Morise, M. Onishi, H. Kawahara, and H. Katayose, "v.morish'09: A morphing-based singing design interface for vocal melodies," *Lecture Notes in Computer Science, LNCS 5709*, pp. 185–190, 2009. (in Japanese).

A STUDY OF SPEAKER DEPENDENT FORMANT SPACE VARIATIONS IN KARAOKE SINGING

Mahnoosh Mehrabani

Center for Robust Speech Systems (CRSS)
University of Texas at Dallas, USA
mahmehrabani@utdallas.edu

John H. L. Hansen

Center for Robust Speech Systems (CRSS)
University of Texas at Dallas, USA
john.hansen@utdallas.edu

ABSTRACT

Previous research has shown that trained and professional singers vary their formant frequencies due to vowel modification while singing. In this study we analyze the formant space of vowels for Hindi amateur singers, singing Karaoke. The first and second formants are examined for each speaker, and the deviation of formant space configuration from speaking to singing is compared for different speakers. It is shown that while singing vowel space varies from one speaker to the other, for all speakers the average distance between vowels in F2/F1 plane decreases from speaking to singing. Correlation of F1 and F2 with F0 is also studied for different vowels, and speaker dependent variations of formant tuning is analyzed. In addition to the scientific value, the results of this study can be applied to singer identification, singing skill evaluation, and singing quality assessment.

1. INTRODUCTION

Scientific studies have shown changes in speech production of vowels for professional singers based on articulatory modification while singing, which result in the deviation of singing vowel space from speaking [1–5]. Vocal tract resonances have been studied based on comparing singing formant frequencies with speaking [6–9]. Compared to the number of studies, which focus on the acoustic analysis of singing for trained and professional singers, only a limited number of studies have explored the spectral characteristics of untrained and amateur singing [10–13]. In this study, we analyze speaker dependent differences in the formant space of sung vowels for Hindi Karaoke singing. We specifically focus on the deviation of first two formant frequencies from neutral speaking to singing.

Formant frequencies are defined as peaks of speech spectrum [14], which represent resonating frequencies of the vocal tract. An early speech analysis study [15] showed that the first and second formants correspond to tongue height and frontness, respectively. The same study indicated that vowels can be distinguished based on their first two formant frequencies. A classical study by Sundberg

on the acoustic of singing [1] showed the deviation of formant frequencies, determined by the shape of vocal tract, from speaking to singing for professional opera singers. It was shown that for male opera singers, the two lowest formant frequencies are lower in singing compared to ordinary speech, and spectral amplitude is considerably higher between 2500 and 3000 Hz. This higher amplitude around 3 kHz is due to the clustering of third, fourth, and fifth formants, known as the "singer's formant" [6]. It was also shown [1] that female opera singers vary their first formant frequency according to the pitch of singing.

One of the characteristics of trained or professional singing is vowel modification based on "formant tuning", defined as "approximating one or both of the two lowest vocal tract resonances to harmonics of the glottal source" [16]. Henrich [5] studied the first two vocal tract resonances (R1,R2) of trained opera singers, while singing four different vowels. It was shown that sopranos tune R1 to the fundamental frequency (F0), and R2 to 2F0, and altos tune R1 to F0 or 2F0, while tenors and baritones generally use R1:2F0 and R1:3F0 tunings. Though formant tuning is a singing skill, and intended to increase the audibility, it affects vowel quality. Carlsson and Sundberg [17] showed that formant tuning in synthesized singing vowels, reduces the vowel quality. Another study [18] on the perceptual aspects of singing related timbre, which includes vowel quality to the vocal tract resonance frequencies, or formants.

During the past few decades, a significant number of studies have analyzed sung vowels and compared their acoustic features to spoken vowels. However, most previous studies have examined isolated vowels. Only a limited number of studies have analyzed singing vowels in context [19–21], which mostly focused on trained singers.

This current study examines formant space deviation of sung vowels from spoken for Hindi Karaoke singing, and analyzes the variations between different speakers. It will be shown that despite the speaker dependent differences in the singing vowel space, the average distance between vowels in F2/F1 plane decreases from speaking to singing for all speakers. Based on previous studies on vowel intelligibility [22], this can be interpreted as less distinctive vowels in Karaoke singing compared to speaking. Furthermore, the relation between the two lowest formant frequencies and fundamental frequency is compared for amateur singers based on studying the correlation between formant and F0 vectors for each vowel. In addition to the scientific value, the results of this study can be applied to develop

quantitative spectral measures for singing quality, as well as speaker identification based on singing voice signatures. The singing speech data is described in the next section. Sec. 3 shows the deviation of singing formant space from speaking, and speaker dependent variations of sung vowel space is analyzed in Sec. 4. Sec. 5 studies the relation between F1, F2, and F0, and finally conclusions are drawn in Sec. 6.

2. SINGING DATA

The speech data for singing and speaking vowel space analysis was selected from the UT-Sing corpus. UT-Sing is a multilingual singing database that we collected for the purpose of singing speech analysis, and studying the effects of singing on various speech systems [23, 24]. UT-Sing consists of singing and speaking components in four languages: English, Farsi, Hindi, and Mandarin. Native speakers of each language selected 5 popular songs in their language to sing and read the lyrics. Each song was approximately 3-5 minutes in duration. The speakers' voice was recorded in a soundbooth using a close-talk microphone while singing as well as reading the lyrics of the same songs. Karaoke system prompts were used for singing data collection. While subjects were listening to the music through headphones, the lyrics were displayed, and only the subjects singing voice was recorded (i.e., no music was captured within the audio stream).

UT-Sing subjects were not professional singers, and had a variety of singing skills. Therefore, this database can be used to analyze amateur singing in different languages. With 5 songs per speaker, it is a more realistic and song independent analysis than databases with the same song for all singers. In addition, subjects selected the songs they wanted to sing to be familiar with the songs they were singing and sing more comfortability.

For this study, six Hindi speakers including three males and three females were selected. The total number of different songs sung by these six speakers were 16. There were some overlap between the songs each speaker selected to sing. All singing and speaking utterances for each speaker were phonetically transcribed. Phonemes were manually annotated by a trained transcriber fluent in Hindi. We analyzed the formant space for five Hindi vowels. Table 1 shows the total number of tokens for each vowel. The first row of the table indicates the International Phonetics Alphabet (IPA), and Devanagari symbols. Our formant analysis is based on sustained vowels: 50% of the middle of sung and spoken vowels with duration more than 0.1 sec. Despite the same phonetic context for singing and speaking, the number of sung vowel tokens shown in Table 1 is larger than spoken tokens because of the increased duration for sung vowels. Fig. 1 [25] illustrates the articulatory space of the five Hindi vowels from Table 1.

3. SINGING FORMANT SPACE

Previous studies have shown the effects of vowel modification while singing on the first two formant frequencies for professional singers [3, 5, 16]. Trained singers tune their F1

Vowel	/e:/ (ए)	/a:/ (आ)	/i:/ (ई)	/u:/ (उ)	/o:/ (औ)
Spoken Tokens	1067	1649	293	239	465
Sung Tokens	1868	2196	488	299	673

Table 1. Hindi vowel counts.

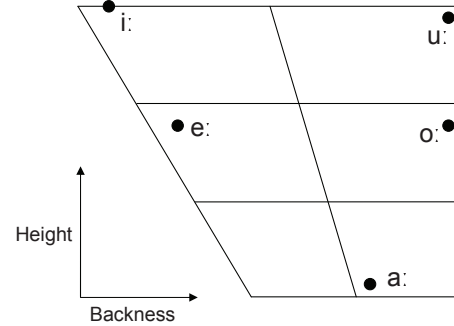


Figure 1. Hindi vowels IPA chart for the five vowels used for formant analysis.

and sometimes F2 to the pitch of singing or its harmonics. This tuning results in the deviation of F2/F1 vowel space of singing compared to speaking. Most previous studies are based on sung isolated vowels. A recent study [19] examined the vowel space of singing for American English speakers with several years of vocal training and singing experience. In this study, we show the variation of F2/F1 vowel plane from speaking to singing for amateur Hindi singers. Wavesurfer speech analysis tool [26] was used for formant estimation. Fig. 2 depicts the F2/F1 vowel plane transition from speaking to singing. The results are averaged for all speakers and all songs.

As shown in Fig. 2, sung vowels are generally closer to each other in the F2/F1 plane compared to spoken. In other words, the formant space occupied by vowels is smaller for singing than speaking. An early study [27] showed that "the linguistic information conveyed by a vowel sound does not depend on the absolute values of its formant frequencies, but on the relationship between the formant frequencies for that vowel and the formant frequencies for other vowels pronounced by that speaker". Therefore, we define the Average Formant Distance (AFD) for our analysis of formant structure of sung vowels:

$$AFD = \frac{1}{N} \sum_{i=1}^N (|F_2' - F_2| + |F_1' - F_1|) \quad (1)$$

where N is the number of all the unique vowel pairs in the analysis vowel set. Mean estimated values of the first two formants are used for AFD calculation. The AFD for Fig. 2 is 989.2, and 587.6 for speaking and singing, respectively. Therefore, there is a 40.6% relative decrease in AFD from speaking to singing. In the next section, formant space of singing vowels is analyzed and compared for each of the three male and three female speakers.

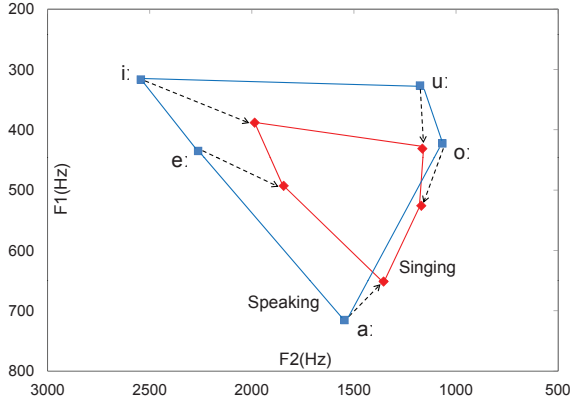


Figure 2. *F2/F1 vowel configuration change from speaking to singing.*

4. SPEAKER DEPENDENT FORMANT ANALYSIS

Formant space deviation from spoken to sung vowels was studied and compared for individual speakers. Fig. 3 and Fig. 4 show the speaker dependent vowel space variation from speaking to singing based on mean of the estimated first and second formant values for male and female speakers, respectively. The AFD changes from speaking to singing was analyzed and compared between speakers. Despite the differences between the song sets for different speakers, the reliability of our analysis is due to averaging over 5 songs per speaker, same song styles for all the speakers (Bollywood Pop songs), and the same text for speaking and singing.

In order to show that variations in Fig. 3 and Fig. 4 are speaker dependent and not song dependent, the F2/F1 space for sung and spoken vowels for two male speakers singing the same song are illustrated in Fig. 5, and for two female speakers singing the same song in Fig. 6. As noted, some overlap exists between songs selected by each speaker. Fig. 5 (a) and (b) correspond to Fig. 3 (a) and (b), respectively (i.e., same speakers). Fig. 6 (a) and (b) correspond to Fig. 4 (a) and (b), respectively. As shown in Fig. 3 and Fig. 4, for all speakers, the average distance between vowels reduces from speaking to singing. For three male speakers the relative AFD decreases are: 17.7%, 17.2%, and 14.4%, while for three female speakers AFD reduces by 88.2%, 49.8%, and 30.6%. On average, for female speakers the AFD reduction is larger with more variability between speakers compared to male speakers. The coefficient of variation (standard deviation divided by mean) for relative AFD reduction from speaking to singing is 0.5 and 0.1 for female and male speakers, respectively. Statistical analysis of the correlation between fundamental frequency and the two lowest formant frequencies for singing, and its relation to the formant space analysis results is presented in the next section.

5. CORRELATION BETWEEN F0 AND FORMANT STRUCTURE

Statistical correlation analysis between singing fundamental frequency and the first two formant frequencies is pre-

sented. As noted, formant tuning in trained singers refers to approximating F1 and/or F2 to harmonics of F0. We examined the relation between the two lowest formants and F0 for untrained amateur singers by calculating the correlation between formant and F0 vectors. F1, F2, and F0 were estimated for each frame (i.e., 10 msec.) for five Hindi vowels sung by three male and three female speakers using wavesurfer. Singing mean F0 of each vowel for six speakers is shown in Table 2. Next, the correlation coefficient for F1 and F2 vectors with F0 vector were calculated for each vowel and each of the three male and three female speakers. The Pearson's correlation coefficient between F1 and F0 is:

$$\rho_{F1, F0} = \frac{E[(F1 - \mu_{F1})(F0 - \mu_{F0})]}{\sigma_{F1} \sigma_{F0}} \quad (2)$$

The correlation coefficient between F2 and F0 is calculated in the same manner. Tables 3 and 4 show correlation coefficients of F0 with F1 and F2, respectively. As shown, F2 of amateur singers does not have significant correlation with F0, while F1 and F0 have correlation coefficients greater than 0.5 and even close to one for some speakers and some vowels. As shown in Table 3, for the first two female speakers with higher mean F0, the first formant of two vowels: /i:/, /u:/ is highly correlated with F0. This is consistent with the previous studies on formant tuning [5]. /i:/ and /u:/ have the lowest F1, which is slightly lower than the singing F0 for these two speakers, and increasing F1 to approximate F0 [1] results in the high correlation between F1 and F0. For the first male speaker there is no significant correlation between F1 and F0 for any vowels. However, for the second male speaker with a similar mean F0, F1 values for /i:/ and /u:/ are highly correlated with F0, and for /e:/ and /o:/ the correlation coefficient is greater than 0.5. For the third male speaker F1 vectors for /e:/, /u:/, and /o:/ are correlated with F0 vector. In addition, for the third female speaker with lower F0 than the other two female speakers, F1 of the vowel /o:/ and F0 are correlated.

The hypothesis is that for the speakers with lower F0, since F1 of all five vowels are higher than the singing F0, they attempt to increase F1 to 2F0. Therefore, there is also some correlation between higher F1 vowels such as /e:/ and /o:/ with F0 for these speakers. To verify this hypothesis, scatter plots of F1 versus F0 was studied for the vowel /i:/ for a female and male speaker with correlation coefficient of 0.9 between F1 and F0. Fig. 7 shows the scatter plots. Next, using linear regression, each scatter plot was approximated by a line. The slope of the linear approximation for the female speaker was 0.8, while that of the male speaker was 1.4. The difference between linear regression slopes confirms the difference between formant tuning in male and female amateur singers. As shown in Table 3 there are also speaker dependent formant tuning differences for male speakers and for female speakers. Comparing Table 3 with Fig. 3 and Fig. 4 shows that for some speakers there is a relation between AFD reduction from speaking to singing and correlation of formants with F0.

As noted, Wavesurfer was applied for formant estimation throughout this study for a statistical analysis of formants. In order to test robustness of Wavesurfer formant

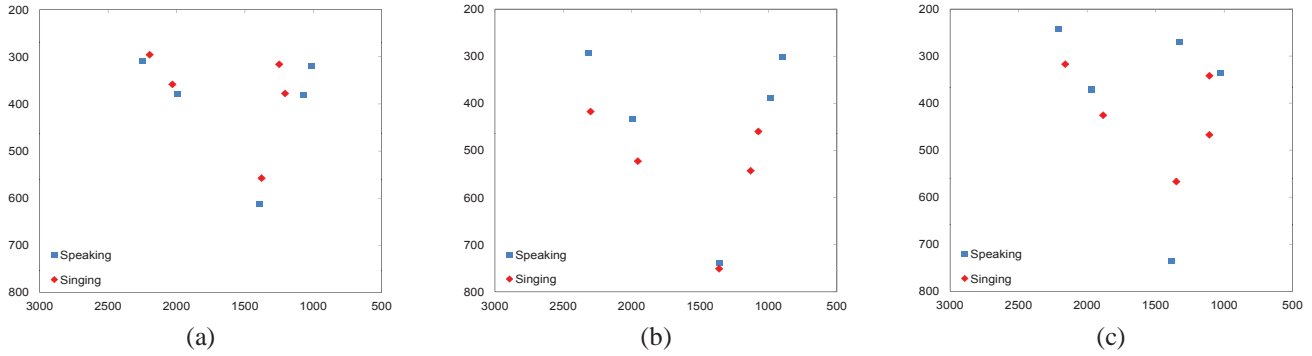


Figure 3. *F2/F1 configuration of sung and spoken vowels for three Hindi male amateur singers.*

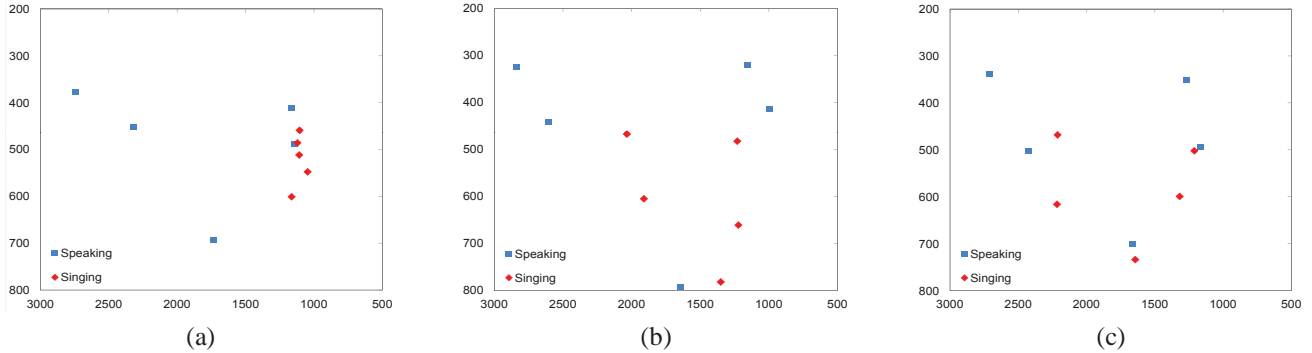


Figure 4. *F2/F1 configuration of sung and spoken vowels for three Hindi female amateur singers.*

Vowel	/e:/	/a:/	/i:/	/u:/	/o:/
Male Speaker 1	224	217.3	221.4	245.6	233.5
Male Speaker 2	220.2	216.4	223.4	238.7	224.6
Male Speaker 3	222.8	224.4	226.5	238.8	243.6
Female Speaker 1	434.1	406.2	450.7	457.6	435.5
Female Speaker 2	427.3	406.4	433.3	437.7	411.7
Female Speaker 3	295.2	301.5	292.9	279.8	283.9

Table 2. *Mean F0 (Hz) of sung vowels.*

Vowel	/e:/	/a:/	/i:/	/u:/	/o:/
Male Speaker 1	0.3	-0.1	0.2	0.4	0.3
Male Speaker 2	0.7	0.2	0.9	0.96	0.6
Male Speaker 3	0.6	0.2	0.4	0.7	0.8
Female Speaker 1	0.4	0.1	0.9	0.7	0.1
Female Speaker 2	-0.01	-0.1	0.9	0.8	0.1
Female Speaker 3	0.4	-0.1	0.3	0.4	0.6

Table 3. *Correlation coefficient of the first formant (F1) of each sung vowel with its fundamental frequency (F0).*

estimation for singing speech, specifically with high pitch, F1 and F2 were estimated based on studying the spectra and were compared to formants estimated by Wavesurfer for one song and one of the female speakers with high F0. Fig. 8 compares average formant frequencies. Wavesurfer estimation error for F1 and F2 was approximately 3% and 9%, respectively. This verifies that statistical analysis results based on Wavesurfer formant estimation are reliable and not due to formant estimation errors.

6. CONCLUSIONS

In this study, we have examined speaker dependent variations for Hindi Karaoke singing. Speaking and singing vowel spaces for three male and three female speakers were compared. It was shown that formant space deviation from spoken to sung vowels varies for different speakers. The

average distance between vowels in F2/F1 plane was calculated for each speaker, and for all speakers the average distance for singing was lower than speaking. Formant tuning differences between speakers were also studied based on the correlation between estimated formants and fundamental frequency for each vowel. It was shown that some speakers don't use any formant tuning, while other speakers tune F1 of some vowels to the harmonics of F0. For male and female speakers with high correlation between F1 with F0, the statistical analysis results were shown to be consistent with previous studies on formant tuning for trained singers.

Speaker dependent differences in singing are partially due to individual singing style, vocal timbre, singing skill, and physiological differences between amateur singers. As UT-Sing database consists of amateur singing with a variety of

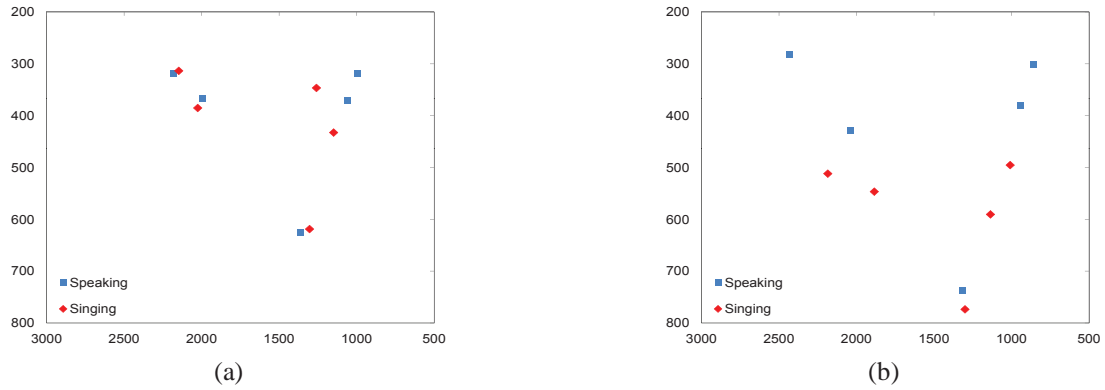


Figure 5. $F2/F1$ configuration of sung and spoken vowels for two Hindi male amateur singers singing the same song.

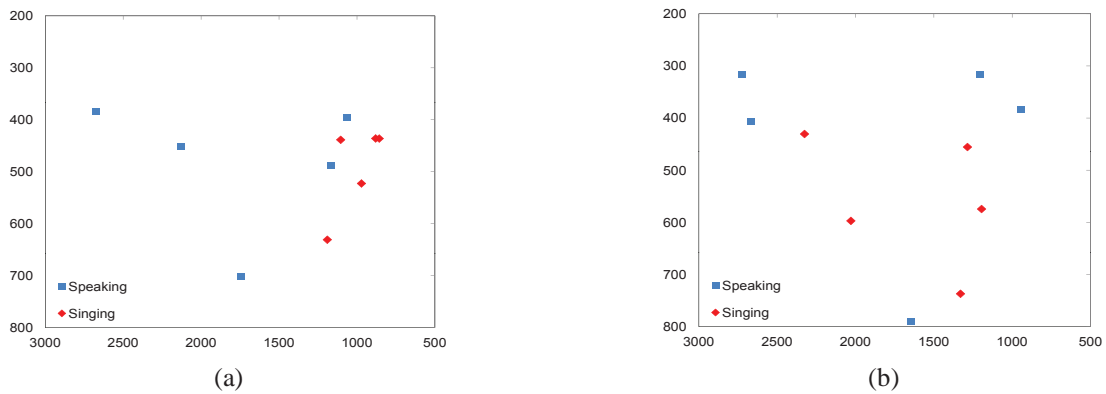


Figure 6. $F2/F1$ configuration of sung and spoken vowels for two Hindi female amateur singers singing the same song.

Vowel	/e:/	/a:/	/i:/	/u:/	/o:/
Male Speaker 1	-0.2	-0.1	-0.1	0.1	0.1
Male Speaker 2	-0.2	0.2	0.03	0.5	0.1
Male Speaker 3	-0.02	0.09	0.1	-0.4	0.5
Female Speaker 1	0.2	0.1	0.3	0.1	0.2
Female Speaker 2	-0.6	-0.2	-0.5	-0.4	-0.05
Female Speaker 3	-0.4	-0.1	-0.4	0.1	0.3

Table 4. Correlation coefficient of the second formant ($F2$) of each sung vowel with its fundamental frequency ($F0$).

skills, future study includes combining formant space variations with formal listener tests for a quantitative acoustic measure for singing quality or skill assessment.

7. REFERENCES

- [1] J. Sundberg, *The acoustics of the singing voice*. Scientific American, 1977.
- [2] G. Bloothoof and R. Plomp, "Spectral analysis of sung vowels. i. variation due to differences between vowels, singers, and modes of singing," *Journal of the Acoustical Society of America*, vol. 75, p. 1259, 1984.
- [3] J. Sundberg, "The science of the singing voice," *Northern Illinois University Press*, 1987.
- [4] B. H. Story, "Vowel acoustics for speaking and singing," *Acta Acustica united with Acustica*, vol. 90, no. 4, pp. 629–640, 2004.
- [5] N. Henrich, J. Smith, and J. Wolfe, "Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones," *Journal of the Acoustical Society of America*, vol. 129, p. 1024, 2011.

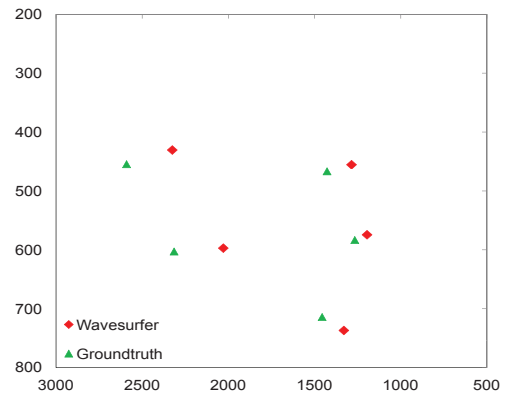


Figure 8. Testing robustness of Wavesurfer formant estimation for singing.

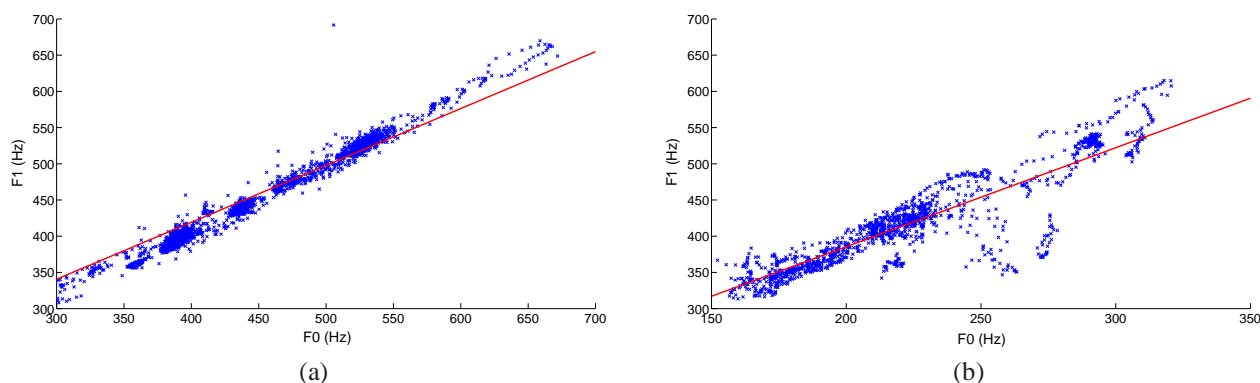


Figure 7. Scatter plots of $F1$ versus $F0$ for the vowel /i:/ for (a) female speaker 1 and (b) male speaker 2.

- [6] J. Sundberg, "Articulatory interpretation of the singing formant," *Journal of the Acoustical Society of America*, vol. 55, p. 838, 1974.
- [7] H. K. Schutte, D. G. Miller, and J. G. Švec, "Measurement of formant frequencies and bandwidths in singing," *Journal of Voice*, vol. 9, no. 3, pp. 290–296, 1995.
- [8] E. Joliveau, J. Smith, and J. Wolfe, "Vocal tract resonances in singing: The soprano voice," *Journal of the Acoustical Society of America*, vol. 116, p. 2434, 2004.
- [9] J. Sundberg, "Formant structure and articulation of spoken and sung vowels," *Folia Phoniatrica et Logopaedica*, vol. 22, no. 1, pp. 28–48, 2009.
- [10] V. M. O. Barrichelo, R. J. Heuer, C. M. Dean, and R. T. Sataloff, "Comparison of singer's formant, speaker's ring, and lta spectrum among classical singers and untrained normal speakers," *Journal of voice*, vol. 15, no. 3, pp. 344–350, 2001.
- [11] K. Omori, A. Kacker, L. M. Carroll, W. D. Riley, and S. M. Blaugrund, "Singing power ratio: Quantitative evaluation of singing voice quality," *Journal of Voice*, vol. 10, no. 3, pp. 228–235, 1996.
- [12] C. Watts, K. Barnes-Burroughs, J. Estis, and D. Blanton, "The singing power ratio as an objective measure of singing voice quality in untrained talented and non-talented singers," *Journal of Voice*, vol. 20, no. 1, pp. 82–88, 2006.
- [13] W. Brown Jr, H. B. Rothman, and C. M. Sapienza, "Perceptual and acoustic study of professionally trained versus untrained voices," *Journal of voice*, vol. 14, no. 3, p. 301, 2000.
- [14] G. Fant, *Acoustic theory of speech production*. Walter de Gruyter, 1970, no. 2.
- [15] G. E. Peterson and H. L. Barney, "Control methods used in a study of the vowels," *The Journal of the Acoustical Society of America*, vol. 24, p. 175, 1952.
- [16] D. Miller and H. K. Schutte, "Formant tuning in a professional baritone," *Journal of Voice*, vol. 4, no. 3, pp. 231–237, 1990.
- [17] G. Carlsson and J. Sundberg, "Formant frequency tuning in singing," *Journal of Voice*, vol. 6, no. 3, pp. 256–260, 1992.
- [18] J. Sundberg, "Perceptual aspects of singing," *Journal of voice*, vol. 8, no. 2, pp. 106–122, 1994.
- [19] E. D. Bradley, "An investigation of the acoustic vowel space of singing," in *Proceedings of 11th International Conference on Music Perception and Cognition*, 2010.
- [20] H. Frostel, A. Arzt, and G. Widmer, "The vowel worm: Real-time mapping and visualisation of sung vowels in music," in *Proceedings of the 8th Sound and Music Computing Conference*, 2011, pp. 214–219.
- [21] T. F. Cleveland, J. Sundberg *et al.*, "Formant frequencies in country singers' speech and singing," *Journal of Voice*, vol. 13, no. 2, pp. 161–167, 1999.
- [22] A. T. Neel, "Vowel space characteristics and vowel identification accuracy," *Journal of Speech, Language and Hearing Research*, vol. 51, no. 3, p. 574, 2008.
- [23] M. Mehrabani and J. H. Hansen, "Language identification for singing," in *IEEE ICASSP*, 2011, pp. 4408–4411.
- [24] M. Mehrabani and J. Hansen, "Singing speaker clustering based on subspace learning in the GMM mean supervector space," *Speech Communications*, vol. 55, no. 5, pp. 653–666, 2013.
- [25] M. Jha and P. Rao, "Assessing vowel quality for singing evaluation," in *Communications (NCC), IEEE National Conference on*, 2012, pp. 1–5.
- [26] K. Sjölander and J. Beskow, "Wavesurfer - an open source speech tool," in *Proceedings of ICSLP*, vol. 4, 2000, pp. 464–467.
- [27] P. Ladefoged and D. E. Broadbent, "Information conveyed by vowels," *Journal of the Acoustical Society of America*, vol. 29, p. 98, 1957.

TEMPORAL COORDINATION IN VOCAL DUET PERFORMANCES OF MUSICAL ROUNDS

Caroline Palmer, Frances Spidle, Erik Koopmans

Department of Psychology
McGill University
1205 Dr. Penfield Ave
Montreal, QC, Canada H3A 1B1
{caroline.palmer, frances.spidle,
erik.koopmans}@mcgill.ca

Peter Schubert

Dept. of Music Research
Schulich School of Music
McGill University
555 Sherbrooke St W.
Montreal, Canada H3A 1E3
erik.koopmans@mcgill.ca

ABSTRACT

How do vocalists coordinate their timing with others in choral ensembles? Performers in many musical ensembles take on musical roles, such as a conductor or leader of the group; these roles may influence their synchronization. Pitch relationships between the singers' musical parts may also influence synchronization, such as the comparison between Unison singing (singing the same pitches) and musical Rounds (singing the same part, delayed in time). We compared the temporal coordination of duet performance in Unison and Round performances of a familiar melody. Vocalists with experience in ensemble performance sang the melody in duet performances; one participant was assigned the role of Leader and one of Follower, in Unison and Round performances. The tone onsets of the Leader preceded those of the Follower by a small but consistent amount. In addition, Rounds were performed at a slower rate than Unison performances, suggesting the increased difficulty of maintaining a different part. Finally, the pattern of correlations across the Leader and Follower's beat durations indicated strong similarity in tempo changes between the singers within each performance. The correlations were reduced in the Rounds relative to Unison performances, but Round performances showed similar tempo changes when the 2-bar delay between parts was taken into account.

A Method of Division of Soprano Ranges and Confirmation of their Voice Transformation Point Based on Harmonics Analysis

Qu Ge

Central Conservatory of Music, Beijing, China

suyuantingge@163.com

Abstract: This study raises a new method to divide the soprano ranges and to find the sound transformation point with the analysis of the fundamental frequency of every sound in the whole range, the numbers of visible harmonics, the intensity of basic sounds, the difference of intensity among the first, second and third harmonic and the types of envelope, etc, besides the judgments of instructor's subjective hearing and the singer's self-sense when the three sopranos singing the vowel /a/.

Key Words: Vocal-range Voice-transformation-point (*abbr.* VTP) Intensity-difference-of 1st-2nd-3rd-harmonics Harmonics-envelope Number-of-visible-harmonics

Introduction

Rubin(1960) et al. found, with the high speed cinematography, that vocal cords vibration falls into three types. Vocal teachers are generally acknowledged that singer's voice range can be divided into three parts: high, middle and low range. Voices in the same range, compared with those in other ranges, have similar timbre, the situation and ways to utter these voices are stable. simultaneously, there are different requirements for the physiological parts and uttering method according to different parts in the whole vocal range. Therefore, singers have to adjust the position of larynx, intensity of air stream, shape of vocal track, opening and shutting of oropharynx, position of voice, resonance, articulation, and so on. And the critical point for adjustment is called voice transformation point by the singing teachers. It is especially important for both teachers and students to divide voice ranges reasonably and confirm voice transformation points correctly because different ranges require different singing methods.

Reasonable voice range division and confirmation of the voice transformation points can lead singers to adjust purposefully according to the special requirements of different voice range, so that the vocal teaching can be conducted correctly. If not, there will be great harm to vocal learners, which will cause the situation that students sing in the middle, or the opposite. Former vocal division in teaching and singing is that the whole range is divided into three parts equally and call them low, middle and high

voice range. Voice transformation point is always considered to be the inflexion which appears unsteadily when students sing the ascending and descending musical scales according to teacher's experience and student's feeling (comfortable or not). An qualified vocal teacher can decide students' voice range division and the voice transformation points accurately in most time according to his acute hearing and rich experience, and it is a basic for vocal teaching smoothly. So, whether is there another way to divide voice range and find the voice transformation points in addition to teacher's subjective judgment on the basis of their professionalism and professional study?

This paper makes analysis of every tone of /a/ in the whole range which is sung by three sopranos with the method of harmonic analysis. And this research will work out an objective new way to find voice ranges and the voice transformation points combining the transforming characteristics of acoustic features such as the number of visible harmonics, intensity of fundamental partial, distribution of harmonic series and its intensity envelope on the basis of data analysis and comparison.

1. Method

Three sopranos are chosen as the singers according to their different ages, time of singing and university. The three singers can sing properly, have no throat disease, are not in their period and follow different vocal teachers. The three sopranos in this paper are called S1, S2 and S3 for short.

Every singer separately utters the vowel /a/ from the lowest voice to the highest voice. This study makes some breathe design and voice repeat to avoid the unfavorable effect caused by insufficient breath. For example, when the musical scales of A major are sung, the first sentence goes from the lowest tone A3

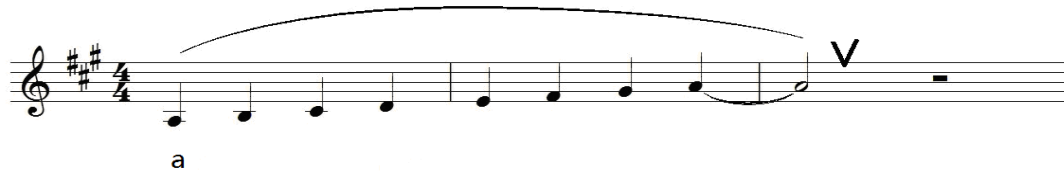
ascending to A4 (music score 1). The second sentence ascends to F#5 from F#4 (music score 2). The tones F#4, G#4 and A4 are repeated. Sentence 3 ascends from E5 to C#6 or D6 (music score 3) and

the tones E5 and F#5 are repeated (music score 1-3). The three singers utter from low voice to high voice in the light of A major (all the three singers think that A3 is their lowest tone. Chart 1).

Chart 1.

Oder	Age	Voice Part	Singing Style	Years of Learning	Selfreport Range	Selfreport VTP	University	Grade
S1	20	soprano	folk	10	A3-D6	E5 or F5	China Conservatory	Sophomore
S2	24	soprano	Bel canto	8	A3-D6	Eb5 or F5	China Conservatory	Graduate student second year
S3	34	soprano	Bel canto	12	A3-C6	Eb5 or E5	Central Conservatory of Music	Teacher in school of arts

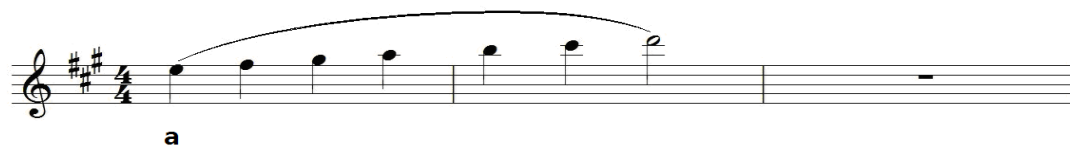
music score 1.



music score 2.



music score 3.



The voice samples are collected in standard studio, the parameters such as reverberation time and uniformity of sound field are in accordance with standard. Corrected German Neumann U87ai condenser microphone is put to use with cardioid direction. Microphone is adjusted to the height of the singer's mouth, the distance between microphone and mouth is 30cm. the singer is told to stand at the assigned position and do not move and move when singing. Singers are required to get familiar with the music before recording and to sing the ascending scale with middle speed and with stable voice and normal intensity. The three singers utter the vowel /a/ in accordance with A major and 50 samples are obtained.

2. Data Analysis

What decides the main vocal features of timbre are the number of its harmonics and its characteristics of intensity distribution. Generally speaking, the less the harmonics, the purer and thinner of the voice; the more the harmonics, the thicker and richer of voice. Compared with other harmonics, the relative intensity of fundamental frequency in the whole harmonics has

important effect on the timbre, for example, whether the fundamental frequency is the strongest, if not, which harmonics exceed it in intensity and go beyond the range, the shape of envelope intensity of the whole harmonic range, etc. This study takes uttering /a/ as the object which is sung from the lowest voice A3 ascending to the highest of A major scale, analyzing comprehensively voices' changes of vocal feature by its frequency, number of visible harmonics, intensity of fundamental sound, intensity difference of 1st, 2nd and 3rd harmonic, types of envelope, the strongest harmonics and its intensity, to decide the critical point or range to provide evidence for voice range division and confirmation of voice transformation point.

2.1. S1's parameter and its analysis

From the analysis of data in chart 2, we can see that there are less visible harmonics, and the number of harmonics show a descending tendency from high to low, the average number is only 3-4 in the 18 test tones. And from that we can see that the timbre of the singer is pure but

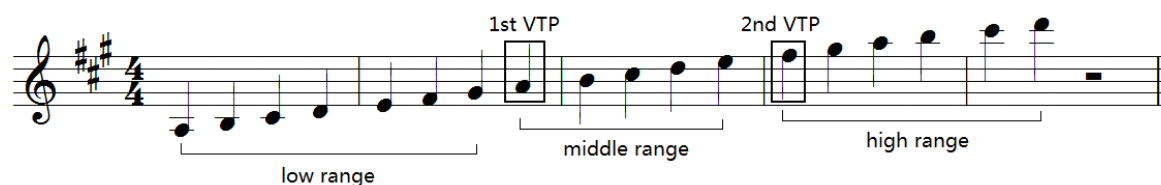
thin, the intensity of the 1st harmonic generally show a phenomenon of becoming stronger with the rising of pitch. From low to high, the intensity envelope types of 1st, 2nd and 3rd harmonic fall into the following: Middle Sunken Type (*abbr.* MST), from A3 to G#4, Middle Bulge Type (*abbr.* MBT), from A4 to E5, Orderly Descending Type (*abbr.* ODT), from F#5 to D6, and their corresponding strongest harmonic is the 1st (A3-G#4), 2nd (A4-E5), 1st (F#5-D6). They show strong consistency in the voice range. According to the above analysis, the author of this paper thinks that the

low range of S1 is A3-G#4, the first voice transformation point is A4. The middle range is A4-E5, the high range is F#5-D6, and the second voice transformation point is F#5. The second voice transformation point is in consistence with the statement of the singers while the first one is in lower range. Singer can succeed in achieving voice range shift without adjusting vocal organ greatly, so that most singers neglect the existing of the first voice transformation point (music score 4).

Chart 2. S1's parameter list of /a/ vowel

Order	Pitch	Frequency (Hz)	Visible harmonics	Intensity (dB)	Intensity difference and envelope type	Strongest harmonics and its intensity(dB)
1	A3-0	219.95	5	27.01	11.02, 1.45/ ODT	1st/27.01
2	B3+20	249.80	5	26.73	8.27, -2.20/ MST	1st /26.73
3	C#4+3	277.79	5	25.58	5.61, -4.37// MST	1st /25.58
4	D4+7	295.02	4	22.92	3.68, -0.83// MST	1st /22.92
5	E4-11	327.62	4	26.03	-0.25, 0.43/MBT	2nd/26.28
6	F#4-9	368.24	3	27.38	3.52,0.86/ ODT	1st /27.38
7	G#4+30	422.60	3	31.23	3.58,4.09/ODT	1st /31.23
8	A4+26	446.83	3	29.62	-0.44, 12.54/ MBT	2nd /30.06
9	B4-23	487.59	3	34.20	-4.78,30.85/ MBT	2nd /38.98
10	C#5-14	550.10	3	33.68	-10.87,34.15/ MBT	2nd /44.55
11	D5-24	579.46	3	37.58	-7.70,42.27/ MBT	2nd /45.28
12	E5-21	651.41	3	39.54	-6.54,41.33/ MBT	2nd /46.07
13	F#5-50	719.24	3	40.70	4.38, 35.97/ ODT	1st /40.70
14	G#5-12	824.92	4	52.66	8.94,21.56 ODT	1st /52.66
15	A5+2	881.24	4	50.85	6.70,13.51/ ODT	1st /50.85
16	B5-21	976.22	3	55.51	7.02,12.93/ ODT	1st /55.51
17	C#6-32	1088.45	2	57.51	9.59/ ODT	1st /57.51
18	D6-26	1157.48	2	58.96	1.02/ ODT	1st /58.96

Music score 4.S1sound division and its sound transformation point



2.2. S2's parameter and analysis

Chart3. S2's parameter list of "a" vowel

Order	Pitch	Frequency (Hz)	Visible harmonics	Intensity (dB)	Intensity difference (dB)and envelope type	Strongest harmonics and its intensity(dB)
1	A3-37	215.35	6	11.94	9.05, -0.84/ MST	1st/11.94
2	B3+16	249.19	5	19.12	11.72, -1.10/ MST	1st /19.12
3	C#4+2	277.55	5	23.51	9.26, -2.43/ MST	1st /23.51
4	D4+20	297.17	7	22.61	6.08, -4.06/ MST	1st /22.61
5	E4+10	331.63	7	19.73	-7.29, 5.45/ MBT	2nd/27.02
6	F#4+31	376.85	7	18.11	-2.98,-9.82/OAT	3rd /30.95
7	G#4+10	417.76	7	25.81	-2.07,4.48/ MBT	2nd /27.88
8	A4+39	450.07	8	30.95	-2.29, 25.33/ MBT	2nd /33.66
9	B4+19	499.59	7	24.97	-5.41,16.86/ MBT	2nd /30.38
10	C#5+30	564.20	6	30.20	-12.41,31.40 /MBT	2nd /42.61
11	D5+37	600.27	7	33.35	-7.64,29.75/ MBT	2nd /40.99
12	E5+16	665.60	5	39.81	6.95,26.76/ ODT	1st /39.81
13	F#5-8	736.95	5	32.86	9.46,22.14/ ODT	1st /32.86

Order	Pitch	Frequency (Hz)	Visible harmonics	Intensity (dB)	Intensity difference (dB) and envelope type	Strongest harmonics and its intensity (dB)
14	G#5-13	824.47	4	32.95	14.07, 13.84 / ODT	1st / 32.95
15	A5-19	870.62	4	35.85	18.77, 10.56 / ODT	1st / 35.85
16	B5-63	952.90	4	41.78	25.19, -0.73 / MST	1st / 41.78
17	C#6-51	1076.74	3	43.79	30.01, -9.01 / MST	1st / 43.79

Comparing to S1, chart 3 shows that S2's visible harmonics increase obviously and the intensity of fundamental harmonic is ascending. The intensity difference of 1st, 2nd and 3rd harmonic and the types of envelope can be generally divided into MST (A3-D4), MBT (E4-D5), ODT (E5-A5) and MST (B5-C#6). The strongest distribution of harmonics in the range falls into the three stages of 1st harmonic (A3-D4), 2nd harmonics (E4-D5), 1st harmonic (E5-C#6). Though the intensity difference of 1st, 2nd and 3rd harmonic and types of envelope can be generally divided into four types and the strongest

distribution of harmonics in the range falls into the three stages, the three parameters have great inner consistency on the musical scale stages. So, for the S2's singing, A3-D4 is low range, D4 is the first voice transformation point; E4-D5 is the middle range, E5-C#6 is the high range and E5 is the second transformation point. The second singer states that her sound transformation point is Eb5 or F5 (actually the second voice transformation point). And the second voice transformation point E5 that this research decides is within the range she stated (music score 5).

Music score 5. S2 sound division and its voice transformation point



2.3. S3's parameter and analysis

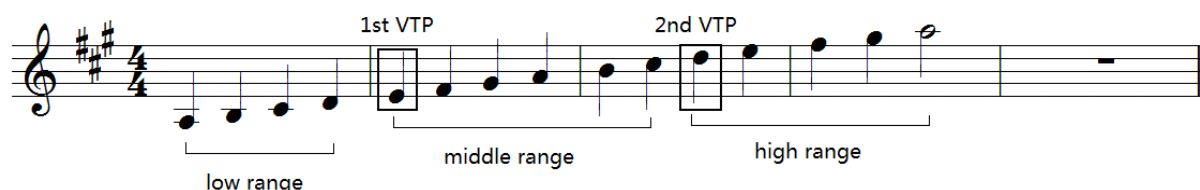
Chart4. S3's parameter list of /a/ vowel

Order	Pitch	Frequency (Hz)	Visible harmonics	Intensity (dB)	Intensity difference (dB) and envelope type	Strongest harmonics and its intensity (dB)
1	A3+12	221.49	5	13.20	2.12, 0.46/ODT	1st / 13.20
2	B3+5	247.64	5	19.72	0.65, 1.84/ ODT	fourth / 22.57
3	C#4+0	277.25	6	22.27	1.65, -4.74/ MST	fourth / 27.18
4	D4-4	293.10	7	23.30	0.48, -4.36/ MST	fourth / 29.91
5	E4-4	328.99	9	24.29	-0.74, -6.18/OAT	3rd / 31.21
6	F#4+4	370.86	8	26.92	-2.38, -9.77/ OAT	3rd / 39.07
7	G#4+19	419.91	6	28.73	-1.89, -2.59/ OAT	3rd / 33.21
8	A4+35	449.14	5	28.18	-2.07, 13.50/MBT	2nd / 30.23
9	B4+25	501.16	6	29.27	-5.71, 24.68/ MBT	2nd / 34.98
10	C#5+40	567.57	6	38.75	-3.83, 32.39/ MBT	2nd / 42.58
11	D5+32	598.34	6	43.30	1.69, 29.26/ ODT	1st / 43.30
12	E5+53	679.86	6	45.37	9.90, 23.80/ ODT	1st / 46.07
13	F#5+42	758.30	5	49.58	22.37, 14.59/ ODT	1st / 49.58
14	G#5+66	862.90	5	50.72	31.08, 4.56/ ODT	1st / 50.72
15	A5+59	910.58	4	50.32	27.66, 6.99/ ODT	1st / 50.32

The parameters in chart 4 show that there are more visible harmonics in S3's voice range and her overtones are rich and thick. The fundamental harmonic become stronger gradually. She has more types of envelope and intensity difference among 1st, 2nd and 3rd harmonic. The four types of envelope all exist in her voice frequency and the unusual type of Orderly Ascending Type (*abbr.* OAT) appears

several times. Her strongest harmonics involve all of 1st, 2nd, 3rd and 4th harmonic. To sum up, we get the result that S3's low range is A3-D4, middle range is E4-#C5, the first voice transformation point is E4. Her high range is D5-A5, the second voice transformation point is D5, which is a bit lower than what she stated: E^b5 or E5 (musical score 6) .

Music score 6.S3's sound division and its voice transformation point



3. Conclusion and Discussion

3.1. Arguments about Whether Voice Ranges and Sound Transformation Points Exist or not

Voice transformation points are the premise of different voice ranges and voice range begins and ends with voice transformation points. In the history of vocal music's development, there are different opinions that voice transformation doesn't exist even that every tone is a point. I also agree with the point of view that every tone needs adjustment of physical and vocal aspects. After acoustic analysis, we found that the three sopranos, with different learning experience, education background, age and singing method, have different timbres when they sing the same musical scale of the same vowel. But after analyzing their voice frequency according to the features and transformation law of parameters in the frequency range, we get to know that three ranges can be divided in their full voice ranges. That is to say, in their whole ranges, there exist three areas

with similar acoustic features, and every tone in the three ranges has tiny difference with specific parameter, but they show the consistency of general acoustic feature and it distinguishes with the consistency in other ranges. Hence, we get the conclusion that different voice ranges really exist. When transforming from low range to middle range or from middle range to high range, singers have to make physical of singing adjustment in large scale for a tone, and the tone is the voice transformation point. And this point is the beginning of next voice range. So, the writer of this paper holds that different sopranos have three ranges of low, middle and high and two voice transformation points exist at the beginning of the middle range and high range (or the ending of the former voice range). The conclusion from analyzing the three singers' voice range and voice transformation point is as follows in chart 5.

Though the first point is not stated by the three sopranos, the second point is in accordance with their self-reported point and the phenomenon is a good note to the objectivity of the new technique.

Chart 5. Statistic table of voice range division and sound transformation points conformation

number	singer	Low range	Middle range	High range	1 st VTP	2 nd VTP	Selfreport VTP
1	S1	A3-G#4	A4-E5	F#5-D6	A4	F#5	E5 or F5
2	S2	A3-D4	E4-D5	E5-C#6	E4	E5	Eb5 or F5
3	S3	A3-D4	E4-C#5	D5-A5	E4	D5	Eb5 or E5

3. 2. Discussion

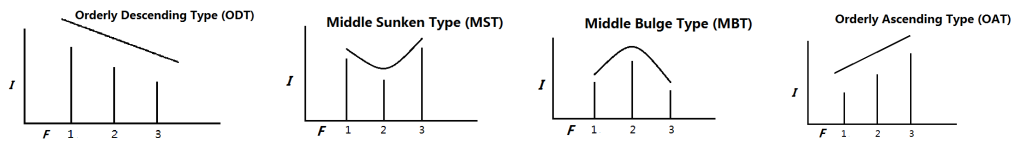
1) Researchers are required to have some new insight of summarization and regularity for the studying object to discover new problems, study new object and summarize new rules. The objective description and definition of the new knowledge need corresponding new parameters. As to the study of singing acoustics, it is certain to have some new and meaningful vocal parameters to participate in the study. Further study should be done to testify the degree of correlation to the object and the rationality of the vocal parameters this paper bring forward such as "number of visible harmonics", "types of envelopes and intensity difference of 1st, 2nd and 3rd harmonics" and so on.

2) This paper simplifies the analysis of data with taking all the tones of vowel /a/ as testing object. Though it has some typical sense, it needs further study whether the result of analyzing other vowel is in accordance with this research or not.

3) Two of the three singers adopted in this research are *bel canto*, and one is *folk style*. The final result shows that S1, the *folk style* has significant difference on the division of voice ranges and confirmation of voice transformation points with S2 and S3 who are *bel canto* (chart 5). What are Reasons causing this situation? Is it only a kind of coincidence or is it a certain result? It needs further exploring.

References:

- [1]. Rubin HJ, Hirt CC. *The falsetto. A high speed cinematographic study*. Laryngoscope. 1960;70:1305–1324.
- [2]. “numbers of visible harmonics” refers to the number of harmonics which are obviously above the abscissa axis under the situation that the strongest harmonics in the list can be showed completely in a window when analyzing the frequency with Gmas2.0. The intensity of visible harmonics is positive numbers.
- [3]. The intensity difference of the first, second and third harmonics refers to the difference between the first harmonics and the second one, and difference between the second and the third. This parameter can briefly reflect the intensity relationship among the three harmonics which greatly affect the quality of voice. It also briefly reflect their intensity envelope.
- [4]. The intensity envelopes of the first, second and third harmonics fall into four types: orderly descending, middle-sunken, middle bulge and orderly ascending (pictures as follows).



DIVERSE RESONANCE TUNING STRATEGIES FOR WOMEN SINGERS

John Smith

Physics, University of
New South Wales, Sydney
john.smith@unsw.edu.au

Joe Wolfe

Physics, University of
New South Wales, Sydney
j.wolfe@unsw.edu.au

Nathalie Henrich

CNRS GIPSA Lab,
Grenoble, France
Nathalie.Henrich@gipsa-
lab.grenoble-inp.fr

Maëva Garnier

CNRS GIPSA Lab,
Grenoble, France
maeva.garnier@gipsa-
lab.grenoble-inp.fr

ABSTRACT

Over a range from 200 to 2000 Hz, the fundamental frequency f_0 of women's singing voices covers the range of the first two resonances ($R1$ and $R2$) of the vocal tract. This allows diverse techniques of resonance tuning. Resonances were measured using broadband excitation at their lips. A commonly noted strategy, used by sopranos, and some altos, is to tune $R1$ close to the fundamental frequency f_0 ($R1:f_0$ tuning) once f_0 approached the value of $R1$ of that vowel in speech. At extremely high pitch, sopranos could no longer increase $R1$ sufficiently and switched from $R1:f_0$ to $R2:f_0$ tuning. At lower pitch many singers of various singing styles found it advantageous to use $R1:2f_0$ tuning. Additionally, many sopranos employed $R2:2f_0$ tuning over some of their range, often simultaneously with $R1:f_0$ tuning.

1. INTRODUCTION

Introductions to phonetic acoustics typically explain how some of the high harmonics of the voice are provided with an acoustic boost by the first two acoustic resonances of the vocal tract, with frequencies $R1$ and $R2$. The resultant formants or maxima in the envelope of the spectrum of the voice have roles in characterizing vowels and some consonants [1]. In singing text, these resonances have important additional functions: because they act as impedance matchers between the glottis and the external radiation field, they enhance the level of sound produced by the voice.

Women's singing voices of different types typically have a fundamental frequency f_0 in the range 160 to 2000 Hz. Singing in the higher part of this range obviously complicates the phonetic role of the tract resonances. However, singers can use either or both of these resonances in strategies to provide high output sound levels with relatively little effort, and perhaps also to assist sound production. This paper looks at some of these strategies.

In normal speech the vibrating vocal folds produce a signal with fundamental frequency f_0 , which is usually unrelated to the particular phoneme being produced.

Copyright: © 2013 Smith et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Different phonemes are associated with different resonance frequencies of the tract. When a harmonic of the voice (an integral multiple of f_0) lies sufficiently close to any one of the R_i , that harmonic is radiated strongly.

Resonance tuning (also known as formant tuning), is the adjustment of the frequency of one or more resonances to match that of one or more harmonics of the voice. Resonance tuning offers singers a technique that can increase loudness with little extra vocal effort [2,3,4]. Furthermore, it has been suggested that the vibration and stability of the vocal folds may be enhanced if they experience an inertive load; i.e. if the resonance frequency is slightly above f_0 [5].

At the low pitch used by most men singers, it is likely that harmonics of f_0 will be reasonably close to $R1$ and/or $R2$, and consequently no widespread resonance tuning strategy is necessary. At the higher pitches used by women, a range of tuning strategies involving both $R1$ and $R2$ become important. These are the subject of the present paper.

2. MATERIALS AND METHODS

2.1 Measuring tract resonances

The measurements were conducted at UNSW in a room treated to reduce reverberation and to reduce external noise. The room has more than 30 dB insulation from the surrounding lab, where sound levels are already low. The average reverberation time over the range used is about 12 ms.

Vocal tract resonances were measured at the lips during singing using broadband excitation at the lips [6,7] – see Figure 1.

At high frequencies, this technique provides much more accurate measurements than those that use the output sound alone; e.g. linear prediction or inverse-filtering. The technique is also less perturbing than approaches that involve external mechanical vibration or that employ various non-periodic phonations. It also avoids the problem of calculating acoustics from geometry that arises if MRI measurements are used.

The technique does, however, have some disadvantages. One is that the vocal gesture needs to be held for a second or so to get good signal to noise ratio (luckily singers are very good at this). It also has the disadvantage that the tract is measured in parallel with

the external radiation field – this means that weak resonances might not be detected.

It is worth mentioning the potential effect of sub-glottal resonances. These are potentially important to the vibration of the vocal folds, because they contribute to the acoustic load on the fold's regeneration mechanism. Could they also affect measurements of the supra-glottal resonances via excitation at the lips? The answer is yes, in the case of respiration, when the vocal folds are widely separated and a new technique [8] allows us to discern them as separate resonances when measuring at the lips. For vocalisation, however, the supra- and sub-glottal tracts are acoustically separated by the inertance of the air between the folds. Its reactance is proportional to frequency, so simple calculations show that, at the frequencies of interest, the glottis impedance effectively seals the tract, which then behaves, to a good approximation, as a closed duct.

There is also a potential slight disadvantage of requiring singers to perform with a device positioned at the lower lip. The flexible mount allows normal jaw movement, but large movements of the body and head are not possible. To date, none of the subjects have reported difficulties.

2.2 The subjects

The data set examined herein is from measurements on 31 volunteer singers; 4 altos and 27 sopranos. Their experience varied from nationally recognized to amateur. Details of the singers are available elsewhere [9,10,11]. Each singer usually sang a sequence of sustained notes in an ascending diatonic scale that ranged from their lowest to highest comfortable pitch.

3. RESULTS AND DISCUSSION

3.1 $R1:f_0$ tuning at high pitch

Sopranos are obvious candidates to benefit from resonance tuning, because $R1$ covers almost all of the standard soprano range (C4-C6; 260-1050 Hz). Thus as f_0 increases and approaches the value of $R1$ for that vowel in speech, a soprano can advantageously increase $R1$ to match f_0 as shown in Figure 2 – the classic tuning of Sundberg [3,4]. This $R1:f_0$ tuning (also known as $F1:H1$ tuning) was used by almost all the sopranos in our

studies once f_0 increased above 400 to 500 Hz. Indeed in the range 500 to 1000 Hz (approx. C5 to C6) this $R1:f_0$ tuning provides the only possible tuning for $R1$. This has an interesting consequence; at high pitch the variation in the values of $R1$ for different singers is greatly reduced in comparison with the values at low pitch – see Figure 2 in [11].

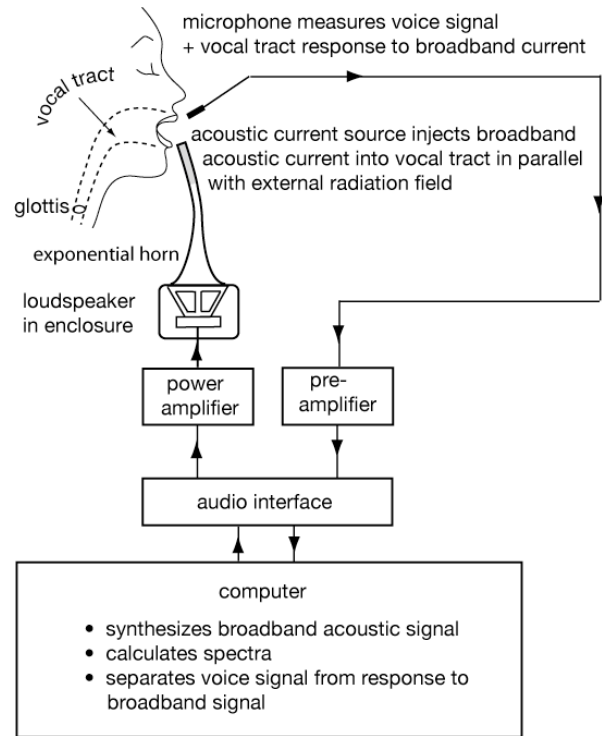


Figure 1. Schematic diagram (not to scale) showing the technique used for real-time measurement of vocal tract resonances. The microphone is normally mounted alongside, and parallel with, the end of the acoustic current source; in this diagram they have been separated for clarity.

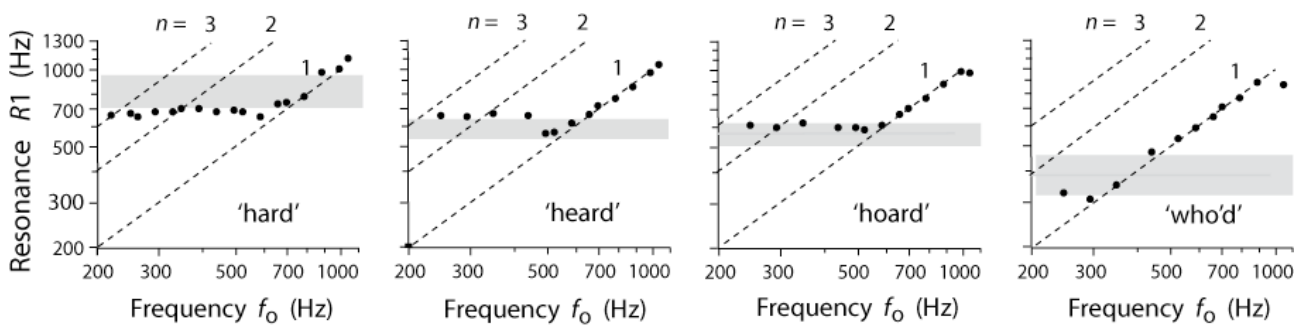


Figure 2. Example of a soprano starting to use $R1:f_0$ tuning when f_0 approaches the value of $R1$ for that vowel in speech. The dashed diagonal lines indicate the relationships $R1 = n f_0$. Shaded areas indicate the mean \pm standard deviation for that vowel in speech measured for sopranos. Data from singer S9 in [9].

3.2 $R2:f_0$ tuning at extremely high pitch

The maximum value of $R1$ in normal speech is typically around 1000 Hz. Some sopranos have learnt to extend this upper limit considerably; Figure 3A shows an extreme example where a soprano has tuned $R1$ to nearly 1500 Hz (around F#6). However, most singers cannot maintain $R1:f_0$ tuning much above 1000 Hz [12]. However because the ranges of $R1$ and $R2$ overlap, it is then possible for a soprano to switch to $R2:f_0$ tuning, and to maintain this for f_0 as high as the upper limit of $R2$ (approx. 2500 Hz or around Eb7) – see Figure 3B.

The use of resonance tuning in the range above 1 kHz may have an importance beyond that of impedance matching the glottis to the radiation field. This is the range of the whistle voice or flageolet register. The mechanism of voice production in this range is not completely understood. Nevertheless, it is possible that a tuned acoustic load could play an important role in determining or stabilising the pitch in this register [11]. It is further possible that learning $R2:f_0$ tuning could be one of the most important steps for a soprano aiming to extend her range to include this register.

3.3 $R1:f_0$ tuning by altos

Altos have a lower value of maximum pitch and so have a smaller range over which $R1:f_0$ tuning would be helpful, particularly for open and mid vowels – see Figure 4. For these vowels it would be possible for sopranos and altos to decrease $R1$ and start $R1:f_0$ tuning at lower values of f_0 , but it appears that singers are generally reluctant to decrease $R1$. Perhaps this is because a decrease in $R1$ would often be achieved by reducing the jaw height, with a consequent decrease in radiation efficiency and hence sound level.

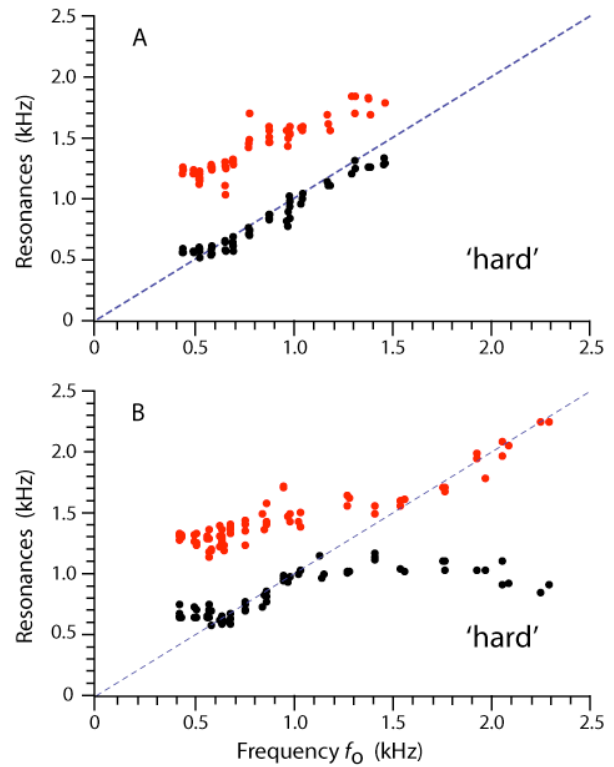


Figure 3. Two examples of a soprano singing the vowel in ‘hard’ at very high pitch above C6. Figure A (top) shows that singer AD4 [11] could maintain $R1:f_0$ tuning as high as 1500 Hz. Figure B (lower) shows that singer NE1 [11] switched from $R1:f_0$ to $R2:f_0$ tuning once $R1$ could no longer be increased. The dashed diagonal line indicates the relationship $R_i = n f_0$. The measured frequencies of $R1$ and $R2$ are indicated by black and red dots respectively.

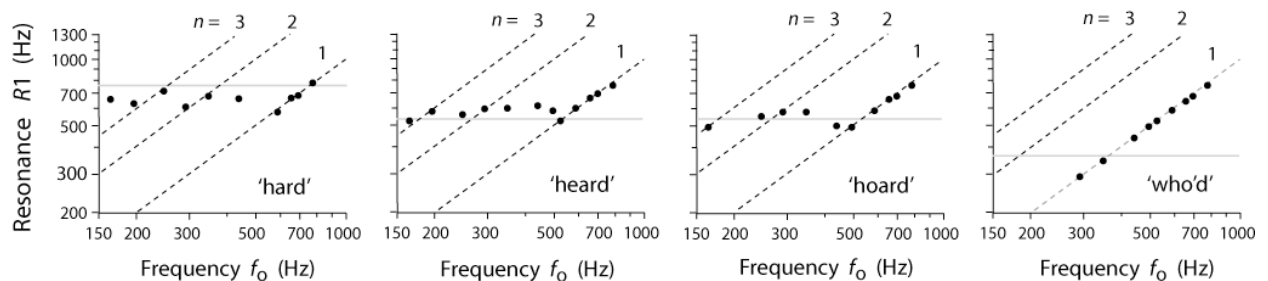


Figure 4. Examples of an alto starting to use $R1:f_0$ tuning when f_0 approaches the value of $R1$ for that vowel in speech. The dashed diagonal lines indicate the relationships $R1 = n f_0$. The horizontal grey lines indicate the value of $R1$ measured for this singer and vowel in speech. Data from alto A1 in [10].

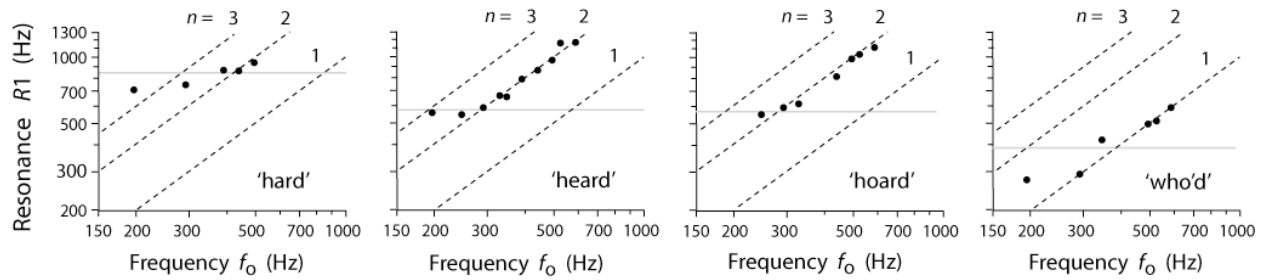


Figure 5. Example of an alto using $R1:2f_0$ tuning for the open and mid vowels once $2f_0$ approaches the value of $R1$ for that vowel in speech. She uses $R1:f_0$ tuning for the closed vowel in 'who'd' where $R1$ in speech is lower. The dashed diagonal lines indicate the relationships $R1 = nf_0$. The horizontal grey lines indicate the value of $R1$ measured for this singer and vowel in speech. Data from alto A2 in [10].

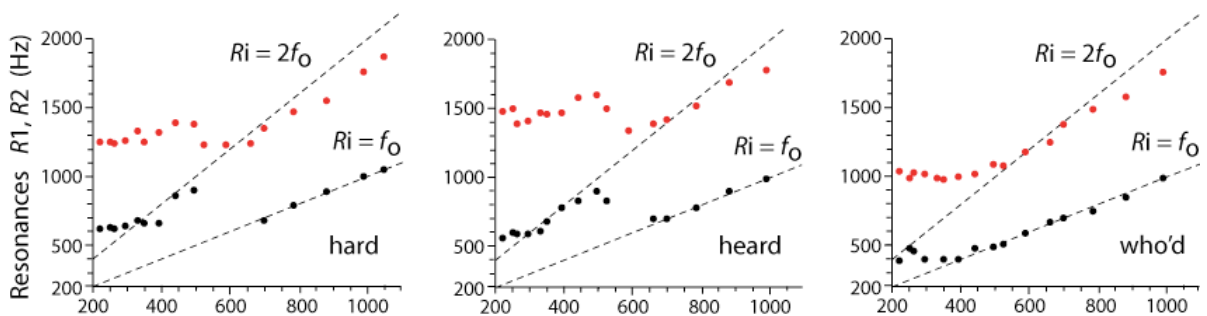


Figure 6. Example of a soprano using $R1:2f_0$ tuning at relatively low pitch. She then switches to simultaneous $R1:f_0$ and $R2:2f_0$ tuning. The dashed diagonal lines indicate the relationships $Ri = nf_0$. The measured frequencies of $R1$ and $R2$ are indicated by black and red dots respectively. Data from singer S7 in [10].

3.4 $R1:2f_0$ tuning by altos and sopranos

An alternative to decreasing $R1$ in order to start using $R1:f_0$ tuning at lower pitches is to use $R1:2f_0$ tuning instead – see Figure 5. This strategy can allow resonance tuning to be employed over a wider range than if only $R1:f_0$ tuning were used.

This $R1:2f_0$ tuning by altos is widely used in the folk music of some cultures. In this frequency range the ear is more sensitive to the second harmonic ($2f_0$) than the fundamental f_0 . Consequently this tuning can produce a very loud sound with an unusual timbre. Both of these features can be heard in a style of Bulgarian women's singing [13] in which $R1:2f_0$ tuning by altos is used. $R1:2f_0$ tuning can also be used in 'belting', a loud theatrical singing style [14].

3.5 $R2:2f_0$ tuning

Figure 6 shows an example where a soprano uses $R1:2f_0$ tuning at low pitch. However once f_0 approaches the value of $R1$ for that vowel in speech she switches to $R1:f_0$ tuning with simultaneous $R2:2f_0$ tuning once f_0 has increased sufficiently.

It was found that many singers exhibited $R2:2f_0$ tuning over at least a part of their range, and this often occurred simultaneously with $R1:f_0$ tuning. This is possible because an increase in $R1$ produced by mouth opening will usually produce an increase in $R2$. A relatively small adjustment could then allow this double tuning.

4. CONCLUSIONS

Diverse strategies of resonance tuning are quite widely used by women singers. Although $R1:f_0$ tuning is the most common, $R2:f_0$ tuning and $R1:2f_0$ tuning can be employed in the upper and lower regions of the singing range. Some singers also use $R2:2f_0$ tuning, often simultaneously with $R1:f_0$ tuning.

Acknowledgments

We thank the Australian Research Council for support. We are also indebted to the singers who volunteered for this study.

5. REFERENCES

- [1] G. Fant, *Acoustic Theory of Speech Production* (Mouton, The Hague), 1970, pp. 15–26.
- [2] B. Coffin, “On hearing, feeling and using the instrumental resonance of the singing voice,” *NATS Bulletin*, 1974, 31, 26–30.
- [3] J. Sundberg, “Formant Technique in a Professional Female Singer,” in *Acustica*, 1975, 32; pp. 89–96.
- [4] J. Sundberg, *The science of the singing voice* (Dekalb, Illinois: Northern Illinois University Press). 1987, pp. 124–129.
- [5] I. R. Titze, “The physics of small-amplitude oscillations of the vocal folds,” *J. Acoust. Soc. Am.*, 1988, 83, pp. 1536–1552.
- [6] J. Epps, J. R. Smith, and J. Wolfe, “A novel instrument to measure acoustic resonances of the vocal tract during speech,” in *Meas. Sci. Technol.*, 1997, 8; pp. 1112–1121.
- [7] A. Dowd, J. R. Smith, and J. Wolfe, “Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time,” in *Lang. Speech*, 1998, 41, 1–20.
- [8] N. Hanna, J. Smith, and J. Wolfe, Low frequency response of the vocal tract: acoustic and mechanical resonances and their losses in *Proceedings of Acoustics 2012 - Fremantle*
- [9] E. Joliveau, J. Smith, and J. Wolfe, “Vocal tract resonances in singing: The soprano voice,” in *J. Acoust. Soc. Am.*, 2004b, 116; pp. 2434–2439.
- [10] N. Henrich, J. Smith, and J. Wolfe, “Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and basses,” in *J. Acoust. Soc. Am.*, 2011, 129; pp. 1024–1035.
- [11] M. Garnier, N. Henrich, J. Smith, and J. Wolfe, “Vocal tract adjustments in the high soprano range,” in *J. Acoust. Soc. Am.*, 2010, 127; pp. 3771–3780.
- [12] E. Joliveau, J. Smith, and J. Wolfe, “Tuning of vocal tract resonances by sopranos,” in *Nature*, 2004a, 427; p. 116.
- [13] N. Henrich, M. Kiek, J. Smith, and J. Wolfe, “Resonance strategies used in Bulgarian women’s singing style; A pilot study,” in *Logoped. Phoniater. Vocol.* 2007, 32; pp. 171–177.
- [14] T. Bourne, and M. Garnier, “Physiological and acoustic characteristics of the female music theater voice,” in *J. Acoust. Soc. America*. 131, 1586–1594.

POWER CONTROL FOR THE SECOND HARMONIC AT HIGH PITCHES IN SOPRANO SINGING: A CASE STUDY

Hironori Takemoto

National Institute of Information
and Communications Technology
takemoto@nict.go.jp

Seiji Adachi

Fraunhofer Institute
Seiji.Adachi@ibp.fraunhofer.
de

Takeshi Saitou

Kanazawa University
t-saitou@ec.t.kanazawa-
u.ac.jp

Kiyoshi Honda

Tianjin University
khonda@sannet.ne.jp

Eri Haneishi

Showa University of Music
haneishi@tosei-showa-
music.ac.jp

Hiroko Kishimoto

Showa University of Music
kishimoto@tosei-showa-
music.ac.jp

ABSTRACT

Sopranos have many singing skills relating to control of the vocal tract shape, especially at high pitches. In the present study, one of the skills called *girare* in Italian opera was examined. This skill is said to be used for producing a mild voice at high pitches. The midsagittal images of the vocal tract of a soprano were scanned by magnetic resonance imaging during singing of the vowel /a/ at G5 (784 Hz) and A5 (880 Hz) with and without *girare*. Acoustic analysis showed that *girare* selectively depressed the power of the second harmonic by more than 10 dB. This result implied that the power reduction could make the auditory impression of singing voice mild. Image analysis indicated that *girare* elongated the vocal tract, widened the oral cavity, and slightly constricted the lips. According to acoustic sensitivity analysis, all these changes tend to decrease the frequency of the second resonance of the vocal tract. Acoustic simulation confirmed that these changes in the vocal tract successfully reduced the power of the second harmonic by more than 10 dB.

1. INTRODUCTION

It is well known that sopranos tune the vocal tract resonances (R1, R2, etc.) to the voice harmonics (f_0 , $2f_0$, etc.) especially at high pitches. The way of tuning was investigated in detail by direct measurements of the vocal tract resonances and the voice harmonics using broadband excitation at the mouth [1, 2]. Garnier *et al.* [1] revealed how sopranos change the relationships among R1, R2, f_0 , and $2f_0$ depending on the pitch. On the other hand, apart from the lips it is still unclear how sopranos control the vocal tract shape to effect changes in the vocal tract resonances, because the vocal tract is an invisible internal space. To overcome this problem, magnetic resonance imaging (MRI) could be a useful tool. If the skill in controlling the vocal tract shape could be visualized, the physiological and acoustic features of the skill would be

revealed.

Recently, we measured the vocal tract shape for sopranos using MRI during singing to examine how sopranos increase R1 at high pitches [3]. In that study, the vocal tract area functions at three different pitches were extracted from MRI data. From these area functions, transfer functions and functions of acoustic sensitivity [4, 5] to R1 were calculated and examined. As a result, it was revealed that sopranos tended to change the area and length of selected parts in the vocal tract which had high sensitivity to R1. In short, sopranos effectively changed the vocal tract shape to increase R1 at high pitches.

In the present study, we focused on one of the skills in soprano singing, called *girare* in Italian opera. According to informal interviews with several Japanese sopranos, this skill contributes to producing a mild voice at high pitches, and it could be similar to the skill called “cover” in English opera and “*decken*” in German opera. Although these sopranos were very interested in how they controlled the vocal tract when singing with *girare*, it is difficult for most of them to sing both with and without *girare*, because they have trained to sing with *girare* for a long time. One of them, however, offered to participate in the acoustic and MRI experiments as a subject, because this soprano could sing with and without *girare* at limited pitches for several seconds. Thus, in this preliminary study, first we analyzed the subject’s singing voice and measured midsagittal images of the vocal tract by MRI during singing with and without *girare*. Then, we examined how the soprano changed the vocal tract when using *girare*. Furthermore, we discussed the acoustic function of this skill based on acoustic sensitivity analysis.

2. MATERIALS AND METHODS

2.1 Subjects and Task

The subject was a lyric soprano who was in her 50s and professional for 27 years. The subject was asked to sustain the pitches of G5 (784 Hz) and A5 (880 Hz) on the vowel /a/ with and without *girare*. To allow natural singing, no restriction was placed on the use of vibrato.

2.2 MRI Experiment and Sound Recording

The midsagittal images of the vocal tract of the subject during singing at the two pitches with and without *girare* were scanned by MRI (Siemens MAGNETOM Verio 3T) installed in the Brain Activity Imaging Center at ATR-Promotions Inc. The subject took a supine position in the MRI gantry and was asked to keep singing approximately for 6 s. During singing, MRI scanning for 5 s was performed. The pixel resolution was 1.0 mm, the slice thickness was 7.0 mm, the echo time was 2.43 ms, and the repetition time was 20 ms. Because the MRI scanner generated loud noise (over 90 dB), the singing voices for the acoustic analysis were recorded by a microphone (Denmark Pro Audio, DPA 4066) and a solid state recorder (Marantz, PMD 671) in an anechoic room just after the MRI experiment. In the recording, the subject took a supine position in the same way as in the MRI experiment.

2.3 Computational Simulation

The area functions of the vocal tract at G5 and A5 for the subject could not be obtained because image distortion occurred around a dental filling in lateral planes. Instead, the area function at G5 for another soprano [3] was used to estimate acoustic effects of changes in the vocal tract caused by *girare*, because the overall shape of the vocal tract for the vowel /a/ might be regarded as common among sopranos. Based on the estimated acoustic effects, physiological and acoustic features of *girare* were simulated by a method proposed by Adachi *et al.* [6] which can modify the vocal tract based on acoustic sensitivity functions.

3. RESULTS AND DISCUSSION

3.1 Acoustic Effects of *girare*

Figures 1 and 2 show long-term spectra at G5 (784 Hz) and A5 (880 Hz) with (red) and without (blue) *girare*. Both figures indicate that the spectrum with *girare* was smaller in power than that without *girare* in the frequency range approximately from 1 to 3 kHz. Especially, the power of 2f0 reduced remarkably when using *girare*: a reduction of approximately 10 dB at G5, and 13 dB at A5. It was possible that this remarkable power reduction of 2f0 made the auditory impression mild. Note that all three sopranos could discriminate singing voices with and without *girare* at both pitches in the informal listening test.

There are possibly two factors for such a remarkable power reduction: changes in the vocal tract resonance and those in the vocal fold vibration pattern. Because non-harmonic components and the overall spectral tilt which reflected the vocal fold vibration pattern were common between the two spectra with and without *girare* at the same pitch, the vocal fold vibration pattern probably did not change much. Thus, changes in the vocal tract resonance around 2f0 could have caused the power reduction.

As R2 existed in the frequency region from 1 to 2 kHz which included 2f0 at G5 and A5, it was reasonable to consider that R2 changes affected the power of 2f0. Ac-

cording to Garnier *et al.* [1], around these two pitches, G5 and A5, sopranos tend to tune R1 to f0 and R2 to 2f0, while individual differences also exist. The tuning prominently enhances the power of both f0 and 2f0. If R2 were not tuned to 2f0, the power of 2f0 would reduce remarkably. Thus, it is hypothesized that *girare* is a skill to intentionally avoid tuning R2 to 2f0 at pitches where R1 and R2 are normally tuned to f0 and 2f0, respectively. At this moment, however, it is unclear whether the soprano increased or decreased R2 for detuning, while keeping R1 at f0.

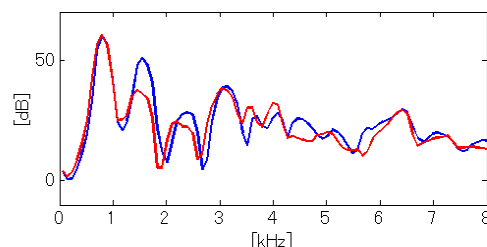


Figure 1. Power spectra for the vowel /a/ sung with (red) and without (blue) *girare* at G5 (784 Hz).

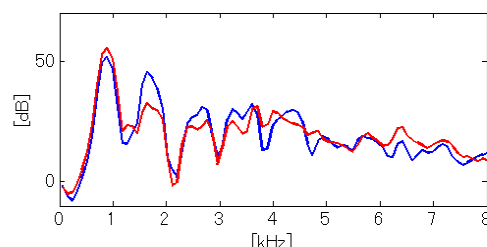


Figure 2. Power spectra for the vowel /a/ sung with (red) and without (blue) *girare* at A5 (880 Hz).

3.2 Changes in the vocal tract shape caused by *girare*

Figures 3 and 4 show outlines of the speech organs extracted from the midsagittal images of the vowel /a/ sung with (red) and without (blue) *girare* at G5 (784 Hz) and A5 (880 Hz). At both pitches, the speech organs were displaced by *girare* in the same way: the tongue, the lower jaw, the hyoid bone, and the larynx were drawn in the postero-inferior direction and the upper and lower lips were slightly protruded and constricted. These displacements would elongate the vocal tract, widen the oral cavity, and constrict the lips. In the following, we will discuss acoustic effects of these changes in the vocal tract shape on the vocal tract resonance.

Figure 5b presents an artificial area function which emulates the vocal tract for the vowel /a/ sung at G5 without *girare*. This area function was obtained by modifying an area function for another soprano [3] to tune R1 to f0 (G5: 784 Hz) and R2 to 2f0 (double frequency of G5: 1568 Hz) using a resonance tuning method proposed by Adachi *et al.* [6]. In this method, small changes in area and length were iteratively applied by referring to area and length sensitivity functions. Note that technically R1 of this area function was tuned to slightly above G5, 800 Hz, because of stable oscillation of the vocal folds.

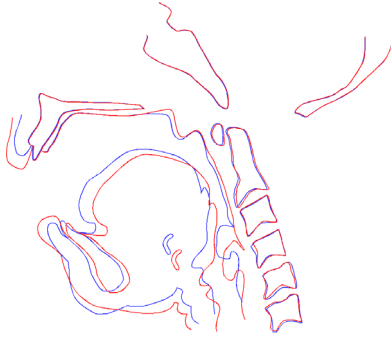


Figure 3. Mid-sagittal outline of speech organs for the vowel /a/ sung with (red) and without (blue) *girare* at G5 (784 Hz).

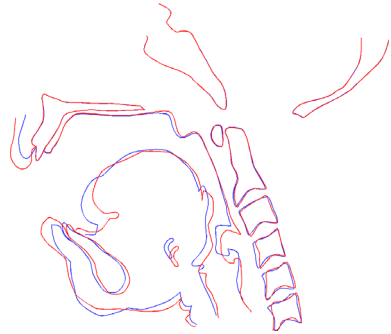


Figure 4. Mid-sagittal outline of speech organs for the vowel /a/ sung with (red) and without (blue) *girare* at A5 (880 Hz).

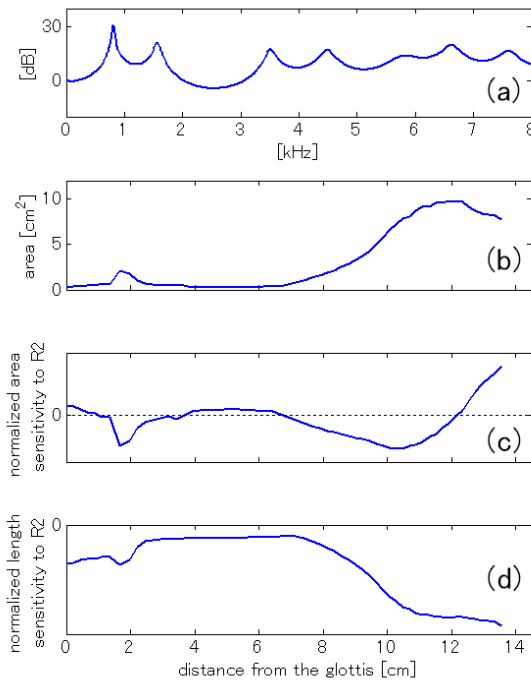


Figure 5. (a) Transfer function calculated from area function (b), (c) normalized area sensitivity function to R2, and (d) normalized length sensitivity function to R2.

Figure 5c presents the area sensitivity function to R2. This function indicates that widening regions with positive value increases R2, while constricting those regions

decreases R2. Conversely, widening regions with negative value decreases R2, while constricting those regions increases R2. The absolute value of the sensitivity function reflects the amount of R2 shift. Thus, widening the oral cavity and constricting the labial end, as caused by *girare*, would most effectively decrease R2. Figure 5d presents the length sensitivity function to R2. Because this function is always negative, elongating any part of the vocal tract decreases R2. Thus, elongating the vocal tract (near the glottis and lips) as caused by *girare* would also decrease R2.

In short, the area and length sensitivity functions indicate that all the estimated changes in the vocal tract caused by *girare* would decrease R2. Consequently, the power of $2f_0$ would remarkably decrease relative to when R2 was tuned to $2f_0$.

3.3 Simulation of *girare*

First, we verified whether the vocal tract modification for decreasing R2 could cause similar changes in the vocal tract shape to those estimated from MRI data. In this simulation, the area function shown in Fig. 5b was used as the original vocal tract shape, i.e., the one without *girare*. As described in the previous section, R1 and R2 of this vocal tract were tuned to 800 Hz and 1568 Hz, respectively. To emulate a vocal tract shape with *girare*, the resonance tuning method of Adachi *et al.* [6] was applied again to the original area function in Fig. 5b, modifying it to maintain R1 at 800 Hz while decreasing R2 from 1568 Hz to 1200 Hz.

Figure 6a presents the original (blue) and modified (red) area functions. As the figure shows, the modification widened the oral cavity, and both constricted and elongated the lip region. These changes agreed with those estimated from the midsagittal images of the vocal tract shown in Figs. 3 and 4. Thus, the vocal tract modification for decreasing R2 successfully simulated the physiological features of *girare*.

Figure 6b shows transfer functions calculated from the original (blue) and modified (red) area functions discussed above. As R2 decreased, the power at 1568 Hz decreased by approximately 20 dB. Thus, if f_0 is set to G5 (784 Hz) and *girare* is used, the power of the second harmonic at $2f_0$ (1568 Hz) would reduce by approximately 20 dB, theoretically.

To examine the power reduction, we simulated vowel production using these area functions and a two-dimensional model of vocal fold vibration [7]. The reason why a two-dimensional model was used for this simulation rather than the two-mass model [8] was that the two-dimensional model can oscillate at high pitches close to R1 where the two-mass model is difficult to oscillate. In this simulation, when the vocal fold resonance frequency was set to 720 Hz and the subglottal pressure was set to 1500 Pa, f_0 became G5 (784 Hz). Other parameters were set to default values of the two-dimensional model.

Figure 7 shows long-term spectra at G5 with (red) and without (blue) *girare*: i.e., the red spectrum was calculated from the modified area function, while the blue spectrum was calculated from the original area function.

Although the spectral tilt is different from that of the soprano voice because the two-dimensional model cannot simulate falsetto register, the power of $2f_0$ for the red spectrum was approximately 12 dB lower than that for the blue spectrum. In short, the vocal tract modification for decreasing R2 suppressed the power of $2f_0$ by 12 dB. This value was close to the observed one described in Sec.3.1. Thus, the vocal tract modification for decreasing R2 quantitatively simulated the acoustic feature of *girare*.

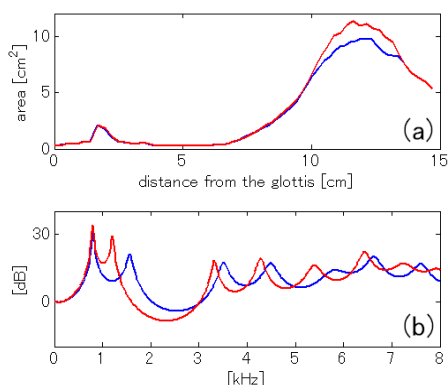


Figure 6. (a) original (blue) and modified (red) area functions, and (b) corresponding transfer functions.

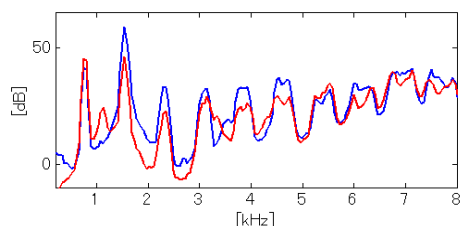


Figure 7. Vowel power spectra calculated by long-term Fourier analysis of vowel waveforms synthesized from the original (blue) and modified (red) area functions shown in Fig. 6 (a).

4. CONCLUSION

In the present study, we examined the mechanism of a soprano skill, called *girare*. The acoustic recording and MRI experiments for a soprano were performed during singing with and without *girare* at G5 (784 Hz) and A5 (880 Hz). At both pitches, the power of $2f_0$ reduced by more than 10 dB when the subject used *girare*. According to midsagittal images of the vocal tract, it was estimated that the vocal tract was elongated, the oral cavity was widened, and the lips were slightly constricted by *girare*. The area and length sensitivity functions indicated that all these changes in the vocal tract contributed to decreasing R2. Then, to examine effects of decreasing R2 on the power of $2f_0$ at G5, the area function whose R1 and R2 were tuned to f_0 and $2f_0$ respectively was modified to decrease only R2 by more than 300 Hz, by the resonance tuning method proposed by Adachi *et al.* [6]. As a result, the power of $2f_0$ in the synthesized vowel reduced by approximately 12 dB. Thus, at least for this subject, it is safe to say that *girare* was a skill which reduced the pow-

er of $2f_0$ by reducing R2 at high pitches where R1 and R2 are normally tuned to f_0 and $2f_0$, respectively.

However, it is unclear whether the skill of this subject revealed in the present study is common among sopranos and can be defined as *girare* in general. Through informal interviews with several sopranos, we found that the definition of *girare* was ambiguous and varied among them; this is understandable, as singing skills including *girare* are basically invisible. Thus, although the physiological mechanism and acoustic function of *girare* for the subject were revealed in the present study, because this is just a case study, more subjects are necessary for generalization. For that purpose, the direct measurement of the vocal tract resonances and voice harmonics [1, 2] could be more convenient than MRI experiments.

Acknowledgments

This study was partly supported by JSPS KAKENHI Grant Numbers 24500233 and 22520156. The authors thank Dr. Parham Mokhtari for helpful comments in revising the manuscript.

5. REFERENCES

- [1] M. Garnier, N. Henrich, J. Smith, and J. Wolfe, "Vocal tract adjustments in the high soprano range," *J. Acoust. Soc. Am.* 127, 2010, pp. 3771-3780.
- [2] E. Joliveau, J. Smith, and J. Wolfe, "Vocal tract resonances in singing: The soprano voice," *J. Acoust. Soc. Am.* 116, 2004a, pp. 2434-2439.
- [3] H. Takemoto, K. Honda, Y. Tanabe, H. Kishimoto, and T. Kitamura, "Changes in the vocal tract shape of sopranos at high pitch," *J. Acoust. Soc. Am.* 131, 2012, p.3304.
- [4] G. Fant and S. Pauli, "Spatial characteristics of vocal tract resonance modes," *Proc. Speech Comm. Sem.* 74, 1974, pp.121-132.
- [5] G. Fant, "Vocal-tract area and length perturbations," Quarterly Progress and Status Report No. 4, *Speech Transmission Laboratory*, Kungliga Tekniska Högskolan, Stockholm, 1975, pp.1-14.
- [6] S. Adachi, H. Takemoto, T. Kitamura, P. Mokhtari, and K. Honda, "Vocal tract length perturbation and its application to male-female vocal tract shape conversion," *J. Acoust. Soc. Am.* 121, 2007, pp. 3874-3885.
- [7] S. Adachi, and J. Yu, "Two-dimensional model of vocal fold vibration for sound synthesis of voice and soprano singing," *J. Acoust. Soc. Am.* 117, 2005, pp. 3213-3224.
- [8] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell Syst. Tech. J.* 51, 1972, pp.1233-1268.

GENERATING SINGING VOICE EXPRESSION CONTOURS BASED ON UNIT SELECTION

Martí Umbert

Jordi Bonada

Merlijn Blaauw

Music Technology Group, Universitat Pompeu Fabra
Barcelona, Spain

{marti.umbert, jordi.bonada, merlijn.blaauw}@upf.edu

ABSTRACT

A common problem of many current singing voice synthesizers is that obtaining a natural-sounding and expressive performance requires a lot of manual user input. This thus becomes a time-consuming and difficult task. In this paper we introduce a unit selection-based approach for the generation of expression parameters that control the synthesizer. Given the notes of a target score, the system is able to automatically generate pitch and dynamics contours. These are derived from a database of singer recordings containing expressive excerpts. In our experiments the database contained a small set of songs belonging to a single singer and style. The basic length of units is set to three consecutive notes or silences, representing a local expression context. To generate the contours, first an optimal sequence of overlapping units is selected according to a minimum cost criteria. Then, these are time scaled and pitch shifted to match the target score. Finally, the overlapping, transformed units are crossfaded to produce the output contours. In the transformation process, special care is taken with respect to the attacks and releases of notes. A parametric model of vibratos is used to allow transformation without affecting vibrato properties such as rate, depth or underlying baseline pitch. The results of a perceptual evaluation show that the proposed approach is comparable to parameters that are manually tuned by expert users and outperforms a baseline system based on heuristic rules.

1. INTRODUCTION

Modeling expressive speech and singing voice has attracted the interest of researchers for many years now. One of the main problems with many of the current singing voice synthesizers that are widely available, such as Vocaloid [1], is that the included models of expression are relatively simple and often are not sufficient for providing a natural-sounding and expressive synthesis “out-of-the-box”. To improve results, users have to manually edit control parameters such as pitch bend, dynamics and vibrato, making the process very time-consuming and requiring expert skills.

To alleviate this problem, this paper aims to automatically generate better expression contours from a high-level,

symbolic input score. The scope of this article is limited to pitch and dynamics (loudness or singing intensity) evolution over time within a local context. Other aspects of musical expression, such as phrasing, timing deviations, timbral variations or ornaments, are not considered.

One of the most basic approaches, and the one that is typically used in current singing voice synthesis systems, is based on heuristic rules. For instance, in [2] a simple parametric model is used, based on anchor points, which are manually derived from a small set of arpeggio recordings. The advantage of these approaches is that they are relatively straight-forward and completely deterministic. On the other hand, the main drawback is that either the models are based on very few observations that don’t fully represent a given style, or they are more accurate but become unwieldy due to the complexity of the rules.

Copy synthesis is another basic type of approach which avoids the need for modeling. In this case, expressive parameters are directly taken from parallel recordings and used to control the synthesis. For instance, [3] applies this concept to singing voice synthesis, where a singing performance directly controls the synthesized expression. The same approach is also applied for generating prosody in speech synthesis [4]. Timbre is set by an emotional database of diphones, and pitch and dihone durations are obtained by copying them from a real utterance. For the Vocaloid synthesis engine, a similar system has been released commercially [5], which has resulted in what is generally considered a very significant increase in synthesis quality. The main disadvantages of copy synthesis approaches are the need for parallel recordings to capture expression controls.

Statistical models, such as HMMs [6], avoid the main drawbacks of the approaches mentioned above. These have been used to model and to produce emotional speech and different speaking styles. Some related applications are style adaptation (from neutral to a target style), interpolation between two or more styles, intensity control and style identification [7]. A well-known project where these techniques have been implemented is HTS [8].

With respect to singing voice synthesis, modeling singing style using HMMs has also been studied. In [9] the singing style is modeled statistically. It focuses on relative pitch, vibrato and dynamics using context dependent HMMs. In this work, notes are considered to contain up to 3 regions (beginning, sustained and end) in order to reflect their expression. HMMs are used to model the singing expression

parameters of these regions or behaviors.

The main drawback of HMM-based synthesis is the over-smoothing of parameters, since statistical averaging affects quality and the perceived emotion [6]. On the other hand, it has proven to be flexible to change voice characteristics, speaking styles and emotions (e.g. interpolation, extrapolation). Another advantage of statistical models is that they are like “black boxes”, simply trained with data, without requiring much prior knowledge about the underlying expressive mechanisms.

Finally, an approach inspired by copy synthesis, but avoiding its main disadvantages, is unit selection. It has been used for synthesizing expressive speech utterances. For instance in [10] unit selection is applied to concatenate variable-size units taken from a large database of emotional speech of a single speaker. In this case the units will contain the audio segments used in synthesis, including inherent prosody. In [11], unit selection is used for the transformation of prosody in intra-speaker emotional voice conversion. First, a mapping is made between parallel neutral and emotional pitch contours, segmented into units. In this case the units used are “accent groups”; an accented syllable and its surrounding syllables. Then, for a given neutral input unit, the closest unit in the map is selected and the mapped emotional prosody is applied in its place.

The main drawback of concatenative-based synthesis is the lack of flexibility with respect to statistical approaches (e.g. interpolation). On the other hand, due to the lack of statistical modeling in unit selection, its best-case examples are generally considered to outperform the HMM-based ones [6].

This paper applies the unit selection approach to expression contour generation for singing voice synthesis. The presented approach is concatenative, taking units from a database of a capella singing voice recordings. Thus, the idea is to capture the original expression from short excerpts of this set of recorded songs by keeping the fine details of the control parameters.

The remainder of this paper is organized as follows. Section 2 explains in detail the proposed system, describing the expression database and the process for the generation of the expression contours. In section 3, the evaluation setup is explained and the results are discussed. In section 4 the conclusions are presented.

2. PROPOSED SYSTEM

2.1 Overview

The system is related to the user’s input and the synthesis engine as shown in figure 1. Given a target score as input, represented in terms of the notes and timing provided by the user, the system (within the gray area) generates expression contours to control the synthesizer. The expression database stores a set of processed and labeled singing voice recordings. The local contexts of this database are the basic elements or units of the presented approach.

Units are melodic contexts of three notes or silences, with their associated dynamics and pitch contours. Choosing units of such length allows to capture a note attack and re-

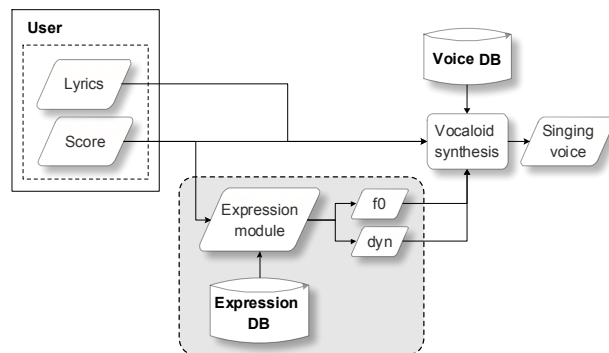


Figure 1. System interaction with the synthesis engine.

lease contours with the previous and following notes contexts. Depending on the target song, a different sequence of units may be retrieved to be transformed and overlapped to generate the expression contours. At synthesis, the engine uses these expression contours to control the samples from the voice database.

The following subsections explain more in detail how the database is created, the optimal sequence of units is selected, the units are transformed (to match target pitch and duration) and concatenated.

2.2 Expression database creation

2.2.1 Recordings

The ideal expression database would need to cover the complete sonic space in terms of possible note durations and pitch intervals. If this was the case, units would require no (or little) transformation to match the target score. Thus, in the proposed approach, just a few songs were recorded and labeled to get an initial idea of the system’s performance. Details on the recorded database are given in section 3.1.

The generated expression contours must control the synthesizer for any target lyrics. Therefore, it is important that the expression contours in the database do not contain unvoiced sections or micro-prosody due to phonetics. That is to say, in order to avoid fluctuations in the recorded pitch and dynamics due to phonetics (not attributable to expression), the lyrics of the recorded songs were modified to only vowels. Thus, any sequence of notes was sung as /ua-i-a-i/, where /a/ and /i/ vowels are alternated at every note change and /ua/ diphthong is used to attack a note from a silence.

2.2.2 Labeling

In the current approach, the recorded songs were labeled in a semiautomatic procedure. The information needed to represent units are the song pitch and dynamics contours, note values and timing as well as vibrato parameters.

Pitch is estimated based on the spectral amplitude correlation (SAC) algorithm described in [12]. In terms of dynamics, the extracted energy sample values are normalized and smoothed using a sliding window of 0.5 seconds. This is to keep the tendency of dynamics instead of the energy at frame level.

The segmentation of the recordings provided the note pitch and timing information. Since recordings were done with the modified lyrics, this task is easier than by score following or detecting pitch changes. Given that notes and vowel changes are strictly related, note segmentation is equivalent to vowel change detection. GMM models were trained for clustering and regression. The data used for training were MFCCs extracted from sustained vowel recordings. The outcome of the segmentation was manually checked.

Note to note transition times are needed to preserve note transition shape during transformation. These are estimated as the time instants when pitch deviates a threshold from the labeled note pitch. The threshold is set to 10% of interval (with a minimum of a quarter semitone).

The vibrato parameters allow resynthesis keeping the shape of the original vibrato at any note pitch and duration. The extracted parameters are depth, rate, baseline pitch and reconstruction error. The estimation of these parameters is semiautomatic, where the first step is to manually indicate the first and last peak or valley for each vibrato. Although the way these parameters are estimated is out of the scope of this paper, their relationship to the reconstructed pitch contour with vibrato $\tilde{F}0(n)$ is:

$$\tilde{F}0(n) = \bar{F}0(n) + d(n)\sin(\varphi(n) + \varphi_{sign}) \quad (1)$$

$$\varphi(n) = \sum_{k=0}^{n-1} 2\pi r(k)\Delta_t + \varphi_{correc}(n) \quad (2)$$

where, in equation 1, $\bar{F}0(n)$ is the estimated baseline pitch (no vibrato) at frame n , φ_{sign} is a constant value that indicates whether the sinusoid's initial phase is 0 or π , $d(n)$ is the pitch deviation (depth) with respect to the baseline, and $\varphi(n)$ is the sinusoid phase. In equation 2, $r(k)$ is the vibrato rate at frame k , Δ_t is the frame shift time and $\varphi_{correc}(n)$ is the reconstruction error.

In figure 2, we show an example of vibrato parameters extraction and resynthesis. The top most subfigure represents the original pitch, its resynthesis and the baseline estimated parameters are plot. In the other three subfigures, depth, rate and reconstruction error are shown respectively.

2.3 Unit selection

Unit selection aims to retrieve short melodic contexts from the expression database that, ideally, match the target contexts or units. Since perfect matches are unlikely, this step retrieves the optimal sequence of units according to a cost function.

The Viterbi algorithm is used to select the set of units that minimize the cost value. This value is the combination of different cost measures. The aim of these costs is that the sequence of units is the least transformed as possible, with units easy to overlap. These costs also consider the introduction of score variations as well as the selection of consecutive units from the database.

2.3.1 Transformation cost

The transformation cost measures how much a source unit u_i has to be modified to match a target unit t_i . It can be

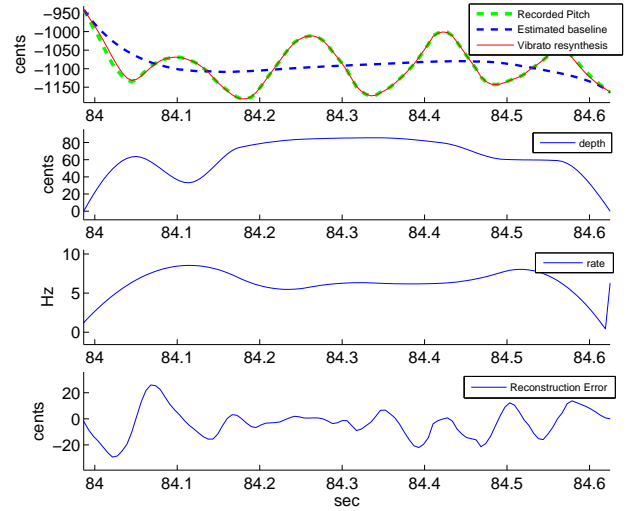


Figure 2. Vibrato resynthesis and parameters: depth, rate, reconstruction error and baseline pitch.

expressed in terms of the mean of two sub-cost functions (amount of pitch shift ps and time stretch ts) as in equation 3:

$$C^t(t_i, u_i) = \frac{1}{2} (C_{ts}^t(t_i, u_i) + C_{ps}^t(t_i, u_i)) \quad (3)$$

Both subcosts functions are a weighted sum of note durations dur ratios (or unit interval pitch values int) between source and target units. The C_{ts}^t cost computation is shown in equation 4:

$$C_{ts}^t(t_i, u_i) = \sum_{n=1}^3 \omega_{ts}(n) \log_2 \left(\frac{dur(u_i(n))}{dur(t_i(n))} \right) \quad (4)$$

where n is the note index within the unit, and time stretch weights ω_{ts} give more relevance to the central unit note transformation:

$$\omega_{ts} = [0.75, 1.5, 0.75] \quad (5)$$

The C_{ps}^t cost computation is shown in equation 6:

$$C_{ps}^t(t_i, u_i) = \sum_{n=1}^2 \omega_{ps}(n) \log_2 \left(\frac{int(u_i(n))}{int(t_i(n))} \right) \quad (6)$$

where n points to the two pitch intervals, and pitch shift weights ω_{ps} give the same importance to both intervals,

$$\omega_{ps} = [1, 1] \quad (7)$$

Besides, an extra rule is applied to avoid selecting some source units. We have assumed that an ascending interval should not be used to generate a descending interval (and vice-versa). Also, silences must be present in the same note in the source and target units, otherwise that unit should not be selected. If this requirements are not met, the transformation cost is set to infinity.

2.3.2 Concatenation cost

The concatenation cost measures how appropriate two units are for overlapping. Consecutive units in the selected sequence share two notes, and crossfading has to be applied to obtain smooth transitions. If the source units are also consecutive in the expression db, the cost is zero. Otherwise, the cost is measured based on the transition to the central unit notes. For this computation, transitions start and end times are used.

2.3.3 Alternative score cost

It measures the possibility to introduce some variations to the original score by erasing or adding silences. Using these variations may offer the possibility to select a sequence of units at a lower cost.

In case of erasing a silence, its length determines the cost. On the other hand, the added silences between notes are very short and not actually synthesized due to synthesizer constraints. Therefore, this variation is not penalized.

2.3.4 Continuity cost

With the three costs used up to this point, it is likely that units are selected from very different songs and contexts. However, the more different the contexts are, the higher impact it has on the resulting contour. At a very local context, this is managed by the concatenation cost, although it only takes into account whether two candidate units are consecutive in the database or not. A higher scope of concatenation is managed by the continuity cost, towards the musical concept of phrasing.

Continuity cost is included to favor the selection of a certain amount L of consecutive source units. Thus, more similar contexts and easy to concatenate (already done by the original singer) can be selected. The starting point is set to a silence or from a point where two units are not consecutive in the database. While L consecutive units are not chosen, selecting non-consecutive units is penalized. When L is reached, a new starting point is set.

2.4 Unit transformation

This step deals with the transformation of the selected sequence of units. Source notes have to match target notes in pitch and duration. Source unit dynamics contour is also stretched according to the target unit duration.

Figure 3 shows the basic idea for the expression contours generation. A target sequence of four notes (bottom image), can be generated by overlapping a couple of source units (A and B) which share two notes. The target pitch contour (pink dashed line) is generated by transforming them in time (according to the target note durations) and frequency (target note pitches). After transformation, crossfading is applied between the pitch contours. Vibratos appearing in the source units are also rendered, preserving the original depth and rate and spanning over the target note duration.

Unit pitch contour is transformed by adding an offset value per note. This offset is the difference between target and source notes. Offset values during note transitions are interpolated linearly in order to have smooth changes.

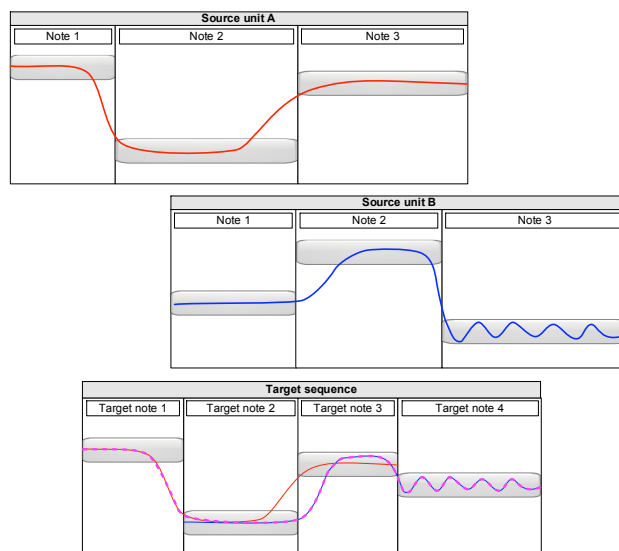


Figure 3. Two overlapping source units (top plots), transformed contours (bottom, solid lines) and concatenated contours (bottom, dashed line)

Unit contours (pitch, dynamics and vibrato parameters) are time stretched in a non-linear way. Most of the transformation is applied during the note sustain in order to preserve note transition shapes. The ratio between source and target note durations determines the amount of transformation.

2.5 Unit concatenation

The expression contours were finally rendered in three steps. First, the basic contours are generated, then the baseline pitch is tuned and finally vibratos are rendered.

2.5.1 Basic contours

The overlapping step of the transformed pitch, dynamics and vibrato parameter contours was handled with a cross-fading mask. This mask was computed per unit in order to determine the samples that contributed to the output contour. More relevance is given to the attack to the central unit note and its sustain, until next unit central note attack time based on note transition start and end times.

2.5.2 Baseline pitch tuning

In order to ensure that sustains were at the right target pitch, the baseline pitch was tuned. A similar process to auto-tuning techniques was followed before rendering the final pitch contour.

This step consists on adding a correction offset to each pitch frame value. First, a sliding window is used to compute local pitch average values through each note duration. The deviation of each frame average value with respect to the target note pitch is weighted in order to get the correction offset. Given the shape of the applied weights (tukey window), boundary note frames are less modified than middle note frames.

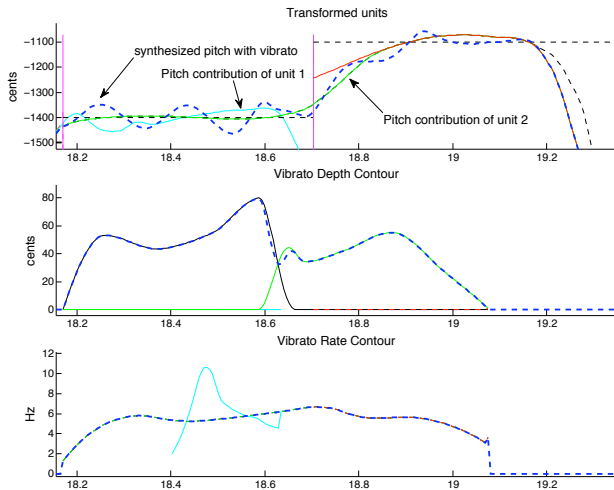


Figure 4. Transformed unit pitches and vibrato control contours concatenation.

2.5.3 Vibrato synthesis

Vibratos were synthesized using the depth, rate and reconstruction error contours generated for the target song. Those frames with depth equal to zero contained no vibrato. Otherwise, the procedure introduced in 2.2.2 was followed for synthesis.

An example of the result is shown in figure 4 (dashed line), with most frames belonging to a vibrato segment. The contributing units contours are represented in continuous lines. The top-most subfigure shows the pitch values of the transformed source units and the resulting pitch with vibrato. This vibrato was synthesized with the depth shown in the second sub-figure, where the two contributing units can also be observed. The vibrato rate is shown in the bottom subfigure.

3. EVALUATION

3.1 Experimental setup

We evaluated the achieved expression by conducting a MOS type test with 16 participants. The subjects rated the synthesized performances from 1-5 in terms of naturalness (rather synthetic or human), expressiveness (very inexpressive to very expressive) and singer skills (very bad or good timing, tuning, overall perception).

Three excerpts of 30 seconds were synthesized. For each of these excerpt, three versions were synthesized using different methods of generating expression contours. These were the baseline method based on heuristic rules, manual tuning of dynamics, pitch bend and vibratos, and finally the synthesis using the proposed system. All versions had background music.

The heuristic rules or default configuration was obtained following the algorithm described in [2]. Pitch and dynamics curves are obtained from the interpolation of a set of points generated by normal distributions (derived from real arpeggio performances). The amount of points used in the interpolation depends on the absolute note duration.

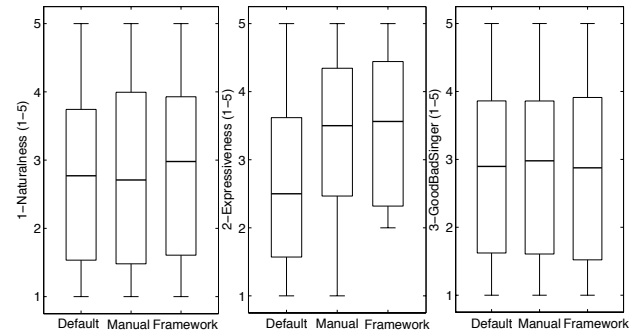


Figure 5. Results of listening tests.

With respect to the manual tuning, it was done by skilled experts who are used to generate singing performances with Vocaloid.

The expression database built for this evaluation contained melodic sections from four recorded songs in soul/pop style. In total, six minutes of a cappella singing voice were recorded by a female trained singer. The target songs were not present in this database.

The subjects first listened the three versions of the song being rated to get an overview of the variability within examples and then listened to them again in order to rate them individually. This was done separately for each song. The order in which songs were listened to was not always the same and versions were presented in a random order. These songs were synthesized using a Spanish voice bank. The rating task took around 15 minutes.

3.2 Results and discussion

In order to evaluate how the three different versions compare to each other, the results are grouped in terms of the control parameter configurations within each rated question. These are shown in figure 5, where the boxplots refer to naturalness, expressiveness and singer skills respectively. The statistics show the mean opinion scores, standard deviations (above and under mean) and minimums and maximums. Paired-samples t-tests were conducted to determine the statistical differences between the evaluated synthesis configurations with respect to a p-value threshold of 0.05.

Concerning naturalness, the three versions have been rated quite similarly. Although the proposed system has a slightly higher mean value, this difference is not statistically significant with respect to the baseline method and the manual tuning.

In terms of expressiveness, it can be observed that the baseline method has the lowest mean rating, followed by the manually tuned version which is slightly improved by our method. In this case, the differences between both the proposed system and the manual configuration with respect to the baseline method are statistically significant ($p=2.64 \times 10^{-6}$ and $p=3.23 \times 10^{-6}$, respectively). On the other hand, no statistically significant difference is observed between the proposed system and the manual configuration ($p=0.76$). Therefore, expression was improved using the proposed system and achieved a similar level to the man-

ual configuration.

Finally, with respect to whether the singer is good or bad, the three versions have a similar mean value. The differences between both the proposed system and the manual configuration with respect to the baseline method are not statistically significant.

The sound files used in the listening tests are online at: <http://www.dtic.upf.edu/~mumbert/smac2013/>.

4. CONCLUSIONS

In this paper we have introduced a new method for generating expression contours for singing voice synthesis based on unit selection. It is worth mentioning that our system does not rely on statistical models and therefore it is capable of preserving the fine details of the recorded expression. First, the steps for the expression database recordings and labeling have been detailed. With respect unit selection process, the four costs that are taken into account have been explained. These costs involve unit transformation and concatenation, alternative score generation and continuity cost. Unit transformation in time and frequency, unit concatenation with the crossfading masks, and contours rendering have been described.

From the listening tests, our system is capable to automatically generate a performance which is as expressive and natural sounding as can be achieved by manual tuning of parameters. Also, its naturalness and perceived singer skills are not worse than the baseline rule-based system.

Automatic generation of expression controls for a given target style has several advantages. It contributes to reducing the time a user spends in providing expression to singing performance. Another advantage is that it provides a richer starting point than the default configuration for manual expression tuning. More importantly, the proposed system paves the way towards modeling all of the aspects of expression for a singer in a particular style.

Concerning the future work, the expression database can be improved by realizing a comprehensive study of the note durations and intervals to cover in a given style. The cost functions can be adapted to new labeled data (type of note figures and strength, timing deviations, lyrics). The tremolo effect could be considered by modeling dynamics in a similar way as vibrato instead of the current smoothing step. We also plan to designing objective evaluation tests for unit selection, transformation and concatenation.

5. REFERENCES

- [1] H. Kenmochi and H. Ohshita, "Vocaloid - commercial singing synthesizer based on sample concatenation." in *INTERSPEECH*, 2007, pp. 4009–4010.
- [2] J. Bonada, "Voice processing and synthesis by performance sampling and spectral models," Ph.D. dissertation, Universitat Pompeu Fabra, Barcelona, 2008.
- [3] J. Janer, J. Bonada, and M. Blaauw, "Performance-driven control for sample-based singing voice synthesis," in *Proc. of the Int. Conf. on Digital Audio Effects (DAFx-06)*, Sept. 18–20, 2006, pp. 41–44.
- [4] M. Schröder, "Can emotions be synthesized without controlling voice quality," *Phonus 4, Forschungsbericht Institut für Phonetik, Universität des Saarlandes*, 1999.
- [5] T. Nakano and M. Goto, "Vocalistener: A singing-to-singing synthesis system based on iterative parameter estimation," in *Proceedings of the 6th Sound and Music Computing Conference. Porto*, 2009.
- [6] H. Zen, K. Tokuda, and A. W. Black, "Statistical parametric speech synthesis," *Speech Communication*, vol. 51, no. 11, pp. 1039–1064, 2009.
- [7] T. Nose and T. Kobayashi, "Recent development of HMM-based expressive speech synthesis and its applications," in *Proc. 2011 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2011)*, 2011.
- [8] H. Zen, T. Nose, J. Yamagishi, S. Sako, T. Masuko, A. W. Black, and K. Tokuda, "The HMM-based speech synthesis system (HTS) version 2.0," in *Proc. of Sixth ISCA Workshop on Speech Synthesis*, 2007, pp. 294–299.
- [9] K. Saino, M. Tachibana, and H. Kenmochi, "A singing style modeling system for singing voice synthesizers," in *INTERSPEECH*. ISCA, 2010, pp. 2894–2897.
- [10] A. Iida, N. Campbell, S. Iga, F. Higuchi, and M. Yasumura, "A speech synthesis system with emotion for assisting communication," in *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*, 2000.
- [11] D. Erro, E. Navas, I. Hernáez, and I. Saratxaga, "Emotion conversion based on prosodic unit selection," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 5, pp. 974–983, 2010.
- [12] E. Gómez and J. Bonada, "Towards computer-assisted flamenco transcription: An experimental comparison of automatic transcription algorithms as applied to a cappella singing," *Computer Music Journal*, In Press.

Woodwinds reeds and flutes

THE PLAYER–WIND INSTRUMENT INTERACTION

Joe Wolfe André Almeida Jer Ming Chen David George Noel Hanna John Smith

School of Physics, The University of New South Wales, Sydney

J.Wolfe@unsw.edu.au andregoiios@gmail.com jerming@phys.unsw.edu.au
david.george@student.unsw.edu.au n.hanna@unsw.edu.au john.smith@unsw.edu.au

ABSTRACT

Players control a range of parameters in the player-instrument system. First we show how loudness and pitch vary over the plane of mouth pressure and force on the reed of a clarinet, and thus how these parameters can be used in compensation to produce trajectories in this plane that have varying loudness and timbre but constant pitch. Next we present impedance spectra for several different types of musical instruments and for the vocal tract, to allow general observations. We report different ways in which the acoustic properties of the player's tract interact with those of the instrument bore to control the frequency of reed vibration in some wind instruments. We also show how vocal tract resonances can influence timbre.

1. INTRODUCTION

The player often is more interesting than the instrument: in general, a good musician on a poor instrument sounds better than a poor musician on a good instrument. Some of what makes a musician good lies outside the realm of science, let alone acoustics, but acousticians can contribute to understanding good performance by researching the player-instrument interaction, aiming to understand how good players achieve their musical goals. Beyond its intrinsic interest, this has possible applications: an improved and explicit understanding of how good players play could help guide students and teachers.

Some of the player-instrument interaction consists in the former doing something to the latter: pushing the right keys in the right sequence, applying certain values of force with the lips, certain pressures from the lungs, etc. These are of continuing interest in our lab, and we begin by presenting one such set of parameters.

Another part of the player-instrument interaction for wind instruments is directly acoustic in that it involves the acoustics of the player's vocal tract.

1.1 The elements

Performance involves the interaction of the principal acoustical elements of the wind instrument–player system: (i) a source of air at positive pressure, (ii) a vibrating element, usually an air jet, a reed or the player's lips, (iii)

the downstream duct, *i.e.* the bore of the instrument and (iv) the upstream duct, comprising the player's airway.

Players control all of these: (i) The air pressure and flow are controlled by muscles of the torso and also, in some cases, by the glottis. On the very short time-scale, flow is also controlled in an almost binary fashion by the tongue, which can cease the flow by contact with the roof of the mouth (in gestures like *ta*, *la*, *da*, *ka* etc) or by contact with the reed. (ii) The valves are diverse: Flutists control lip aperture size and geometry. Brass players vary the geometry and mechanical properties of the lips. Reed players choose or make their reeds and vary their mechanical properties with their lips. (iii) The geometry of the downstream duct is varied with valves and slides in brass instruments and by opening or closing covering tone holes in woodwind. (iv) In many cases, players vary the shape and position of the tongue, palate and the opening of the glottis, to control the acoustic properties of the upstream tract.

Two teams discussed the acoustics of the upstream duct at the last SMAC [1,2] and a friendly rivalry began. Scavone [1] presented circuit models of an upstream resonator and a downstream waveguide. We reported results using physical models of the vocal tract and either a cantilever valve or water filled latex 'lips' to represent player's lips [2]. We next used a broad-band signal and the capillary technique to measure the acoustic impedance in the mouth of a player while he played the didjeridu [3]. Later, we applied the technique to saxophone, clarinet and trumpet [4-6]. Meanwhile, Scavone developed a different technique to study the tract involvement: microphones inside bore and mouth give the ratio of the two impedances for harmonics of the note [7]. The two techniques are somewhat complementary and the teams progressed in parallel: Scavone's technique uses the vibrating reed as the (large) signal, which gives large signal:noise ratio and so allows rapid measurements and the ability to track rapid changes in time. However, because it only samples the frequency domain at multiples of the playing frequency, involvement of the vocal tract support must be inferred from a sparse representation (it does not directly measure tract resonances). Our technique gives the impedance spectrum in the mouth but, because our probe signal's energy is spread over hundreds of frequencies, we need windows of tens of ms up to seconds.

In this paper, we review aspects of the musician's control in all four areas and present new work. The review disproportionately cites our own work, in part because the player-instrument interaction has been one of the main lines of research in our lab since the last SMAC.

2. COMPENSATING CONTROL PARAMETERS

In many instruments, increasing pressure in the mouth, with all other parameters held constant, changes the pitch. On a recorder, the only parameter that the player can use to control the jet is the mouth pressure. Consequently, it is difficult to play a *decrescendo* on a sustained note without a fall in pitch. Some accomplished players compensate for these pitch changes using the downstream duct by partially opening and closing tone holes.

In other instruments, the effect of changing pressure may be compensated with other parameters to keep the pitch constant. In reed instruments, for a given configuration of the bore, the pitch is also dependent on the forces applied to the reed and to the configuration of the upstream duct. Consequently, a player can play a *decrescendo* at constant pitch by following an iso-frequency contour through the space of control parameters.

The clarinet is the 'lab rat' of wind instruments: many scientific studies and models exist [e.g. 8-10]. One way to study control parameters is to measure some of them in the player as in [11] or in a blowing machine, as in [12]. Here we use a systematic study of how pitch, loudness and spectrum depend on mouth pressure, lip force and reed parameters. To control these parameters and to maintain a constant configuration of the upstream duct, we use an automated clarinet playing system.

2.1 The automated clarinet player

The clarinet player was built partly for this purpose [13] and partly for a competition (Artemis International) for such automata. A film clip of the player in concert is at www.phys.unsw.edu.au/jw/clarinetrobot.html. A pump provides air whose pressure is regulated in the long term by controlling the pump and on short time scales via a controlled leak. In this experiment, the reed force was applied by the weight of suitable masses through a soft sorbothane pad, which provided losses somewhat like those from a human lower lip. The reeds are synthetic (Légère, Barrie, Ontario) with 'hardness' ratings 2 and 3.5, the clarinet a Yamaha YCL250 with a Yamaha CL-4C mouthpiece.

2.2 Pitch and loudness control

Figure 1 plots the frequency f and sound level L as functions of the gauge pressure P in the 'mouth' and the force F applied by the 'lip'. The force is applied at 10 mm from the reed tip. The shaded region represents the parameter combinations that produce notes, except that the upper pressure limit near 7 kPa is imposed by the pump used. The reed has a manufacturer's 'hardness' rating of 2, the note played is F4 (written G4) and the frequencies are a little low because the instrument is playing at 19 ± 1 °C. This note is in the clarinet's 'throat register': it is an octave and a minor third above the lowest note, it is still in the first mode of vibration, and 12 tone holes are open.

Above the shaded region, no sound is produced because the combination of lip force and pressure push the reed closed against the mouthpiece. This extinction line has a

negative slope with a magnitude of a few cm^2 . Left of the edge of the playing region, with a rather steeper slope, either no sound or squeaks are produced. At very low F , the damping is insufficient to prevent squeaks: i.e. sounds with frequencies close to the resonance of the reed, rather than the resonance of the bore.

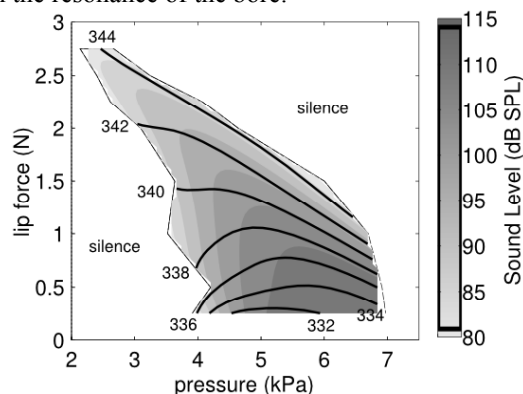


Figure 1 In the space of pressure in the artificial mouth and force exerted on the reed, we plot frequency in Hz (black lines) and sound pressure level (shading) measured at 50 mm from the bell on the axis of the clarinet.

f is higher at higher F and also, over most of the playing regime, at higher P . The iso- f lines have negative slopes of a few cm^2 over most of the playing regime. This is qualitatively explained by a simple effect: Both F and P tend to close the reed against the mouthpiece, which reduces the compliance of the air in the mouthpiece and also the effective compliance of the reed. This raises the frequency of the peaks in the parallel impedance of reed and bore.

Figure 1 shows that, over part of the range, a player could vary sound level and constant frequency by blowing harder and making a compensating reduction in lip force, and adjusting no other parameters. At high P and F , however, the iso- f lines and the iso- L contours are nearly parallel. So, if no other control parameters were used, the player would need to fall to a lower pitch (by relaxing the jaw (lower F)) to play more loudly. Going flat when playing loudly is a fault often identified by clarinet teachers.

In most instruments, vibrations of increased amplitude produce greater nonlinearities in the valve, with the consequence that the amplitudes of higher harmonics increase more rapidly than those of the fundamental. A simple quantification of this is the spectral centroid, the frequency weighted average of the amplitude of all spectral components. The spectral centroid is strongly correlated with the perceived brightness of timbre. Contours of the spectral centroid on the P, F plane have shapes similar to those of sound level (data not shown).

F4, the note shown in Figure 1 is one of the easiest to sound on the clarinet. For notes one octave below and one octave above, the playing regime is rather smaller. Lip force application at 10 mm from the reed tip facilitates sound production: in experiments with the lip force applied 5 mm either side of this position, the playing regimes are smaller. Finally, in similar experiments using a stiffer reed, the results are qualitatively similar, but notes are only produced in the higher range of F , above about

3 N. Space precludes including these data here: they have been submitted for publication elsewhere [14].

3. THE UPSTREAM AND DOWNSTREAM DUCTS

In some cases, there is a limited symmetry between the two ducts [15]. Call the total and acoustic pressures immediately upstream from the reed, inside the player's vocal tract, P_{mouth} and p_{mouth} and use the subscript bore for those immediately downstream, in the duct provided by the instrument. A clarinet reed tends to bend inwards towards the mouthpiece, thus tending to close the flow pathway, when the pressure difference $P_{\text{mouth}} - P_{\text{bore}}$ is positive. Under some playing conditions, $P_{\text{mouth}} - P_{\text{bore}}$ acts to bend a brass player's lips outwards into the instrument, opening the flow pathway [16,17]. Defining the acoustic impedances Z_{mouth} and Z_{bore} as p/U , where U the acoustic volume flow into the duct then, if the flows out of the mouth ($-U_{\text{mouth}}$) and into the instrument (U_{bore}) are equal, say U , then $p_{\text{mouth}} - p_{\text{bore}} = -UZ_{\text{mouth}} - UZ_{\text{bore}} = -U(Z_{\text{mouth}} + Z_{\text{bore}})$. In words, Z_{mouth} and Z_{bore} are in series with regard to the mechanisms described. It is therefore worth comparing and contrasting the impedance spectra of some of the ducts involved, which we do in the next section.

It can also be shown that the passive impedance of a reed or lip is in parallel with the series combination mentioned above. Further, the pressure difference $P_{\text{mouth}} - P_{\text{bore}}$ across the valve is not the only source of force acting to open or close it: for instance, the dynamic or 'Bernoulli' pressure $\frac{1}{2}\rho v^2$ can also play a role, so the two ducts are not necessarily in series with respect to all possible regeneration mechanisms in the valve.

3.1 Impedance measurements

With one exception, the impedance spectra shown in Figure 2 are made using the three-microphone technique calibrated with three non-resonant loads, one of which is an acoustically infinite duct [18]. The smallest microphone spacing in this impedance head is 40 mm, which limits the frequency response to about 4 kHz. The lower limit is about 10 Hz, giving a range of 9 octaves. The use of non-resonant calibration loads and the iterative optimisation of the measurement signal together allow a dynamic range of more than 90 dB. The frequency resolution depends on the period of the measurement signal.

3.2 Ducts, resonances and antiresonances

Figures 2, 3 and 4 show the impedance spectra Z of some simple ducts, several instruments and a vocal tract, to allow some general discussion. Figure 2 shows the measured impedance spectra of a number of ducts. (a) shows the impedance of an open cylinder with an internal diameter of 15 mm. (b) shows an open cone with a half angle of 1.74° (equal to that of a soprano saxophone) [19]. The cone is truncated at the small (input) end to allow flow into the cone via the impedance head used to measure it, and the truncation replaced with a cylinder of equal volume. Their effective lengths are 325 mm, cho-

sen so that the first maximum in Z of the cylinder occurs at C4 and the first minimum at C5.

3.3 The downstream duct: woodwinds

The player's control of the downstream duct often involves complicated co-ordinations, such as manipulating different keys in the same transition. Departures from simultaneity are sometimes systematic [20], which raises interesting acoustic and pedagogical questions, which we shall not pursue further here, but instead concentrate on quiescent states.

Below the cylinder in Figure 2 we show Z for a flute [21] and clarinet [22], whose bores include cylindrical sections. (To save space, phase is not shown for these curves.)

The array of open tone holes creates a cut-off frequency, which is about 1.5 kHz for the clarinet. Below this frequency, sound waves are reflected at a point near the first of the open holes. Well above this, they travel down the whole length of the bore, which explains why the spacing of the peaks in Z decreases above that frequency. That cut-off for the flute is around 2 kHz but the more closely spaced peaks are less evident, because at several kHz the bore is short circuited by a shunt provided by the Helmholtz resonator in the head joint [21].

The observation that there is only a little similarity between Z for these instruments and that for a cylinder explains why highly simplified models fail: for example, the statement that clarinets have weak even harmonics is not usually true.

Below the cone are shown Z for soprano and tenor saxophones [19]. The soprano sax, the clarinet, the flute and the two simple geometries all have the same effective length, so all play C5 in this configuration (dashed circles) except the clarinet, which plays C4.

The tenor saxophone is shown on an expanded scale to show that it is roughly a 1:2 scale model, playing an octave below, although the cone half-angles are different: 1.74° and 1.52° respectively.

For the saxophones and clarinet, the dotted line shows the bore impedance in parallel with the compliance of a typical reed. This is important in intonation: a soft reed (as much as twice the compliance used here) lowers the peak of the parallel impedance and the instrument would play flat unless the tuning slide were shortened.

Note that the saxophones have only two large peaks in Z . The clarinet has three and, above the cut-off frequency, further sharp peaks, whose narrower frequency spacing indicates that they are standing waves over the whole length of the bore. The conical bores of saxophones give little reflection for high frequencies, and so only weak standing waves. As we'll see later, this makes the vocal tract especially important for the high range of the saxophone.

3.4 The downstream duct: lip valves

The Z for four lip valve instruments are shown in Figure 3. (To conserve space, phase is not displayed.)

The first is a didjeridu. The first several extrema are qualitatively consistent with a slightly flared duct, as the

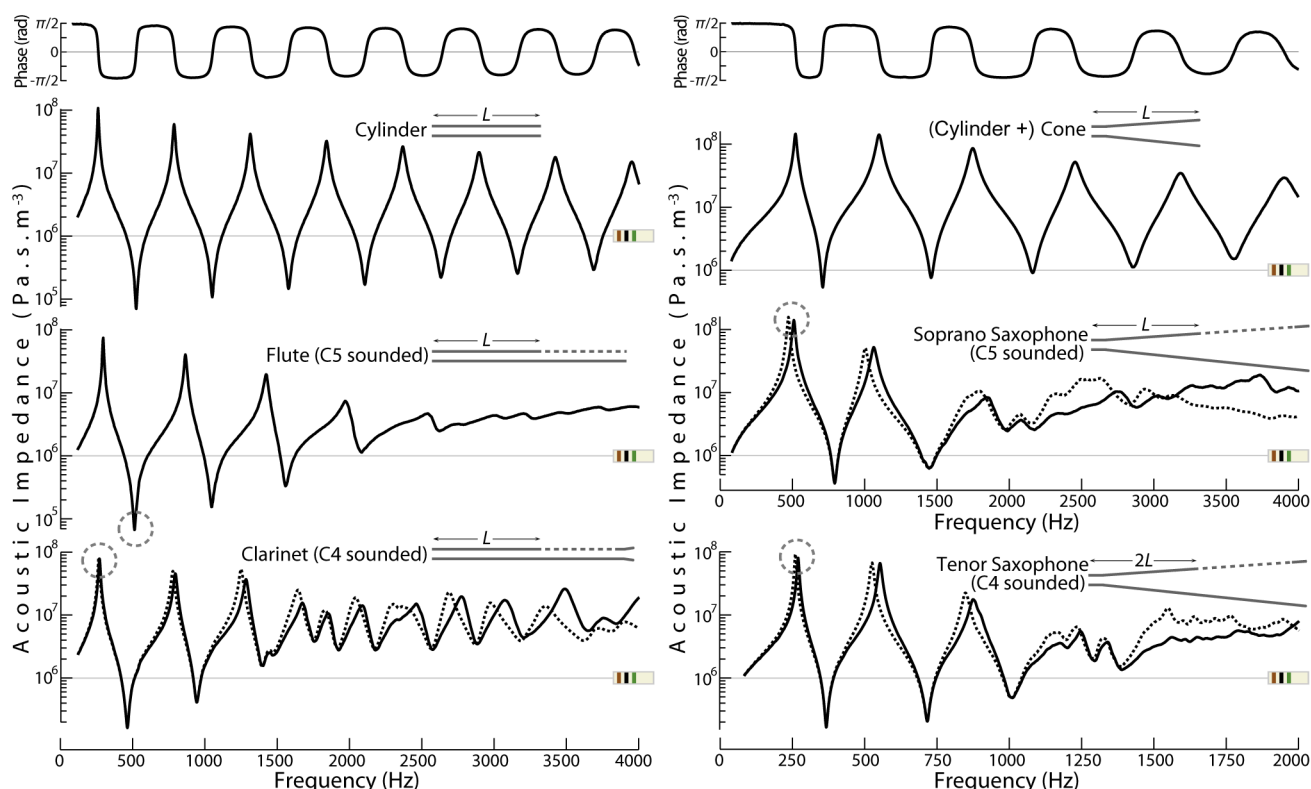


Figure 2. Impedance spectra. At top, a cylinder (left) and a truncated cone (right). These have the same effective length as the flute and soprano saxophone (fingering for C5) and the clarinet (C4). The tenor saxophone has the same fingering as the soprano but is plotted on an expanded frequency axis to show that it is roughly a 1:2 scale model. (Some phase plots are omitted to save space.)

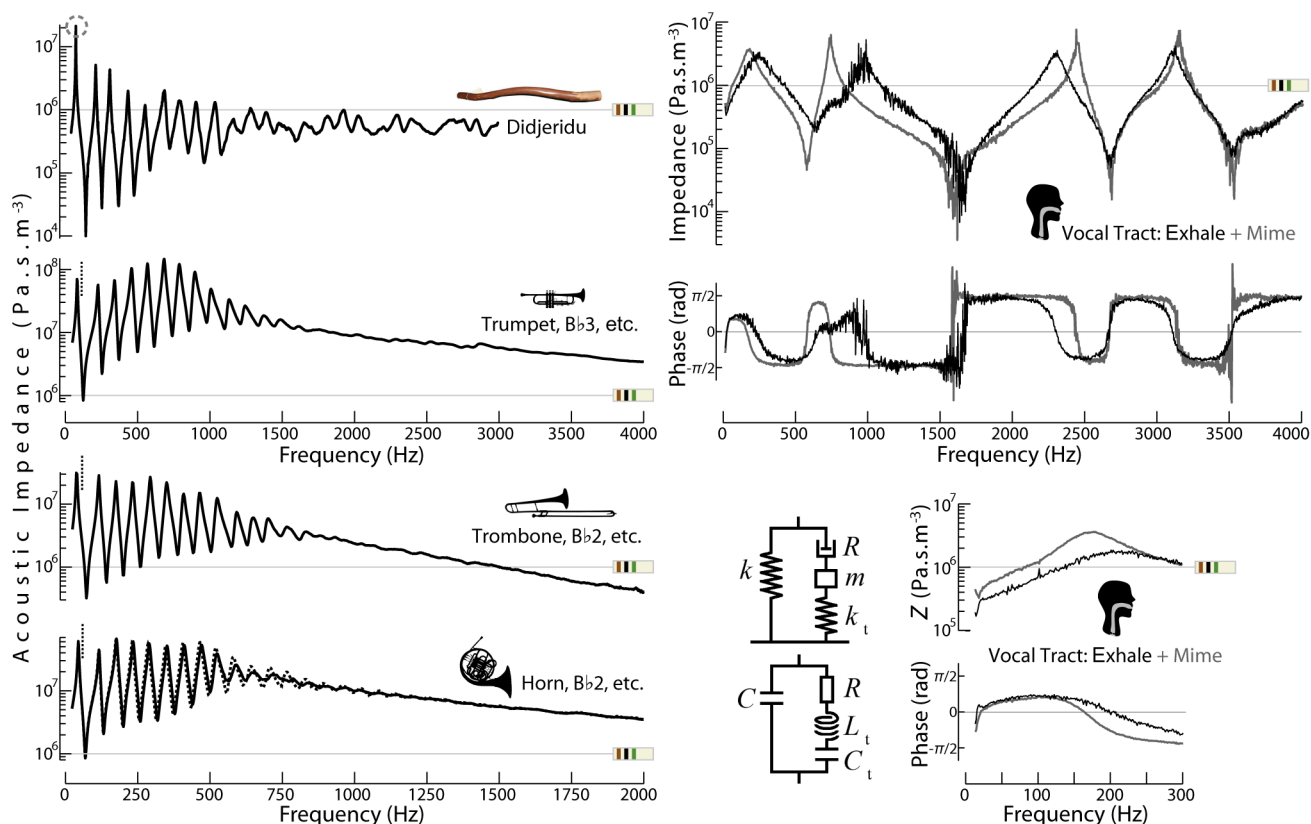


Figure 3. Impedance spectra of lip-valve instruments: a didjeridu, trumpet, trombone and horn: the last two on expanded scales as they are approximately 2:1 scale models of the trumpet. On Figures 2, 3 and 4, the 1 MΩ line is marked.

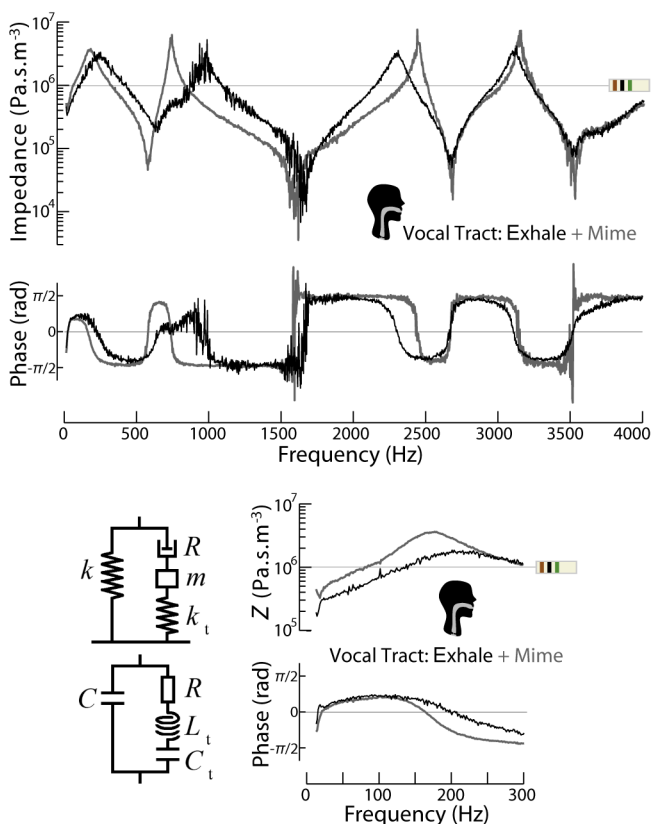


Figure 4. Impedance spectra of a vocal tract measured at the lips over 370 ms: glottis closed (pale) and exhaling (black). The compact-object equivalent circuit is shown next to the plots of low frequency behaviour on an expanded axis.

outer shape suggests. The impedance peak at which it plays is indicated (dashed circle). The irregular features in Z and the absence of strong resonances at high frequencies are due to the irregular interior, which is made by termites in a branch or trunk of a eucalypt tree.

The Z for trombone (slide in) and horn (B \flat side, no valves depressed) are shown on an expanded scale to show that they are, very approximately, scale models of the trumpet (no valves depressed). In all three cases, the cut-off frequency is provided by radiation from the bell, which becomes important when the wavelength is comparable with the radius of curvature of the bell. For the horn, the hand in the bell (dotted line) reduces the efficiency of radiation and so raises the cut-off frequency and the height of the higher frequency peaks.

The three brass instruments each have about 10 or 12 impedance peaks. Starting from the second, these are approximately equally spaced. In this configuration, notes near all of those peaks are played. They correspond to a harmonic series whose fundamental is half the frequency of the second peak. A note with that pitch can be played as a pedal note (dashed vertical line). (The first peak, well below the pedal pitch, is not played.)

Comparing these plots, we might expect that considerably more control over the valve is required of a brass player, whose impedance peaks are more closely spaced. In contrast, when the saxophonist and clarinetist wish to use the second (or third), weaker peak, they can simply use register key(s) to weaken and to detune the lower peak(s).

3.5 The upstream duct: the vocal tract

Figure 4 shows Z for a vocal tract. The impedance head has a diameter of 26 mm and the lips are sealed around that, as described in [23]. The tongue is in the position to pronounce /æ:/ as in 'heard'. For the grey curve, the glottis is closed, so there is no DC flow. For the black curve, the subject is exhaling into the impedance head, which has a downstream vent for this purpose.

Nine octaves covers the first five resonances. Because of the long period of the first resonance, the measurements were made over a single window of 370 ms ($= 2^{14}/44.1$ kHz). For the exhalation case, this shows a measurement limitation: noise due to the turbulent air flow is superposed over the broad-band measured signal. Integration over longer time-scales can improve this, but the subject must sustain the gesture for longer.

Opening the glottis (going from a closed to a somewhat open pipe), one expects extrema in impedance to rise slightly in frequency. For example, whisper uses a larger glottal opening than normal speech, and measurements of the tract resonances from normal phonation to whisper in the same gesture show an increase in frequency [24]. The measurements in Figure 4 were not measured in the same gesture, so other geometrical changes may also contribute, particularly at high frequencies. It is believed that players use a slightly open glottis [25], so a playing configuration could be between these two conditions.

If the vocal tract were a rigid cylinder, 170 mm long and closed at the glottis, we should expect minima at 0.5,

1.5, 2.5 and 3.5 kHz, and maxima at about 1, 2 and 3 kHz, which is roughly what we observe.

In a closed, rigid cylinder, Z would be very large at very low frequencies. Human tissues are not infinitely rigid, of course, and this gives rise to the zeroth minimum and maximum, shown in the inset in Figure 4. The maximum at about 200 Hz is due to the mass of the tissue surrounding the tract and the 'spring' of the air inside it, which is approximately a compact object at this wavelength (2 m).

The minimum at about 20 Hz is due to the mass of the tissue and the 'spring' of its own elasticity. Because the two resonances are three octaves or more apart, impedance maxima and minima occur at $\omega \sim 1/(LC)^{1/2}$ and $1/(LC_t)^{1/2}$ respectively, where L is the inertance of the tissue, C_t the compliance due to the supporting tissues, and C the compliance of the air in the upper tract. Taking a volume of 100 ml for the air in the tract and a surface area of the surrounding tissues of 10^{-4} m², L gives a tissue thickness of ~ 1 cm and C_t a spring constant of $\sim 1/\text{N.cm}^{-1}$.

3.6 Varying the resonances of the upstream duct

How much can the upstream resonances be adjusted by articulation? Opening and closing the glottis makes little difference at high frequency, because the inertance of air in the glottis effectively seals it. Of course, other geometrical changes associated with this variation also may make a difference. In speech, the first resonance is varied primarily by varying the opening at the lips, which connects the tract to the very low impedance of the radiation field. This option is not available to reed and brass players, whose mouths make an airtight seal to the instrument. So the first resonance cannot be varied much and, in our measurements on a range of instruments, we usually see an acoustic maximum between about 200 and 400 Hz, and a minimum a few hundred Hz above that.

The shape of the vocal tract can be used to adjust the resonant frequencies. This is most effective once the half wavelength becomes comparable with the length of the upper vocal tract – when the frequency approaches 1 kHz, but even the first maximum can be varied. As a general rule: a constriction near a pressure node lowers the frequency while one near a pressure antinode raises it.

Let's compare the magnitudes of impedance peaks in figures 3, 4 and 5: those in the vocal tract, with jaw low and tongue neutral, are several M Ω . This already approaches the values of the instrument impedance at high frequency. Further, players can raise this value by raising the tongue near the instrument, which makes the front of the mouth act like an impedance matching cone and produces rather higher impedances, as we shall see. It is easier for an impedance peak in the tract to compete with one in the bore at high frequencies, where the bore resonances are weaker.

4. VOCAL TRACT EFFECTS

4.1 Tract-bore series combination can control pitch

A reed and the air that flows past it are driven by the pressure difference between mouth and bore, which

means that, as described above, the tract and bore act in series.

For the saxophone, this solves a problem: Figure 2 shows that, because of effective radiation and weaker reflection from the large end of the cone, the instrument's Z has only two strong impedance peaks. Although the first register is extended downwards in pitch with the use of extra keys, and the second similarly upwards, the range of the instrument using these first two peaks is just 2 octaves and a musical fifth.

Figure 5 contrasts the situation for a note, G4, in the second register of a tenor saxophone and A#5, in the *altissimo* range. Even beginners can play the first note, with relatively little attention to control parameters. The *altissimo* notes require expertise, because they are only possible when a large amplitude peak in Z_{mouth} is produced and tuned to select one of the instrument's weaker higher resonances [4].

Advanced saxophonists can also use large magnitude peaks in Z_{mouth} to select and to adjust the amplitude of notes in multiphonics (chords) and also to 'bend' the pitch of notes [26].

Because its bore is partly cylindrical, the higher resonances in the clarinet produce stronger peaks in Z_{bore} (Figure 2), so less assistance from the upstream duct is required for *altissimo* playing. However, like saxophonists, clarinetists use peaks in Z_{mouth} in multiphonics and in pitch bending, including the famous *glissando* in *Rhapsody in Blue* [5].

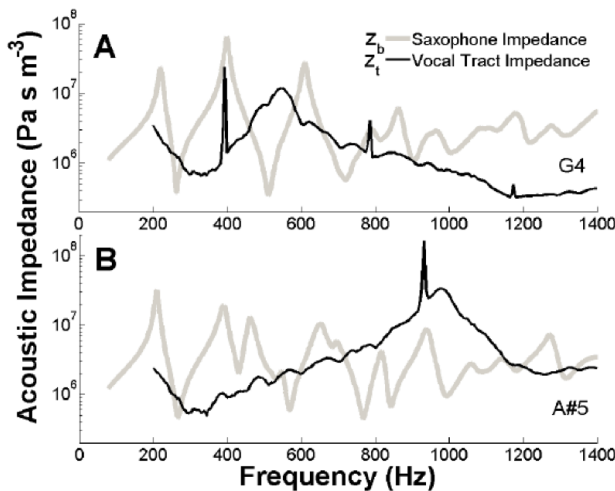


Figure 5. The bore (grey) and vocal tract impedance (black) for two notes played on a tenor saxophone. The narrow peaks are harmonics of the notes played. G4 is in the instrument's second register. A#5 is in the *altissimo*, and is accessible only by tuning a peak in Z_{mouth} to select one of the weak, higher resonances. Reproduced from [4].

Figure 6 shows the vocal tract resonances of saxophonists, clarinetists and trumpeters. In the normal range of the saxophone, neither advanced nor less advanced players show tuning of peaks in Z_{mouth} to the played note. In the *altissimo* range, however, peaks are tuned either to or slightly above the note to be played and those unable to do this cannot play in this register.

For clarinetists, peaks in Z_{mouth} are tuned to the note to be played during pitch bending. In normal playing, however, peaks in Z_{mouth} are tuned a couple of hundred Hz above the note to be played, which means that the upstream impedance is inertive [5].

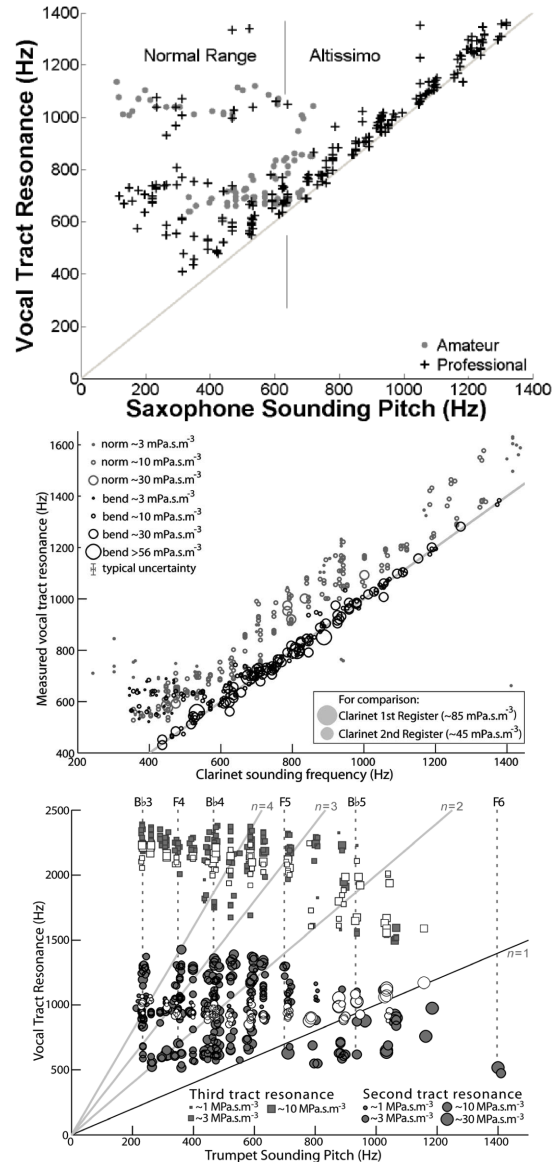


Figure 6. Vocal tract resonance frequency against pitch frequency. Top: tenor saxophone in normal (left) and *altissimo* ranges, showing resonance tuning in the latter. Middle: clarinet pitch bending (black) and normal playing (gray): resonance tuning for pitch bending, and the resonance held somewhat above the pitch in normal playing. Bottom: the seven trumpet players in this study show no consistent resonance tuning. Reproduced from [4,5,6].

What about trumpet players? Some players can play in the upper part of the third octave, where the peaks in Z_{bore} are weak. Do they tune a peak in peaks in Z_{mouth} to assist playing in this *altissimo* range? When we started this study [6], we expected to find vocal tract tuning like that of saxophonists. But the answer, in general, is no. Figure 6 shows no consistent tuning by the seven players in our study.

So, what are the effects of Z_{mouth} in players of brass and other lip-valve instruments? At the previous SMAC, we presented Figure 7, the results of a trombone-playing system using, as the 'lip-valve', an outward swinging cantilever spring with a mass added to give a suitable natural frequency. Upstream, we used different shaped simple cavities as models of the mouth with tongue low in the mouth and high in the mouth. The trombone slide was then moved into its standard positions for notes on the chromatic scale. While the valve operated on the same peak in Z_{bore} (i.e. played in the same register), the pitch decreased by approximately a semitone for each step increase in the slide position, as expected. Depending on the natural frequency of the valve, the system would jump from one register to the next at a particular position.

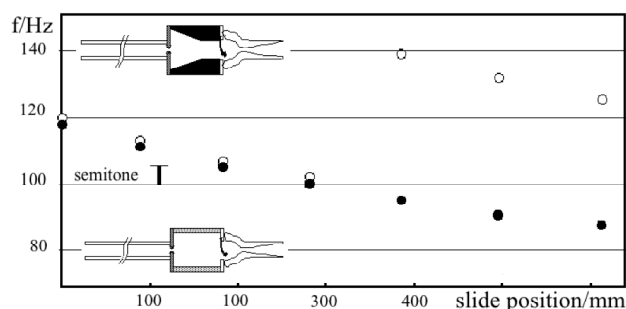


Figure 7. A trombone is 'played' with different slide positions by an artificial valve with two upstream cavities with different shapes.

Without the constriction in place ('low tongue'), the system played the same register flatter than with the constriction in place ('high tongue'). Further, the jump between registers occurred at lower pitches without the constriction. This is consistent with players' observations that lowering the tongue either lowers the pitch slightly or else causes the instrument to jump to a lower register. More recent measurements of the player-trombone interaction are given in another paper in this volume [27].

4.2 Vocal tract interactions with higher harmonics

In the saxophone and clarinet, impedance peaks in Z_{mouth} can contribute to reed vibration, either on their own or in collaboration with peaks in Z_{bore} . Peaks in Z_{mouth} can also, of course, inhibit acoustic flow. Which of these applies?

Players of various wind instruments report that they use different configurations of the vocal tract to control timbre. This is most spectacularly evident on the didjeridu, an indigenous Australian instrument, which is played almost entirely at its lowest resonance. Its musical interest comes from rhythmic variations in timbre produced, *inter alia*, by varying vocal tract shapes, including those used for the cyclic breathing that allows the didjeridu to be played continuously without pauses for breath.

At the last SMAC, we also presented this sound file www.phys.unsw.edu.au/jw/sounds/dij_trombone.wav. It was produced by a pair of artificial lips (water-filled latex cylinders) with a cylindrical pipe downstream modelling the didjeridu. Upstream was a model vocal tract with continuously variable resonances.

Figure 8 shows the spectrum of sound radiated from a cylindrical pipe being played as a didjeridu by a human player. While the subject played, a small impedance head placed between the lips measured Z_{mouth} using the capillary method. On the same figure, vertical ticks indicate the harmonics of the note played and the resonances of the pipe. Unlike a typical clarinet note, the first eleven odd harmonics in this note really are all weaker than their even neighbours, though of course this is not the case with a real didjeridu. The spectral envelope shows minima around 1.5 and 2 kHz. The lower part of the figure shows that Z_{mouth} has maxima at these frequencies: at these frequencies, very little current can pass between the lips, so there is little power input to the instrument, and so little power in the radiated sound.

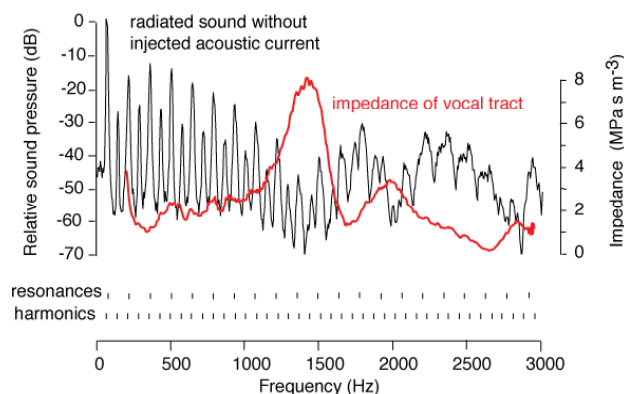


Figure 8. A cylindrical pipe played like a didjeridu. The spectrum of the radiated sound, the frequencies of the pipe resonances and the harmonics and the impedance spectrum of the vocal tract.

The didjeridu has no constriction comparable to that in the mouthpiece of a typical brass instrument and the magnitude of its Z_{bore} is, overall, rather less than those of typical woodwind or brass instruments (see Figures 2-4). Are similar effects observed on other wind instruments? In other papers in this volume, we report Z_{mouth} measurements for the saxophone and the trombone [27,28].

5. CONCLUSIONS

Players must simultaneously control several parameters simultaneously in order to produce desired contours of pitch, loudness and timbre. Features of the acoustic impedance of the vocal tract contribute to pitch, especially when an impedance peak lies near the playing frequency, and to timbre, when peaks fall close to higher harmonics.

Acknowledgments

We thank the Australian Research Council for support, Yamaha for instruments and our volunteer subjects.

6. REFERENCES

- [1] G.P. Scavone, "Modeling vocal-tract influence in reed wind instruments," in Proc. Stockholm Music Acoustics Conf. (SMAC 03) 2003, Stockholm, pp. 291–294.
- [2] J. Wolfe, A.Z. Tarnopolsky, N.H. Fletcher, L.C.L.

- Hollenberg, and J. Smith, "Some effects of the player's vocal tract and tongue on wind instrument sound," in *Proc. Stockholm Music Acoust. Conf. (SMAC 03)*, Stockholm, 2003, pp. 307-310.
- [3] A. Tarnopolsky, N. Fletcher, L. Hollenberg, B. Lange, J. Smith, and J. Wolfe, "Vocal tract resonances and the sound of the Australian didgeridu (yidaki) I: Experiment", in *J. Acoust. Soc. America*, 2006, **119**, pp. 1194-1204.
- [4] J.-M.,Chen, J. Smith, and J. Wolfe, "Experienced saxophonists learn to tune their vocal tracts," in *Science*, 2008, **319**, p726.
- [5] J.-M.,Chen, J. Smith, and J. Wolfe, "Pitch bending and glissandi on the clarinet: roles of the vocal tract and partial tone hole closure," in *J. Acoust. Soc. Am.*, 2009, **126**, pp. 1511-1520.
- [6] J.-M.,Chen, J. Smith, and J. Wolfe, "Do trumpet players tune resonances of the vocal tract?" in *J. Acoust. Soc. Am.*, 2012, **131**, pp. 722-727.
- [7] G.P. Scavone, A. Lefebvre, and A.R. Da Silva, "Measurement of vocal-tract influence during saxophone performance," in *J. Acoust. Soc. Am.*, 2008, **123**, pp. 2391-2400.
- [8] J. Backus, "Small-Vibration Theory of the Clarinet," in *J. Acoust. Soc. Am.*, 1963, **35**, pp. 305-313.
- [9] J. Gilbert, J. Kergomard, and E. Ngoya, "Calculation of the steady-state oscillations of a clarinet using the harmonic balance technique", in *J. Acoust. Soc. Am.*, 1989, **86**, pp. 35-41.
- [10] J.-P. Dalmont and C. Frappé, "Oscillation and extinction thresholds of the clarinet: Comparison of analytical results and experiments", in *J. Acoust. Soc. America*, 2007, **122**, pp. 1173-1179.
- [11] Ph. Guillemain, Ch. Vergez, D. Ferrand and A. Farcy, "An instrumented saxophone mouthpiece and its use to understand how an experienced musician plays" in *Acta Acust.*, 2010, **96**, pp.622-634.
- [12] T.A.Wilson, and G.S. Beavers, "Operating modes of the clarinet", in *J. Acoust. Soc. America*, 1973, **56**, pp. 653-658.
- [13] A. Alméida, J. Lemare, M. Sheahan, J. Judge, R. Auvray, K.S. Dang, S. John, J. Geoffroy, J. Katupitiya, P. Santus, A. Skougarevsky, J. Smith, and J. Wolfe, "Clarinet parameter cartography: automatic mapping of the sound produced as a function of blowing pressure and reed force" in *ISMA2010*, Sydney & Katoomba.
- [14] A. Almeida, D. George, J. Smith, and J. Wolfe, "The clarinet: how blowing pressure, lip force and reed hardness affect pitch, loudness and timbre," in *J. Acoust. Soc. America*, submitted.
- [15] A.H. Benade, "Chapter 35: Air column, reed, and player's windway interaction in musical instruments," in *Vocal Fold Physiology, Biomechanics, Acoustics, and Phonatory Control*, 1985, Denver, pp. 425-452.
- [16] S.J. Elliott, and J.M. Bowsher, "Regeneration in brass wind instruments," in *J. Sound Vib.*, 1982, **83**, pp. 181-217.
- [17] S. Yoshikawa, "Acoustical behavior of brass player's lips" in *J. Acoust. Soc. Am.*, 1965, **97**, pp. 1929-1939.
- [18] P. Dickens, J. Smith, and J. Wolfe, "High precision measurements of acoustic impedance spectra using resonance-free calibration loads and controlled error distribution," in *J. Acoust. Soc. Am.*, 2007, **121**, pp. 1471-1481.
- [19] J.-M.,Chen, J. Smith, and J. Wolfe, "Saxophone acoustics: introducing a compendium of impedance and sound spectra, " in *Acoustics Australia*, 2009, **37**, pp. 18-23.
- [20] A. Almeida, R. Chow, J. Smith, and J. Wolfe, J. "The kinetics and acoustics of fingering and note transitions on the flute," in *J. Acoust. Soc. Am.*, 2009, *J. Acoust. Soc. Am.*, **126**, pp. 1521-1529.
- [21] J. Wolfe, J. Smith, J. Tann, and N.H. Fletcher, "Acoustic impedance of classical and modern flutes" in *J. Sound Vib.*, 2001, **243**, pp. 127-144.
- [22] P. Dickens, R. France, J. Smith, and J. Wolfe, "Clarinet acoustics: introducing a compendium of impedance and sound spectra", in *Acoustics Australia*, 2007, **35**, pp. 17-24.
- [23] N. Hanna, J. Smith, and J. Wolfe, "Low frequency response of the vocal tract: acoustic and mechanical resonances and their losses" in *Proc. Acoustics 2012 Fremantle*, Australia.
- [24] Y. Swerdlin, J. Smith, and J. Wolfe, "The effect of whisper and creak vocal mechanisms on vocal tract resonances", in *J. Acoust. Soc. America*, 2010, **127**, pp. 2590-2598.
- [25] M.S. Mukai, "Laryngeal movements while playing wind instruments," in *Proc. Int. Symp. Music Acoust.*, 1992, Tokyo, pp. 239-242.
- [26] J.-M.,Chen, J. Smith, and J. Wolfe, "Saxophonists tune vocal tract resonances in advanced performance techniques," in *J. Acoust. Soc. Am.*, 2011, **129**, pp. 415-426.
- [27] H. Boutin, N. Fletcher, J. Smith and J. Wolfe "The playing frequency of the trombone and the impedances of the upstream and downstream ducts" in *SMAC13*, 2013, Stockholm.
- [28] Li, W.-C., J.-M. Chen, J. Smith, and J. Wolfe, "Vocal tract effects on the timbre of the saxophone" in *SMAC13*, 2013, Stockholm.

Determination of 2D Quasi Incompressible Flow around a Recorder Labium: a Comparison between Different Methods

R. Auvray, B. Fabre

LAM - D'Alembert,
UPMC Univ Paris 06, UMR CNRS 7190,
11 rue de Lourmel, 75015 Paris, France
auvray@lam.jussieu.fr
benoit.fabre@upmc.fr

P.-Y. Lagr  e

CNRS, D'Alembert,
UPMC Univ Paris 06, UMR CNRS 7190,
4 place Jussieu, 75005 Paris, France
pierre-yves.lagree@upmc.fr

ABSTRACT

The shape of the labium has important consequences on the sound produced by flute-like instrument. This statement is well known by instrument makers who take extremely care with this precise part. The sharp edge of the labium modifies both acoustic and hydrodynamic properties. Non-linear acoustic phenomena and intricate vortex structures might strongly depend on the shape of the labium. A first step in the study of the labium is to consider the acoustic part only. This paper presents a comparison of different numerical methods to estimate the linear part of the acoustic flow around the labium. Results are discussed with respect to the sharpness and the angle of the labium, two main features of the labium that have already been studied experimentally. Finally, one of the methods is proposed as a good candidate to include in a sound production model.

1. INTRODUCTION

In flue instruments a planar air jet is blow towards a sharp edge called the labium. The shape of the labium is known by instrument makers to be crucial in timbre and attack response of flute-like instruments. Whether it is a recorder, a flute or an organ flue pipe, the sound properties of such instruments are not intended to be the same, and so are their labium shapes.

From a physical point of view, the presence of a relatively sharp edge within the resonator and also within the jet flow will have consequences on the acoustics and the aeroacoustics of the instrument, respectively.

The constriction of the window and the sharp edge affect the acoustic transverse flow around the labium. Intuitively, the acoustic velocity is expected to be accelerated while approaching the labium. For high amplitude of oscillation, the flow separation may occur at the labium [1]. The mechanisms of vorticity generation at the labium have been modeled by Howe [2], and identified as strong acoustic damping mechanisms by Howe [3] and Fabre *et al.* [4]. The sharpness and the angle of the labium are associated

with the generation of vorticity, and thus have a great influence on this damping mechanism.

Conversely to the generation of vorticity at the labium, the generation and modulation of the vorticity of the jet within the window is often discussed as the core of sound production modeling. Based on works of Howe [2] and Nelson *et al.* [5], Dequand *et al.* [6] proposed a *discrete-vortex* model in which sound production is ensured by the interaction of propagating discrete vortices with the acoustic flow. According to this model, the source mechanisms depend on the orientation of the acoustic flow near the labium and might be modified by a change in the labium geometry.

Experimental observations of Dequand *et al.* [6] confirm the influence of the labium angle on the sound spectra. A further validation of the *discrete-vortex* model would be to include in the model the labium geometry influence through the modification of the acoustic streamlines and to check the resulting sound spectra. A first step is to develop a consistent description of the acoustic flow within the window. This is the aim of the present study.

Four different methods to estimate the flow acoustic around the labium are presented in section 2 and compared in section 3. The effect of two main characteristics of the labium, the sharpness and the angle, are then presented in section 4. Results are discussed in section 5.

2. PRESENTATION OF THE METHODS

In an aeroacoustical analysis of sound production in such instruments, where a jet interacts with an acoustic resonator, the exact definition of the acoustic flow is one of the key points. Following Howe [2], the acoustic flow is here defined as the fluctuating part of the potential component of the total flow.

This section presents four methods to study such a potential flow. Three of them are based on the assumption that the flow within the window is incompressible since the length of the window W is much smaller than the acoustic wavelength. This is therefore a low frequency approximation. In this case, the two components of the flow are denoted u and v . In the last method, the compressibility is taken into account and the two components of the flow are then denoted u' and v' .

The two first methods (complex potential and Schwarz-

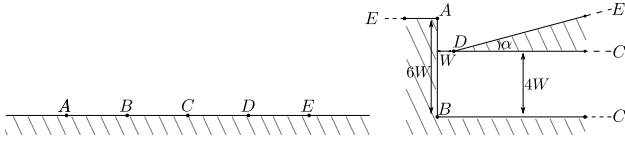


Figure 1. Transformation of the upper half plan (left) to the recorder window geometry (right) with the Schwarz-Christoffel transformation.

Christoffel transformation) are usual methods of fluid mechanics. It can be shown that they are equivalent for simple cases, even if the second can be used for more intricate geometry. The first actually corresponds to the theoretical study of a flow around an angle. The two last methods use the Finite Element Method (FEM) to solve either an incompressible or a compressible flow equation.

2.1 Incompressible flow around an angle: complex potential

For an incompressible potential flow, the velocity satisfies the two relations:

$$\nabla \cdot \mathbf{u} = 0 \text{ and } \nabla \times \mathbf{u} = 0. \quad (1)$$

A description of the two-dimensional flow is possible thanks to the potential ϕ and the streamfunction ψ defined as

$$\begin{cases} u = \frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y} \\ v = \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x} \end{cases}, \quad (2)$$

that automatically satisfy the condition in Eq. (1). The study of the flow can be reduced to the complex analysis of the complex potential f defined as:

$$f = \phi + i\psi. \quad (3)$$

The complex velocity $w = u - iv$ is then given by

$$w = \frac{df}{dz}. \quad (4)$$

The complex analysis of the singularities of f yields to well known flows. Among others, the case [7]

$$f = z^n \quad (5)$$

has been identified as the flow around a angle $\alpha = \pi(2 - 1/n)$. The characteristics of the flow are:

$$\begin{cases} \phi = r^n \cos n\theta \\ \psi = r^n \sin n\theta \end{cases} \text{ or } \begin{cases} u = nr^{n-1} \cos(n-1)\theta \\ v = -nr^{n-1} \sin(n-1)\theta \end{cases}, \quad (6)$$

where r and θ are the cylindrical coordinates in a referential where the tip of the angle is at $r = 0$. Note that for angles α smaller than π the velocity diverges while approaching the tip of the angle.

2.2 Incompressible flow through the recorder window: Schwarz-Christoffel transformation

Another way to obtain a two dimensional flow in a given geometry under potential assumptions is to use the Schwarz-Christoffel transformation of the upper half complex plane

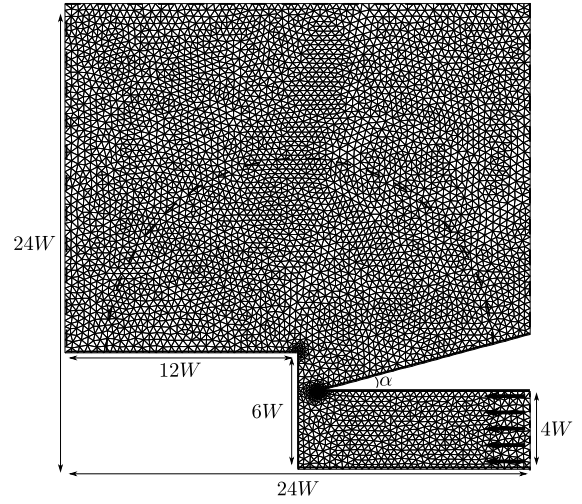


Figure 2. Domain of computation for the Finite Element Method. Solid lines are walls. Outer dashed line are out-flow boundary. The half circle in dashed line represents the beginning of the PML condition for the compressible case. The lower dashed line is the inflow condition. The mesh generated by the software *FreeFem++* is refined near the tip of the labium with an adaptive algorithm whose error criteria is based on the modulus of the velocity.

into a given geometry. The transformation used in this paper is illustrated on figure 1. The computation of the transformation is made thanks to the numerical toolbox developed by Driscoll [8]. The conformal mapping allows to obtain the streamlines at any required position, i.e. it allows to obtain the streamfunction ψ and the velocity \mathbf{u} .

This method has already been applied to study jet receptivity in a recorder [9] and *discrete-vortex* model [6], although it was used in a ideal case with an infinite plate as the labium and no wall facing it.

2.3 Incompressible flow through the recorder window: FEM

It is also possible to solve the flow around the labium by direct numerical computation. The incompressibility equation $\nabla \cdot \mathbf{u} = 0$ is rewritten:

$$\Delta \psi = 0, \quad (7)$$

where Δ is the Laplacian operator and ψ the streamfunction defined in Eq. (2). This equation is solved with the Finite Element Method (FEM) whose only difficulties lie in handling the mesh and the boundary conditions. The domain of computation is shown on figure 2. The constant velocity u_0 on the inflow boundary yields to the condition:

$$\psi = u_0 y. \quad (8)$$

The normal velocity on the wall is zero, which leads to the condition:

$$\begin{cases} \psi = 0 & \text{for the bottom} \\ \psi = 4W u_0 & \text{for the labium} \end{cases}, \quad (9)$$

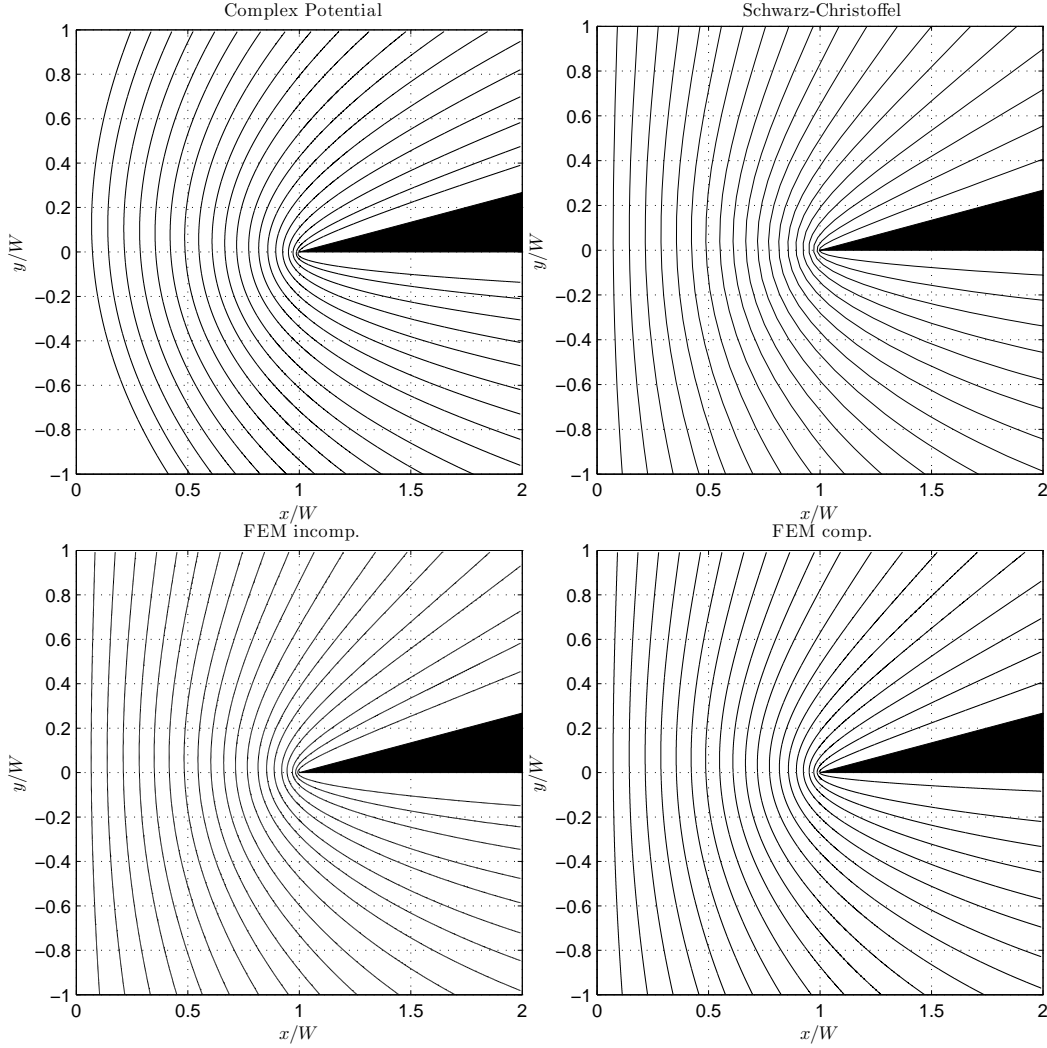


Figure 3. Streamlines computed with the different methods. Except for the complex potential method, the vertical wall stands at $x = 0$ and the bottom wall at $y = -4W$). The Mach number for the FEM computation is $M \approx 10^{-4}$. The Helmholtz number of the compressible FEM computation is $He = 0.004$.

accordingly to the inflow condition (see Eq. (8)). The outflow condition is handled as follow:

$$\psi = u_0 4W \frac{\theta_r - \pi}{\theta_0 - \pi} \quad (10)$$

where the angle $\theta_r = \tan^{-1}((y - y_0)/(x - x_0))$ is the angle of the coordinate on the boundary with respect to the center $(x_0, y_0) = (W, 0)$. This corresponds to the incompressible outflow of a source ($f(z) = \log(z)$), located at the tip of the labium. The mesh is generated by the same software used to solve the FEM: *FreeFem++* [10]. The mesh is automatically refined near the region of interest –as shown on figure 2– with an adaptive algorithm [11] whose error criteria is based on the modulus of the velocity.

2.4 Compressible flow through the recorder window: FEM

Even if the incompressible assumption is widely justified since the window W is much smaller than the wavelength at low frequencies, it is worth computing the compressible flow around the labium. Besides, the FEM requires no

more computation cost than for the incompressible case.

The Helmholtz equation on the pressure p

$$\Delta p + k^2 p = 0 \quad (11)$$

is solved for one wave number $k = 2\pi$. The same domain and mesh as for the incompressible case are used. The boundary condition are now:

- walls: $\frac{\partial p}{\partial n} = 0$, with n the normal direction
- inflow: $p = 1$
- outflow: to ensure the outgoing wave with no reflection, a Perfectly Matched Layer (PML) is used [12].

The PML consists in artificially adding damping while the wave approaches the boundary. A new wave number is defined as

$$k' = k(1 + i\epsilon), \quad (12)$$

where ϵ is a small control parameter which is zero to describe normal propagation and non-zero (~ 0.1) to describe the damped propagation. The subsequent outflow condition is written:

$$\frac{\partial p}{\partial n} = -ik(1 + i\epsilon)p, \quad (13)$$

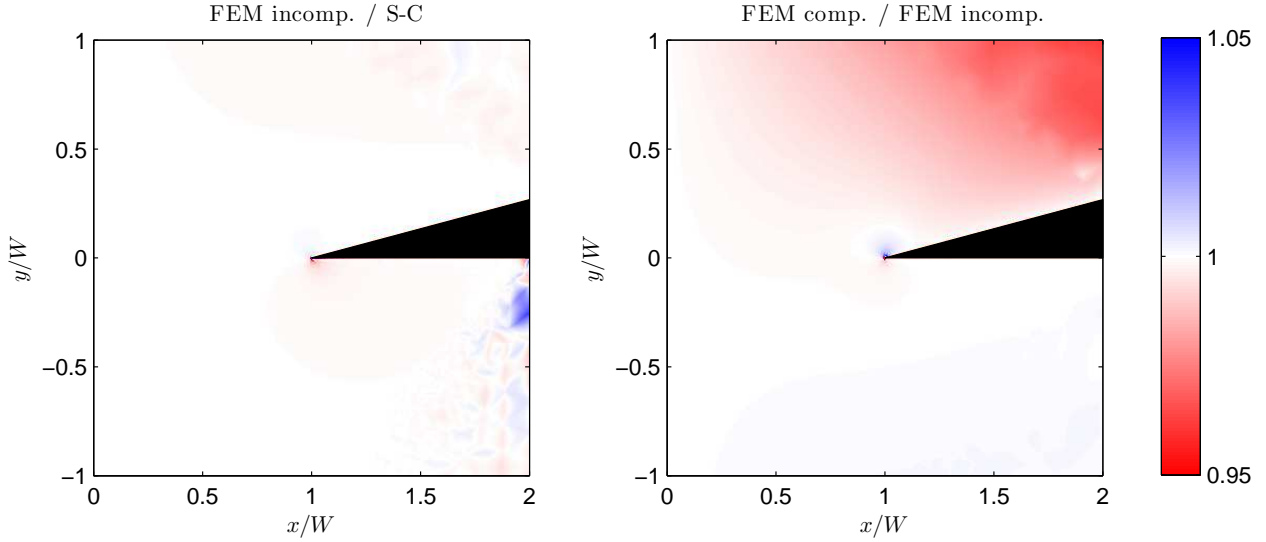


Figure 4. Ratio of the angle of the velocity field θ_i/θ_j , where $\theta = \arctan(v/u)$ and where the indices are relative to the methods (either Schwarz-Christoffel transformation or incompressible FEM or compressible FEM). Left: FEM incompressible over Schwarz-Christoffel. Right: FEM compressible over FEM incompressible.

corresponding to the damped wave propagation through the boundary of normal direction n . The acoustic velocity is then deduced from the pressure field and Euler's equation:

$$\begin{cases} u' = \frac{ik}{\rho_0 c_0} \frac{\partial p}{\partial x} \\ v' = \frac{ik}{\rho_0 c_0} \frac{\partial p}{\partial y} \end{cases}, \quad (14)$$

where ρ_0 is the air density and c_0 the speed of sound in the air.

3. COMPARISON OF THE METHODS

Since the complex potential and the Schwarz-Christoffel methods required a very basic geometry, the four previous methods are compared within the same (and ideal) configuration, i.e. with a sharp labium.

Figure 3 shows the streamlines (isovalue of the stream function ψ) around the sharp labium for the four methods. All the methods provide the same global trend of the flow around the labium. The complex potential does not account for the walls and the streamlines are not bended as it is the case for the other methods.

Every numerical method diverge while approaching the tip of the labium. Thus comparison is made impossible at this precise location. However, within the digit precision allowed by the computer, the power law of the divergence for the three numerical methods has been check to be as expected by the the complex potential. This point will be discussed later in section 4.

Rather than the absolute value of the velocity, the pertinent information is the direction of the velocity. This can be described by the angle θ that accounts for the ratio of the y-component over the x-component of the velocity through

$$\theta = \arctan \frac{v}{u}. \quad (15)$$

This scalar definition allows to compare the field in a two dimensional plot (see figure 4). The Schwarz-Christoffel

transformation and the incompressible FEM give almost the same results, at least near the labium. This gives support to the FEM method.

Then, the comparison between incompressible and compressible is made with the FEM method. The incompressibility is characterized by the Mach number $M = 4u_0/c$ and the Helmholtz number $He = W/\lambda = fW/c$. In both cases the inflow velocity $u_0 \approx 0.01\text{m/s}$ yields a Mach number $M \approx 10^{-4}$. The compressible case is computed with a “large” wavelength ($k = 2\pi$) yielding a Helmholtz number $He = 0.004$. For a soprano recorder, the Helmholtz number would remain under the value corresponding to the highest note (D7, 2350Hz): $He \approx 0.027$. The two methods show some little discrepancies in the upper area, i.e. the outward area. This can be due to the hypothesis made on the outflow condition in Eqs. (10) and (13) for the incompressible and compressible cases, respectively. However, this does not exclude that the incompressible outflow and the compressible radiation behave differently within the outward area. This is interpreted as the visible difference between incompressible and compressible flow.

Besides, it must not be forgotten that the flow is assumed incompressible since W is much smaller than the wavelength, but an oscillating flow is expected to occur on the streamline computed under this assumption.

4. APPLICATION TO A MORE REALISTIC CASE

The FEM method allows the previous and idealized case of the sharp labium to be extended to the more realistic case of a round labium. The previous configuration is modified by introducing a curvature R at the tip of the labium as sketched on figure 5. Results for this new configuration are discussed in the two following sections in which both parameters R and α are varied, respectively.

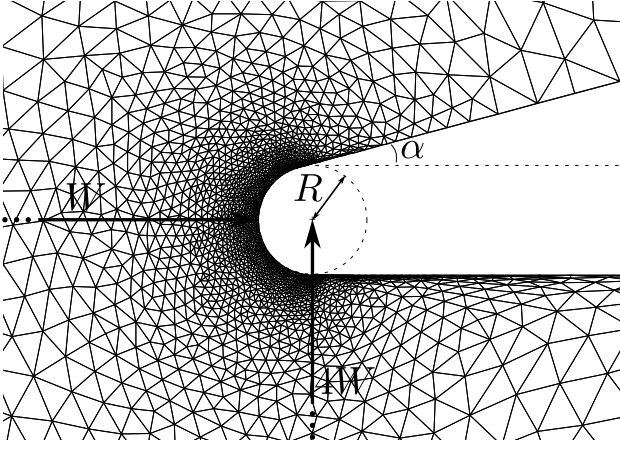


Figure 5. Modification of the tip of the labium to a rounded tip.

4.1 Sharpness of the labium

The radius R has been varied within a wide range of values from $R = W/1000$ to $R = W/10$, for one angle $\alpha = 15^\circ$. The divergence law of the velocity within the window is compared with the complex potential case (see Eq. (6)) on figure 6.

For all the value of R , the velocity shows the same trend according to three distinct areas: close to the wall ($x \lesssim W/2$), close to the labium ($x \gtrsim R$) and between these two areas. The middle area shows the same power law as expected by the complex potential method. When going from the wall to the middle area, the y-component of the velocity v goes from a non zero and constant value to the expected behaviour. When getting closer to the labium, the velocity stops increasing to reach a finite value: the velocity does no longer diverge and this effect starts at a distance $\sim R$ from the tip of the labium.

Besides stabilizing an unrealistic and diverging case, the

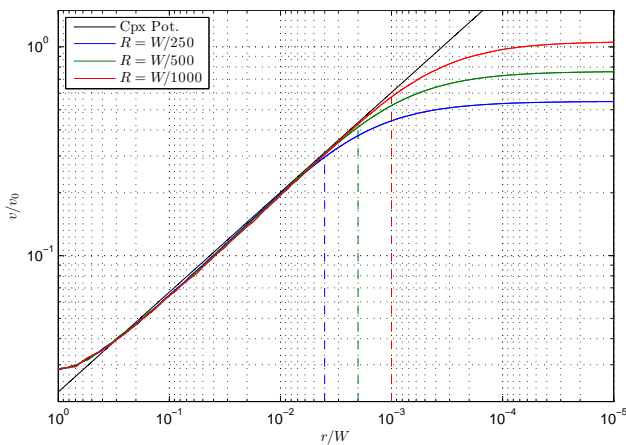


Figure 6. y -component of the velocity v in the window ($y = 0$) versus dimensionless radius r/W for $\alpha = 15^\circ$ and for different values of R . The radius r is defined as $r = W - x$. Note that the r axis is reversed in order to match to the orientation of the other figures. Vertical dashed lines correspond to the different radii.

curvature triggers a slower increase of the velocity while approaching the labium. The bigger the curvature R , the larger the distance at which the velocity stops increasing and the lower the amplitude of the velocity at an arbitrarily close distance to the labium.

4.2 Angle of the labium

The angle α has been varied within a wide range of values from $\alpha = 0^\circ$ to $\alpha = 60^\circ$, for one radius $R = W/500$. The divergence law of the velocity within the window is compared with the complex potential case (see Eq. (6)) on figure 7.

Results are similar to those discussed in the previous section in terms of the global trend according to the three specific areas. As expected, only the value of the divergence exponent n in Eq. (6)) is modified by the angle of the labium: the slope of the velocity v in the $(\log r, \log v)$ plane depends on α only.

Thus, at a distance arbitrarily close to the labium, for a same inflow u_0 the y -component of the velocity v increases as the angle decreases.

5. DISCUSSION AND CONCLUSION

Different methods to study the acoustic flow around the labium of a recorder have been compared for one simple case. The use of incompressible flow methods is justified at low frequencies since the distance W of the window is much smaller than the acoustic wavelength. The Finite Element Method (FEM) is validated in comparison to the complex potential method, in the case of an incompressible flow. The assumption of having an incompressible flow within the window is checked using the FEM: compressible and incompressible FEM give almost the same results. Little discrepancies arise within the outward area. This is interpreted as an effect of the compressibility.

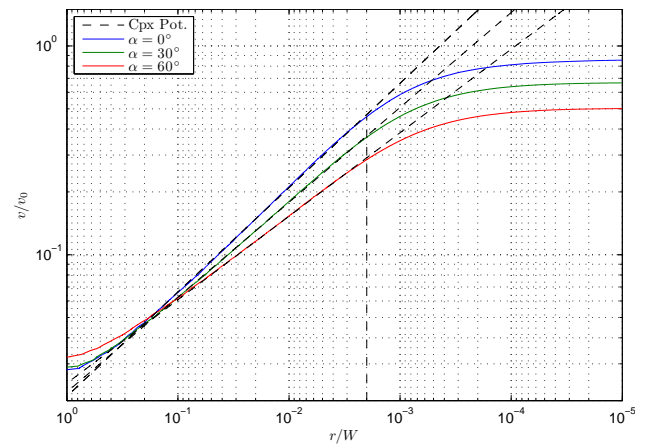


Figure 7. y -component of the velocity v in the window ($y = 0$) versus dimensionless radius r/W for $R = W/500$ and for different values of α . The radius r is defined as $r = W - x$. Note that the r axis is reversed in order to match to the orientation of the other figures. The vertical dashed line correspond to the radius $R = W/500$.

The FEM method presented in this paper is a little more useful than the previous Schwarz-Christoffel transformations made by other authors on the same issue, since it can be applied to more realistic configurations. However, the present study only provides solutions for the potential component of the flow and reaches its limit when the whole flow might be required. This is the case for aeroacoustic analogies, such as Lighthill's one, where the non potential component of the flow is interpreted as sources for the potential component. Besides, other phenomena are also not caught by the potential description: neglecting the viscosity is an other important limitation. Viscous effects are though dominant at the singularity of the potential flow, at the tip of the labium. When the gradient of the velocity increases, boundary layer might grow due to the viscosity and the (acoustic) flow might shed. This has been observed in flute-like configuration, for high amplitude of oscillation [4]. This is usually described as the formation of a free jet that occurs every half period [1, 13] and dissipates energy. Despite all these limitations, the study of the potential flow around the labium still provides usable results for more complex studies that would include both potential and non potential descriptions, and it provides some insights about the point discussed above.

The FEM confirms intuitive results about the growth of the velocity near the labium. When far enough from the wall, the velocity grows as expected for an ideal case. When close enough to the labium (at approximately one radius of curvature of the labium), the velocity ceases growing. It provides quantitative behaviours of the y-component of the velocity v with respect to both the sharpness and the angle of the labium: v increases with the sharpness and with a decrease of the angle. A higher y-component velocity is expected to trigger non-linear effects sooner. From there onwards, it may be possible to find a criteria based on other studies [14] to link the present linear description to non-linear phenomena. This may find application in reducing the trigger of this non-linear effect that is known to be a limiting factor in the growth of the amplitude of oscillation [4] and thus of the acoustic power.

These results come from a numerical computation. It would be interesting however to compare them to an analytical solution of the flow. The complex potential corresponding to the more realistic geometry (walls, labium with round edge) might be difficult to find. However, the present Schwarz-Christoffel transformation that already acknowledges for the walls can be adapted to localized round edges [15]. This would provide an analytical framework from which the results about the growth of the velocity near the edge should be confirmed.

The shape (angle and sharpness) of the labium has never been investigated through modeling. Dequand *et. al* proposed a *discrete-vortex* model that can include the orientation of the acoustic field near the labium. In such modeling, the sound production is ensured by the interaction of discrete vortices with the acoustic streamlines. It is common to consider the vortices close to the labium only: they have a greater contribution than far ones, since the acoustic velocity is greater near the tip of the labium. From the present

study, this assumption can be refined by considering vortices in an area whose characteristic length is of same order than the curvature radius of the labium. The *discrete-vortex* approach is a poor approximation of the flow so that an accurate estimation of the acoustic flow is an overkill when combining with these models. However, the combination would still provide a first tool to study the effect of the angle and/or the sharpness on the sound production.

6. REFERENCES

- [1] U. Ingard and H. Ising, "Acoustic Nonlinearity of an Orifice," *J. Acoust. Soc. Amer.*, vol. 42, no. 1, pp. 6–17, 1967.
- [2] M. S. Howe, "Contributions to the theory of aerodynamic sound, with application to excess jet noise and the theory of the flute," *Journal of Fluid Mechanics*, vol. 71, pp. 625–673, 9 1975.
- [3] —, "The dissipation of sound at an edge," *J. Sound Vib.*, vol. 70, pp. 407–411, 1980.
- [4] B. Fabre, A. Hirschberg, and A. P. J. Wijnands, "Vortex Shedding in Steady Oscillation of a Flue Organ Pipe," *Acust. Acta Acust.*, vol. 82, pp. 863 – 877, 1996.
- [5] P. A. Nelson, N. A. Halliwell, and P. E. Doak, "Fluid Dynamics of a Flow Excited Resonance, Part II: Flow Acoustic interaction," *J. Sound Vib.*, vol. 91, no. 3, pp. 375–402, 1983.
- [6] S. Dequand, J. F. H. Willems, M. Leroux, R. Vullings, M. Van Weert, C. Thieulot, and A. Hirschberg, "Simplified models of flue instruments : Influence of mouth geometry on the sound source," *J. Acoust. Soc. Amer.*, vol. 113, no. 3, pp. 1724–1735, 2003.
- [7] J. Darrozes and C. Francois, *Mecanique des Fluides Incompressibles*. Springer, 1982.
- [8] T. A. Driscoll, "Algorithm 756; a MATLAB toolbox for Schwarz-Christoffel mapping," *ACM Transactions on Mathematical Software*, vol. 22, pp. 168–186, Jun. 1996.
- [9] M. Verge, R. Caussé, B. Fabre, A. Hirschberg, A. Wijnands, and A. van Steenberg, "Jet oscillations and jet drive in recorder-like instruments," *Acta acustica*, vol. 2, no. 5, pp. 403–419, 1994.
- [10] F. Hecht, O. Pironneau, K. Ohtsuka, and A. L. Hyaric, <http://www.freefem.org/ff++>, 2013.
- [11] F. Hecht, "The mesh adapting software: bamg," *INRIA report*, 1998.
- [12] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Physics*, vol. 114, pp. 185–200, Oct. 1994.
- [13] J. H. M. Disselhorst and L. V. Wijngaarden, "Flow in the Exit of Open Pipes During Acoustic Resonance," *J. Fluid Mech.*, vol. 99, pp. 293–319, 1980.

- [14] J. M. Buick, M. Atig, D. J. Skulina, D. M. Campbell, J.-P. Dalmont, and J. Gilbert, "Investigation of non-linear acoustic losses at the open end of a tube," *J. Acoust. Soc. Amer.*, vol. 129, no. 3, pp. 1261–1272, 2011.
- [15] P. Henrici, *Applied and Computational Complex Analysis Vol. I*. John Wiley, 1974.

COMPARISON OF TWO METHODS OF SOUND POWER MEASUREMENTS OF FLUE ORGAN PIPES

Judit Angster

Fraunhofer IBP, Stuttgart, D
Angster@ibp.fhg.de

Katrin Hoge

Fraunhofer IBP, Stuttgart, D

Andras Miklos

Steinbeis TC Applied Acoustics
AkustikOptik@t-online.de

ABSTRACT

The Fraunhofer IBP (Stuttgart, Germany) realized with 10 organ builder companies from 8 countries a research project, which was supported by the European Union and dealt with the matching of the pipe organ to the room. The aim was to assist organ builders already during the planning phase of an instrument. An important parameter for acoustic design is the radiated sound power of the instrument, and thus of the pipes. It was planned to find out a relatively simple method allowing the serial determination of the sound power of numerous pipes. With regard to the specific features of the sound of an open flue pipe the reverberation room procedure was selected from possible measurement methods to determine the sound power. Measurements were carried out to show the influence of wind pressure, pitch and geometry (scaling) of the flue organ pipes on the sound power. The measurements were significantly influenced by the strong interaction of the harmonic spectrum of the pipe sound with the eigenmodes of the reverberation room and by the coherent radiation of both point-like sound sources at the mouth and open end of the flue organ pipe. In a second measuring procedure the measurements were performed by a microphone-array system in the anechoic room. The results as well as the advantages and disadvantages of the measuring methods and a view of other possibilities for the detection of the sound power of flue organ pipes are discussed.

1. INTRODUCTION

For several years, the Research Group of Musical Acoustics at the Fraunhofer Institute for Building Physics in Stuttgart has been doing research in the field of organ acoustics [1-10]. Handcraft skills and long-time experience are the necessary prerequisites to manufacture, preserve and restore these intricate and traditional instruments. And this is why still nowadays the majority of organ builders make use of this conventional and traditional knowledge and experience; there is hardly any scope for experiments within the narrow economic framework for SME (Small and Medium-sized Enterprises) in particular. Thus, researchers at the Fraunhofer IBP made it their business to examine the various components of the complete system of the 'pipe organ' by means of their scientific and technical know-how. The objectives and requirements of organ builders and researchers are discussed in close cooperation with organ builder companies throughout Europe. The topics comprise the wide

range from the wind system of the organ to the physics of the organ pipe. One objective is certainly the profound understanding of the physical processes during sound generation, but great importance is also attached to the transformation of the scientific knowledge with practical relevance. As a result, novel wind systems for the organ [11] as well as scaling and analysis software were developed within the framework of several projects funded by the EU in the last few years [12-14]. Moreover, training for organ builders was offered to learn and get acquainted to some theoretical details of organ building.

In recent times, there was a rising demand for research on the matching of the organ to the room, where it was intended to be finally played. Pipe organs are no longer only set up in naves, but also in very large multipurpose facilities and small community centres. The matching of the pipe organ is urgent, as size and furniture as well as the architectural design have an influence on the acoustics of a room and make various demands on the organ.

The influencing aspects of room acoustics and the effects on the pipe organ have been investigated in a research project (EUCRAFT 017712 – Innopipeorg' Innovative Design Method for Matching the Pipe Organ to the Acoustics of the Room') in cooperation with European organ builder companies, and methods are developed to pre-match the organ to the room [15]. The organ builders are directly involved in the European research projects. They build and deliver the pipes to the laboratory investigations and participate in the measurements by carrying out voicing adjustments in the laboratory.

With regard to the various dimensions of the sites the radiated sound power of the organ is of specific interest as well as the problem which physical parameters influence the sound power. Although some basic references on the sound radiation of pipes exist [16-19] only one paper deals with the measurements of the sound power of organ flue pipes [20]. Nevertheless in this publication only two pipes have been tested (one open and one closed organ pipe) which were placed in the wall between two reverberation rooms. On one hand no detailed description of the pipes is given in the paper, on the other hand the method deviated significantly from the standardized reverberation room method [21]. The sound source (the pipe opening) was placed very near to the wall of the test room and therefore only one source position was applied. So the sound power radiated by different organ pipes has not yet been investigated up to now, and no

verifications of the theoretically calculated assessments by means of measuring the sound power do exist.

Two measurement methods to determine the sound power of flue organ pipes are discussed in this paper. The problems of the already existing methods are demonstrated by measurements, and possible solutions of reliable and time-saving measurement methods are described. The measurements give information on the influence of the geometry of the pipe and the behaviour of the radiated sound power in dependence of the wind pressure¹. In addition, particular attention is given to the suitability of the measurement method for the flue organ pipe. Moreover, a relatively simple method is proposed allowing the serial determination of the sound power of numerous pipes.

2. THE FLUE ORGAN PIPE

To generate a large variety of sounds, the design of the pipe varies for example in material, shape of the resonator or opening of the resonator end (open, stopped or partly stopped (chimney pipe)): This paper investigates open and stopped metal flue organ pipes.

Figure 1 shows the structure of an open flue organ pipe, which basically consists of three different parts: the *pipe foot* leading in the compressed air, the *mouth* generating the sound and labium, flue and cut-up are part of it, and finally the *pipe body* or *pipe resonator*, where the air column oscillates.

The sound of a pipe is generated, because the air column inside the resonator begins to oscillate and standing waves develop with constant air supply. These pressure oscillations are radiated to the ambient air by the openings of the pipe at the resonator end and at the labium [2].

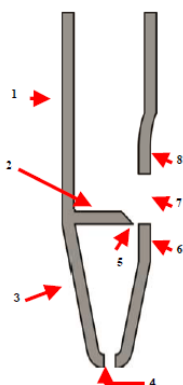


Figure 1. Structure of a flue organ pipe 1. resonator, 2. languid, 3. pipe foot, 4. foot hole, 5. flue, 6. lower lip, 7. mouth (cut-up), 8. upper lip

3. MEASUREMENTS

3.1 Measurement according to the reverberation room procedure

For determining the sound power of flue organ pipes first the standardized reverberation room procedure was chosen [21]. 6 source positions and 8 different rotating microphone paths ($r = 1.2\text{m}$) were applied. Six different flue pipes were selected for the measurement in the reverberation room; two-two open pipes of a diapason and a flute stop, a string pipe and a stopped pipe. The most important scaling parameters are listed in Table 1 and Table 2 and a photo of three pipes is shown in Figure 2. It must be mentioned that the influence of these parameters on the sound power of flue pipes have not been investigated in this project. Particular attention is given to find a relatively simple method allowing the serial determination of the sound power of numerous pipes.

Nevertheless, the effect of the wind pressure on the sound power is investigated for the case when the wind pressure is changed but other parameters of the flue pipe stay unchanged. The measurement is carried out with various wind pressures (50 to 90mm water column). The pipes were voiced at the nominal wind pressure of 70mm water column.

	Parameter		
	stop	diameter (D_i) [mm]	fundamental frequency [Hz]
R5_low	flute	76,0	196
R4_low	diapason	63,7	196
M_low	Viola da Gamba (string stop)	38,4	196
R5_high	flute	43,3	470
R4_high	diapason	34,2	470
G_low	stopped	67,0	196

Table 1. Pipes selected for measurement

pipe	L_p [mm]	D_i [mm]	h [mm]	b [mm]	D_F [mm]	d_f [mm]
R5_low	682	76	14	47.2	11.7	0.6
R4_low	724	63.7	12.2	49.9	9.2	0.6
M_low	720	38.4	7.7	30.4	4.5	0.75
R5_high	281	43.3	7.2	27.3	8.1	0.3
R4_high	312	34.2	6.8	26.8	7.5	0.4

Table 2. Scaling of the pipes selected for measurement
 L_p - pitch length, D_i - inner diameter, h - cut-up height, b - mouth width, D_F - foot hole diameter, d_f - flue width

¹ Wind pressure denominates the operating pressure of the organ. It is indicated in millimeter water column (10mm water \approx 100Pa).



Figure 2. Pipes for measurement: R5_high, G_low and R4_high

Four examples of measured sound power spectra are shown in the following figures: Figure 3 (**pipe G_low (stopped pipe)**), Figure 4 (**R5_high**), Figure 5 (**R5_low**) and Figure 6 (**R4_low**). A summary of the measurement results is given in Table 3.

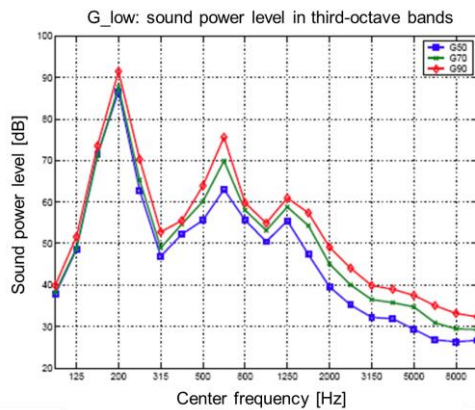


Figure 3. G_low (196Hz): Sound power level in 1/3 octave bands as a function of the wind pressure (measurement in the reverberation room)

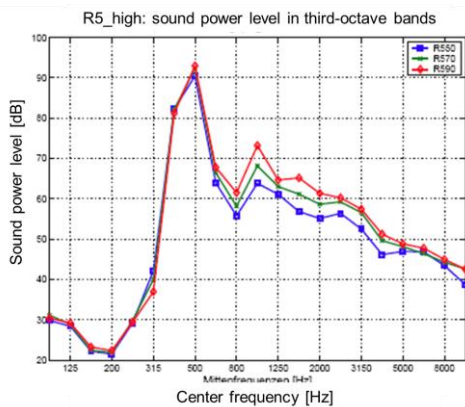


Figure 4. R5_high (470Hz): Sound power level in 1/3 octave bands as a function of the wind pressure (measurement in the reverberation room)

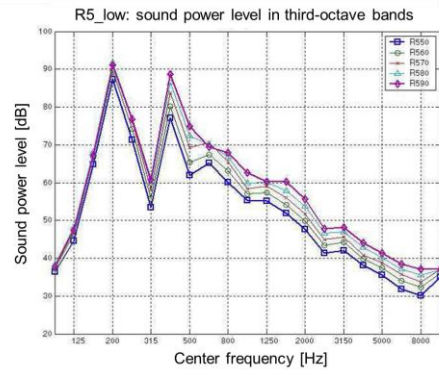


Figure 5. R5_low (196 Hz): Sound power level in 1/3 octave bands as a function of the wind pressure (measurement in the reverberation room)

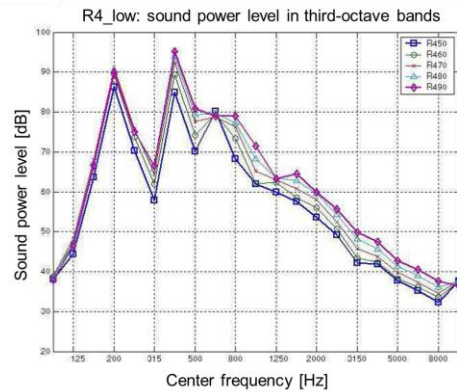


Figure 6. R4_low (196 Hz): Sound power level in 1/3 octave bands as a function of the wind pressure (measurement in the reverberation room)

	wind pressure [mm water]		
	50	70	90
R5_low	88,0 dB	92,0 dB	93,0 dB
R4_low	89,4 dB	95,1 dB	96,7 dB
M_low	84,0 dB	88,6 dB	89,4 dB
R5_high	91,1 dB	92,8 dB	93,3 dB
R4_high	89,8 dB	92,3 dB	94,6 dB
G	86,8 dB	88,9 dB	91,7 dB

Table 3. Overall sound power levels in the reverberation room

The presented sound power level spectra and the overall sound power levels determined according to the ISO 3741 standard show that an increase in wind pressure results in a higher sound power.

To assess the suitability of the measurement for flue organ pipes, it is useful to consider the standard deviation of the sound pressure level. Figure 7 shows the standard deviation for the investigated pipes after averaging by all source and microphone positions and for a wind pressure of 50mm water. The black line marks the maximum allowable deviation. Whereas the pipes of a higher frequency comply with these conditions, the deviation of the lower pipes is very high around the significant fundamental in particular. This fact is primarily due to the low

number of eigenmodes excited by low frequencies as well as to the tonal spectrum of the pipes. Additional source and microphone positions do not yield any improvement. Furthermore, open flue organ pipes approximately radiate like two dependent point sources developing an interference field, which also influences the accuracy of measurements [22]. Figure 8 shows the standard deviation for pipe **G_low** after averaging by all source and microphone positions at three measured wind pressure values.

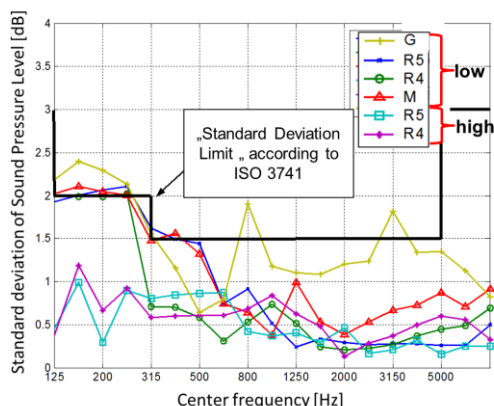


Figure 7. Comparison of the standard deviation of the sound pressure level of all pipes at 50mmwater after averaging by all source and microphone positions

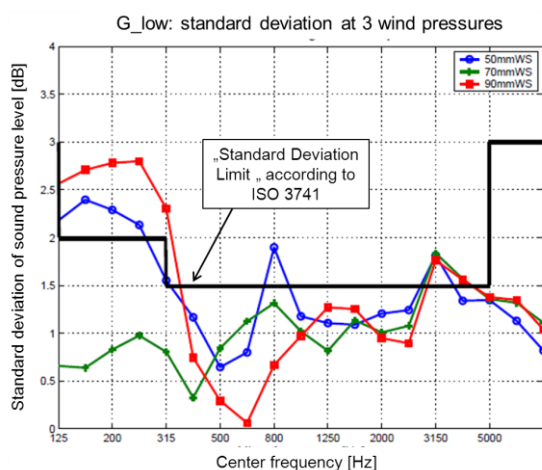


Figure 8. Comparison of the standard deviation of the sound pressure level of pipe **G_low** at three wind pressure values after averaging by all source and microphone positions

3.2 Microphone array measurements in the anechoic room

Measurements in the reverberation room are very time-consuming and labour-intensive due to the large number of source and microphone positions. Moreover, they do not provide any satisfactory accuracy of measurements for the low frequencies. Therefore, the possible determination of sound power by means of a microphone array sound intensity measurement in the anechoic room is taken into consideration. The array consists of 96 microphones arranged in a rectangle for the respective frequency range. The sound intensities are calculated by means

of SONAH (statistically optimized near-field acoustical holography) [23,24]. Afterwards, the sound power is evaluated by means of an area selection tool in the software. Two pipes, R4_low and G, are selected for this measurement. In a first set-up, two quadratic envelopes, open at the top and at the bottom, are put around the mouth and the open resonator end (only R4_low) and then the sound power is calculated. During a second measurement the significance of a single measurement directly in front of the mouth is assumed (Figure 9). This approximation of the sound power should simplify the input in case of serial measurements. The results are listed in Table 4.

As expected, the sound power averaged over the enveloping surfaces is somewhat higher than the power measured in front of the mouth. However, the difference of about 1.5-1.8 dB for R4_low and 1.1-1.2 dB for G_low is comparable with the allowed tolerance of ISO 3741.

The sound power levels of the enveloping surfaces of R4_low comply with the values of the measurements in the reverberation room with the exception of the value for 50mmWS. At this wind pressure the sound power level of G_low in the reverberation room is 5.7 dB less than at 70 mmWater, while the level decrease between 90 mmWater and 70 mmWater is only 1.6 dB. The sound level spectrum in Fig. 5 also shows more than 5 dB level drop in the 400 Hz 1/3-octave band. Since the overall sound power level is strongly influenced by the highest peak of the spectrum, the decrease of the 400Hz peak may explain the observed decrease of the overall sound power level.

Both measurements by microphone array show good compliance for G_low. The influence of the directional characteristics seems to make out about 1 dB difference. In comparison to the values measured in the reverberation room the calculated sound powers in this case are clearly higher. Significant measurement inaccuracies of the reverberation room procedure are certainly responsible for that fact (see explanation also in ‘Summary and Conclusions’).

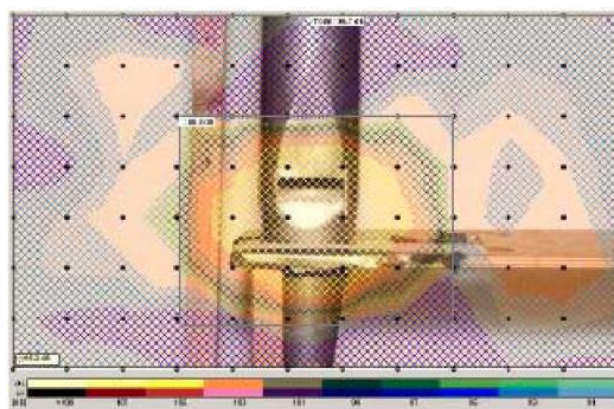


Figure 9. Characteristic calculation result of the sound power measured in front of the mouth of a stopped pipe

the

Pipe	Method	wind pressure [mm water]		
		50	70	90
R4_low	envelope	93.1 dB	95.0 dB	96.3 dB
	mouth	91.3 dB	93.3 dB	94.8 dB
	Rev. room	89.4 dB	95.1 dB	96.7 dB
G_low	envelope	89.4 dB	91.2 dB	92.3 dB
	mouth	88.2 dB	90.1 dB	91.2 dB
	Rev. room	86.8 dB	88.0 dB	91.7 dB

Table 4. Overall sound power levels measured by the microphone array and by the reverberation room method

The most important advantage of the microphone array method is indeed an enormous saving of time and the possibility of graphical processing of the data and results.

4. SUMMARY AND CONCLUSIONS

The matching of a pipe organ to the acoustic properties and geometric dimensions of the room, where it is planned to be played, is significant. The great variety of room volumes and furniture of various sites require the matching of the loudness and defined sound characteristics of the organ. Performing the matching afterwards is very difficult and expensive.

The sound power of an organ as well as of the individual pipes is one important parameter for this kind of matching. This property was investigated by three measurement methods and the dependence of the sound power on the wind pressure were determined. It was obvious that the increase of the wind pressure resulted in an increased radiated sound power of the pipe. Thus it is possible to save costs for materials by an increased wind pressure in dimensioning the organ.

Two different methods to determine the sound power were described: the reverberation room procedure according to [21] and the measurement of sound intensity by means of a microphone array. Both methods have some advantages and some disadvantages.

The reverberation room procedure is not suited for coherent sources. Since open flue pipes may be represented as two coherent simple sound sources, radiating in-phase for odd partials and out-of-phase for even partials, it is obvious that an interference field develops in the room. Additional problems occur with the development of a diffuse field in the reverberation room: The pipe sound is narrow-band and harmonic and thus dependent on the essential mixing by the reflections and superimpositions of the eigenmodes. However, if the rotating microphone is positioned in an unfavourable place, the narrow frequency bands of the pipe sound may not be sufficiently mixed.

Moreover, it is necessary to consider the measurable lowest frequencies in the reverberation room, because the measurements of the pipes with a low fundamental frequency showed high measurement error, whereas the pipes with a fundamental frequency of approx. 470 Hz complied with the requirements of the standard for measurement error.

The prerequisites for the development of a diffuse field are

- the existence of sound energy components at many different frequencies,
- the excitation of many eigenfrequencies of the room, and good mixing of the standing wave pattern,
- the more eigenfrequencies exist and are excited in the observed frequency range Δf , the more diffuse is the sound field.

This means, there is a lower cut-off frequency, where a diffuse field can barely develop in the reverberation room with volume V . According to Meyer and Kuttruff [25] this frequency is $f_{gr} \sqrt[3]{V} = 700 \dots 1000$. The cut-off frequency for the reverberation room of the Fraunhofer IBP with a volume of $V=392\text{m}^3$ is approx. between 96Hz and 137Hz for broadband noise. The results presented here show that the cut-off frequency must be above 200 Hz for narrow-band excitations by organ pipes.

To determine the sound power by means of a microphone array is a novel approach. But due to the lack of knowledge and experience so far, procedures and calculation methods cannot be assessed. The most important advantage is an enormous saving of time and the possibility of graphical processing of the data and results.

A future perspective to determine the sound power of flue organ pipes quickly and in a reliable way is to carry out a microphone array measurement in the case of closed organ pipes at the pipe mouth and in the case of open pipes once at the pipe mouths and once against the open end. As it is technically not easy to place the microphone array above the pipe, the pipe should be placed horizontally into the anechoic chamber with conducting the wind to the pipe foot by means of a wind duct.

It is worth to mention yet that from the point of view of room acoustic design the radiated power of the source is important. Through the sound power level of the source, the reverberation time and the volume of the room the sound pressure level at the listeners can be determined. Certainly the properties of the hearing of human being should be also taken into account; less acoustic power is needed in frequency domains where the human hearing is more sensitive and vice versa.

Naturally the sound power radiation of pipe organs and the direct application of results by organ builders is a very complicated issue and with the presented results only the first step has been made in the investigation of this topic.

Acknowledgments

This research was supported by the European Commission in the frame of a CRAFT (Cooperative Research Action for Technology) research project (Contract No: EUCRAFT 017712) and by the following European organ builder firms:

Werkstätte für Orgelbau Mühleisen, Leonberg, Germany, Manufacture d'orgues Muhleisen, Strasbourg, France, Flentrop Orgelbouw, Zaandam, Netherlands, Orgelbau Schumacher, Baelen, Belgium, Blancafort, Orgueners de

Montserrat, Collbato, Spain, Oficina e Escola de Organasia, Esmoriz (Porto), Portugal, Fratelli Ruffatti Pipe organ Builders, Padova, Italy, Johannes Klais Orgelbau, Bonn, Germany, Pécsi Orgonaépítő Manufaktura KFT, Pécs, Hungary, Orgelmakerij Boogaard, Rijssen, Netherlands.

5. REFERENCES

- [1] J. Angster, A. Miklós: "Documentation of the sound of a historical pipe organ," *Applied Acoustics*. 46 (1995) 61-82
- [2] A. Miklos, J. Angster: "Properties of the Sound of Flue Organ Pipes," *Acustica united with acta acustica*, Vol. 86, S.611-622; S. Hirzel Verlag GmbH & Co., Stuttgart (2000)
- [3] A. Miklós, J. Angster, S. Pitsch, T.D. Rossing: "Reed vibration in lingual organ pipes without the resonators," *J. Acoust. Soc. Am.* 113, Issue 2, (2003), S. 1081-1091
- [4] A. Miklós, J. Angster, S. Pitsch and T.D. Rossing: "Interaction of reed and resonator by sound generation in a reed organ pipe," *J. Acoust.Soc.Amer.* No.119 (5), 3121-3129 (2006).
- [5] G. Paál, J. Angster, W. Garen & A. Miklós: "A combined LDA and flow-visualization study on flue organ pipes," *Experiments in Fluids*. 40, H.6: 825-835 (2006)
- [6] J. Angster, S. Pitsch, A. Miklós: "Design of new wind systems for pipe organs," *CD-Rom: Int. Symp. on Musical Acoustics*, ISMA2007, Barcelona, Spain, 3-S2-1
- [7] H. Außerlechner, T. Trommer, J. Angster and A. Miklós: "Experimental jet velocity and edge tone investigations on a foot model of an organ pipe," *J. Acoust. Soc. Amer.* 126 (2), August 2009 S. 878-886
- [8] T. Trommer, J. Angster and A. Miklós: "Roughness of organ pipe sound due to frequency comb," *J. Acoust. Soc. Amer.* Vol. 131, Issue 1, 739-748 (2012)
- [9] G.R. Plitnik, J. Angster: "The influence of pipe organ reed curvature on tone quality," *J. Acoust. Soc. Amer.* Vol. 132, Issue 5, 3502-3511 (2012)
- [10] P. Rucz, T. Trommer, J. Angster, A. Miklós, F. Augusztinovicz, "Sound design of chimney pipes by optimization of their resonators," *J.Acoust. Soc. Am.* 133 (1), January 2013 (529-537).
- [11] S. Pitch, "Entwicklung von neuartigen Windsystemen fuer Kirchenorgeln," PhD Thesis, Fraunhofer IBP, Stuttgart and University of Siegen (2005)
- [12] T. Wik, "Einschwingvorganganalyse von Orgel- und Gitarrenklängen mit Hilfe moderner Messtechniken," Diplom Thesis, Fraunhofer Institut für Bauphysik and University Stuttgart (2004)
- [13] C. Taesch, "Messung und Charakterisierung von Klängen der Lippenorgelpfeife, Diplom Thesis, Fraunhofer Institut für Bauphysik and University Stuttgart (2003)
- [14] <http://www.ibp.fraunhofer.de/en/Expertise/Acoustics/Musical-Acoustics.html>
- [15] Sixth Framework Program of the European Union (Framework Program on Research, Technological Development and Demonstration CRAFT) EU-CRAFT 017712 - INNOPIPEORG "Innovative Design Method for Matching the Pipe Organ to the Acoustics of the Room" Contract No: COOP-CT-2005-017712 (2005-2008)
- [16] M.P. Verge, B. Fabre, W.E. Mahu, A. Hirschberg, "Feedback excitation mechanism in organ pipes," *J. Acoust. Soc. Am.* 95 (1994), p. 1119-1132.
- [17] N.H. Fletcher, T.D. Rossing, "The physics of musical instruments," Springer-Verlag New York (1991).
- [18] N.H. Fletcher, "Sound production by organ flue pipes," *J. Acoust. Soc. Am.* 60 (1976), p. 926-936.
- [19] L.E. Kinsley, A.R. Frey, A.B. Coppens, and J.V. Sanders, "Fundamentals of acoustics," 4th ed., John Wiley & Sons, Hoboken, New York, etc., (2000)
- [20] G. Franz, H. Ising, P. Meinusch, „Schallabstrahlung von Orgelpfeifen“ *Acustica*, Vol. 22, S.226-231; S. Hirzel Verlag GmbH & Co., Stuttgart (1969/70)
- [21] DIN EN ISO 3741:1999: "Bestimmung der Schallleistungspegel von Geräuschquellen aus Schalldruckmessungen: Hallraumverfahren der Genauigkeitsklasse 1"
- [22] K.Hoge: "Schallleistungsbestimmung von Lippenorgelpfeifen: Messverfahren und Einflussfaktoren," Internship report, Fraunhofer IBP, Stuttgart, University of Dresden, 2006
- [23] J. HALD, "Non-stationary STSF", Brüel & Kjaer Technical Review, Naerum Denmark, (2000)
- [24] E. G. Williams, "Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography"; Academic Press, London (1999)
- [25] H. Kuttruff, "Room Acoustics", Elsevier Applied Science, London, UK. 3rd edition (1991)

Prediction of the dynamic oscillation threshold of a clarinet model: Comparison between analytical predictions and simulation results

B. Bergeot, A. Almeida, B. Gazengel

LUNAM Université, Université du Maine,
UMR CNRS 6613, Laboratoire d'Acoustique,
Le Mans, France
baptiste.bergeot@univ-lemans.fr

C. Vergez

Laboratoire de Mécanique et Acoustique,
LMA, CNRS UPR7051,
Marseille, France
vergez@lma.cnrs-mrs.fr

ABSTRACT

Simple models of clarinet instruments based on iterated maps have been used in the past to successfully estimate the threshold of oscillation of this instrument as a function of a constant blowing pressure. However, when the blowing pressure gradually increases through time, the oscillations appear at a much higher value than what is predicted in the static case. This is known as bifurcation delay, a phenomenon studied in [1] for a clarinet model. In numerical simulations the bifurcation delay showed a strong sensitivity to numerical precision.

This paper presents an analytical estimation of the bifurcation delay of the simplified clarinet model taking into account the numerical precision of the computer. The model is then shown to correctly predict the bifurcation delay in numerical simulations.

1. INTRODUCTION

The oscillation threshold of the clarinet has been extensively studied in the literature [2,3] assuming that the blowing pressure is constant. In this context, the *static* oscillation threshold γ_{st} is defined as the minimum value of the blowing pressure for which there is a periodic oscillating regime. This value of the threshold is obtained by applying a constant blowing pressure, allowing enough time to let the system reach a permanent regime (either static or strictly periodic), and repeating the procedure for other constant blowing pressures. Most studies using iterated maps are restricted to static cases, even if transients are observed. They focus on the asymptotic amplitude regardless of the history of the system and of the history of the control parameter.

A recent article [1] studied the behavior of a clarinet model when the blowing pressure increases linearly. The model starts its oscillations for a much higher value of the blowing pressure than the *static* oscillation threshold. An analytical expression of this dynamical threshold has been derived and its properties studied: the dynamic threshold does not depend on the increase rate of the blowing pres-

sure (for sufficiently low increase rates), but is very sensitive to the value of the blowing pressure at which the increase is started.

The article [1] ends with a comparison between the analytical predictions and numerical simulations (Fig. 10 in [1]), revealing an important sensitivity to the precision used in numerical simulations. Indeed, numerical results only converge towards theoretical ones when the model is computed with hundreds or thousands of digits (i.e. approximating an infinitely precise simulation). Otherwise, the observations of numerically simulated thresholds are far from the theoretical ones, they depend on the increase rate of the blowing pressure and are independent of the starting value of the blowing pressure. The conclusion is that theoretical results obtained in [1] cannot explain the behavior of the model simulated in the common double-precision of a modern CPU.

The aim of this paper is to explain and predict the start of the oscillations in simulations performed with the usual double-precision of computer CPUs (around 15 decimal digits). A model introduced in a journal article [1] is modified to predict the oscillation threshold in limited precisions. The expression of this oscillation threshold is given in section 3. In the same section, the theoretical thresholds (static and dynamic ignoring or taking into account the precision) are compared to the thresholds observed in numerical simulations. The influence of the speed at which the blowing pressure is increased is discussed, as well as that of the initial value of the blowing pressure. A similar analysis is made for the second control parameter of the model (related to opening of the embouchure at rest). The clarinet model and major results from [1] are first briefly recalled in section 2.

2. STATE OF THE ART

2.1 Clarinet Model

This model divides the instrument into two elements: the exciter and the resonator. The exciter is modeled by a nonlinear function F also called nonlinear characteristic of the exciter, which relates the pressure applied to the reed $p(t)$ to the flow $u(t)$ through its opening. The resonator (the bore of the instrument) is described by its reflection function $r(t)$. p and u are two non-dimensional state variables that are sufficient to describe the state of the instrument.

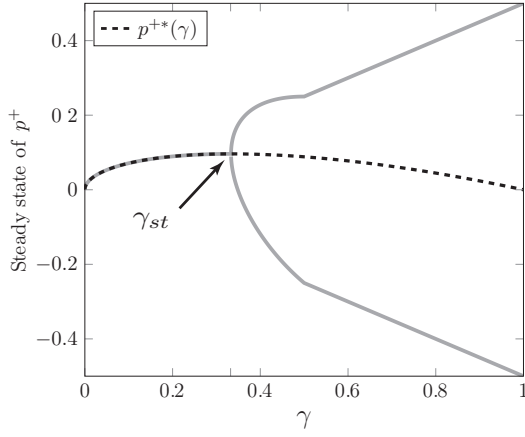


Figure 1. Graphical representation of the static bifurcation diagram for $\zeta = 0.5$. In gray, the stable solutions, in dashed black, the non-oscillating solution.

The solutions $p(t)$ and $u(t)$ depend on the control parameters: γ proportional to the mouth pressure P_m according to

$$\gamma = \frac{P_m}{P_M} = \frac{P_m}{kH} \quad (1)$$

where P_M represents the pressure needed to close the reed entrance (also used to normalize the pressure $p(t)$) a product of $1/k$ the acoustic compliance of the reed and H its distance to the lay at rest. The other parameter is ζ which is related to the opening of the embouchure at rest according to the formula

$$\zeta = Z_c U_A / P_M = Z_c w H \sqrt{\frac{2}{\rho P_M}}. \quad (2)$$

Here, Z_c is the characteristic impedance at the input of the bore, w the effective width of the reed, and U_A the maximum flow admitted by the reed valve. For most of the analysis below, this parameter is fixed at 0.5, a typical value observed in musicians, but the analysis can easily be reproduced for other values of ζ . The nonlinear characteristic is provided by the Bernoulli equation describing the flow in the reed channel [4, 5].

The model is extremely simplified by considering a straight resonator in which the eventual losses are independent of frequency. In the current work, losses are neglected in all calculations. The reed is considered as an ideal spring [3, 6–10]. With these assumptions, the reflection function becomes a simple delay with sign inversion. Using the variables p^+ and p^- (outgoing and incoming pressure waves respectively) instead of the variables p and u , the system can be simply described by an iterated map [6]:

$$p_n^+ = G(p_{n-1}^+, \gamma). \quad (3)$$

An explicit expression for this function is given by Tailard [11] for $\zeta < 1$. This function depends on the control parameters γ and ζ . The time step n corresponds to the round trip time $\tau = 2l/c$ of the wave with velocity c along the resonator of length l .

Using the universal properties of the iterated maps [12, 13], useful information about the instrument behavior can

be drawn from the study of the iteration function. So far, these studies come from the *static* bifurcation theory, which assumes that the control parameter γ is constant. For instance, it is possible to determine the steady state of the system as a function of the parameter γ , and to plot a bifurcation diagram shown in figure 1. When no losses are considered, the oscillation threshold γ_{st} is:

$$\gamma_{st} = \frac{1}{3}, \quad (4)$$

For all values of the control parameter γ below γ_{st} the series p_n^+ converges to a single value p^{+*} corresponding to the fixed point of the function G , i.e. the solution of $p^{+*} = G_\gamma(p^{+*})$. When the control parameter γ exceeds γ_{st} the fixed point of G becomes unstable and the steady state becomes a 2-valued oscillating regime. Figure 1 shows an example of the bifurcation diagram with respect to the variable p^+ .

An iterated map approach can be used to predict the asymptotic (or *static*) behavior of an ideal clarinet as a function of a constant mouth pressure. This procedure avoids the phenomenon of *bifurcation delay* which is observed in numerical simulations when the control parameter γ is increased.

2.2 Slowly time-varying mouth pressure

2.2.1 Dynamic bifurcation

A control parameter γ increasing linearly with time is taken into account by replacing eq. (3) by eqs. (5a) and (5b):

$$\begin{cases} p_n^+ = G(p_{n-1}^+, \gamma_n) \\ \gamma_n = \gamma_{n-1} + \epsilon. \end{cases} \quad (5a)$$

$$(5b)$$

The parameter γ is assumed to increase slowly, hence ϵ is considered arbitrarily small ($\epsilon \ll 1$). When the series p_n^+ is plotted with respect to parameter γ_n the resulting curve can be interpreted as a *dynamic* bifurcation diagram and it can be compared to the *static* bifurcation diagram (fig. 2).

Because of the time variation of γ , the system (5) is subject to the phenomenon of bifurcation delay [14, 15]: the bifurcation point is shifted from the *static oscillation threshold* γ_{st} [3] to the *dynamic oscillation threshold* γ_{dt} [1]. The difference $\gamma_{dt} - \gamma_{st}$ is called the *bifurcation delay*.

The techniques used in dynamic bifurcation theory are now required to properly analyze the system. Article [1] provides an analytical study of the dynamic flip bifurcation of the clarinet model (i.e. system (5)) based on a generic method given by Baesens [14]. The main results of this study, leading to a theoretical estimation of the dynamic oscillation threshold of the clarinet are recalled below.

2.2.2 Dynamic oscillation threshold of the clarinet model without noise

A possible theoretical estimation of the dynamic oscillation threshold consists in identifying the value of γ for which the orbit of the series p_n^+ escapes from a neighborhood of arbitrary distance of an *invariant curve* $\phi(\gamma, \epsilon)$. More precisely, the dynamic oscillation threshold is reached when the distance between the orbit and the invariant curve becomes equal to ϵ .

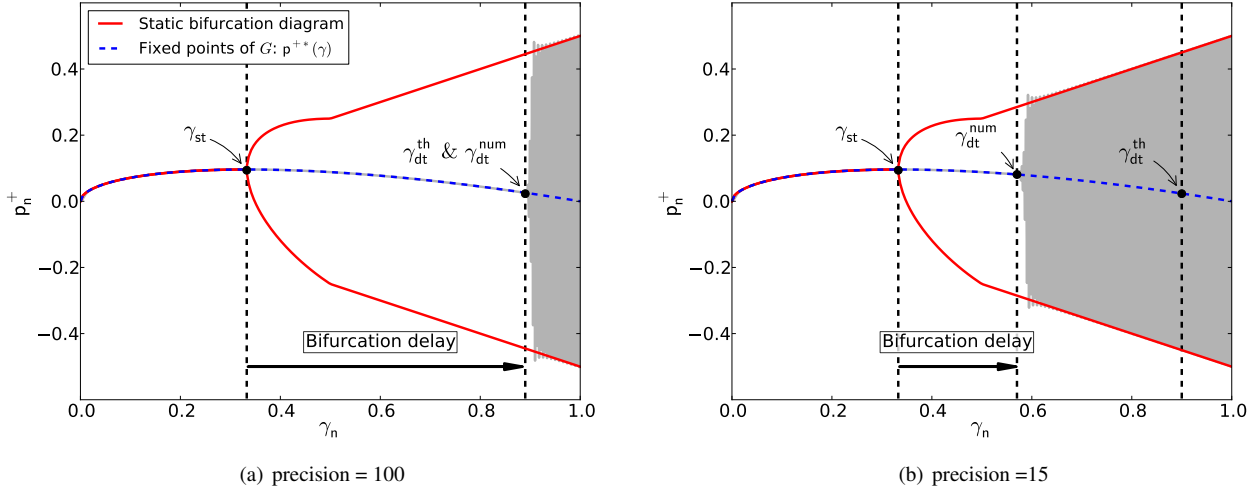


Figure 2. Comparison between *static* and *dynamic* bifurcation diagram as functions of γ_n . $\epsilon = 2 \cdot 10^{-3}$, $\zeta = 0.5$ and the numerical precision is equal to 100 (figure 2(a)) and 15 (figure 2(b)) decimal digits. The thresholds γ_{st} , γ_{dt}^{th} and γ_{dt}^{num} are represented.

The invariant curve (i.e. invariant under the mapping (5), described for example in [16]) can be seen as the equivalent of a fixed point in static regimes, functioning as an attractor for the state of the system. It satisfies the following equation:

$$\phi(\gamma, \epsilon) = G(\phi(\gamma - \epsilon, \epsilon), \gamma). \quad (6)$$

The procedure to obtain the theoretical estimation γ_{dt}^{th} of the dynamic oscillation threshold is as follows: a theoretical expression of the invariant curve is found for a particular (small) value of the increase rate ϵ (i.e. $\epsilon \ll 1$). The system (5) is then expanded into a first-order Taylor series around the invariant curve and the resulting linear system is solved analytically. Finally, γ_{dt}^{th} is derived from the analytic expression of the orbit.

The analytic estimation of the dynamic oscillation threshold γ_{dt}^{th} is defined in [1]:

$$\int_{\gamma_0 + \epsilon}^{\gamma_{dt}^{th} + \epsilon} \ln |\partial_x G(\phi(\gamma' - \epsilon), \gamma')| d\gamma' = 0, \quad (7)$$

where γ_0 is the initial value of γ (i.e. the starting value of the linear ramp). The main properties of γ_{dt}^{th} are (Fig. 6 of [1]):

- γ_{dt}^{th} does not depend on the slope of the ramp ϵ (provided ϵ is small enough)
- γ_{dt}^{th} depends on the initial value γ_0 of the ramp.

3. NUMERICAL SIMULATIONS: THE PRECISION CANNOT BE IGNORED

3.1 Problem statement

The above theoretical prediction ignores the round-off errors of the computer. The bifurcation delays γ_{dt}^{num} ob-

served in simulations¹ are seen to converge to the theoretical ones for very high numerical precision, typically when hundreds or thousands digits are considered in the simulation (cf. figure 2(a) where a precision² of 100 was used). However, in standard double-precision arithmetic (precision close to 15 decimals), theoretical predictions of the dynamic bifurcation point γ_{dt}^{th} are far from the thresholds γ_{dt}^{num} observed in numerical simulations. An example is shown in figure 2(b). In particular, the numerical bifurcation threshold depends on the slope ϵ , unlike the theoretical predictions γ_{dt}^{th} . Moreover, the dependence of the bifurcation point on the initial value γ_0 is lost over a wide range of γ_0 .

As a conclusion, because they ignore the round-off errors of the computer, theoretical results obtained in [1] fail to predict the behavior of numerical simulations carried out at usual numerical precision. In particular this is problematic when studying the behavior of a synthesis model, or simply when trying to understand the large delays in the threshold of oscillation in simulations of real systems. Following a general method given by Baensens [14], we show in section 3.2 how the dynamic oscillation thresholds of simulations with finite precision can be analytically predicted.

3.2 Theoretical estimation of the dynamic oscillation threshold in presence of noise

Following usual modeling of quantization as a uniformly distributed random variable [17], round-off errors of the computer are introduced as ξ_n (referred as an additive white noise). Therefore, system (5) becomes:

¹ In simulations, γ_{dt}^{num} is estimated as the value for which the distance between the simulated orbit and the invariant curve becomes equal to ϵ .

² The choice of the precision is possible using *mpmath*, the arbitrary precision library of *Python*.

$$\begin{cases} p_n^+ = G(p_{n-1}^+, \gamma_n) + \xi_n \\ \gamma_n = \gamma_{n-1} + \epsilon, \end{cases} \quad (8a)$$

$$(8b)$$

where ξ_n is a white noise with an expected value equal to zero (i.e. $\mathbb{E}[\xi_n] = 0$) and variance σ defined by:

$$\mathbb{E}[\xi_m \xi_n] = \sigma^2 \delta_{mn}, \quad (9)$$

where δ_{mn} is the Kronecker delta. The definition of the expected value \mathbb{E} is provided in [18]. Equations (8) are used for the analytic study. In later sections, the results of this analytical study will be compared to numerical simulations of the system (5) using a numerical precision of 15 decimals. As a consequence, the noise level σ will be equal to 10^{-15} .

The method to obtain the theoretical estimation of the dynamic oscillation threshold which take into account the precision (noted $\hat{\gamma}_{dt}^{th}$) is the same as to obtain γ_{dt}^{th} (cf. section 2.2.2). In addition, because of the noise the bifurcation delay is reduced so that the dynamic oscillation threshold γ_{dt} is assumed to be close³ to the static oscillation threshold γ_{st} . Using this approximation, the expression of $\hat{\gamma}_{dt}^{th}$ is:

$$\hat{\gamma}_{dt}^{th} = \gamma_{st} + \sqrt{-\frac{2\epsilon}{K} \ln \left[\left(\frac{\pi}{K} \right)^{1/4} \frac{\sigma}{\epsilon^{5/4}} \right]}, \quad (10)$$

which is the theoretical estimation of the dynamic oscillation threshold of the stochastic systems (8) (or of the system (5) when it is computed using a finite precision). K is a constant that depends on the slope of $\partial_x G(p^+(\gamma), \gamma)$, the derivative of the iteration function at the fixed point.

A summary table of different notations of the oscillation thresholds is provided in table 1.

Table of Notation	
γ_{st}	static oscillation threshold
γ_{dt}^{th}	theoretical estimation of the dynamic oscillation threshold of the clarinet model without noise
$\hat{\gamma}_{dt}^{th}$	theoretical estimation of the dynamic oscillation threshold in presence of noise
γ_{dt}^{num}	dynamic oscillation threshold calculated on numerical simulations

Table 1. Table of notation for thresholds of oscillation.

3.3 Benchmark of theoretical estimators for the dynamic oscillation threshold

This section compares the theoretical estimation of the dynamic oscillation threshold $\hat{\gamma}_{dt}^{th}$ for a standard deviation $\sigma = 10^{-15}$ with the thresholds observed in numerical simulations using the regular 64-bit double-precision of a CPU (about 15 decimal digits). The comparison is carried out as a function of the increase rate (ϵ) of the blowing pressure,

³ This hypothesis could be questioned because according to figures 2(a), even in the presence of noise, the bifurcation delay can be large. However, this hypothesis is required to carry out calculations.

the initial value γ_0 and the embouchure parameter ζ . The estimations of the theory with noise ($\hat{\gamma}_{dt}^{th}$) are plotted simultaneously with γ_{st} and γ_{dt}^{th} .

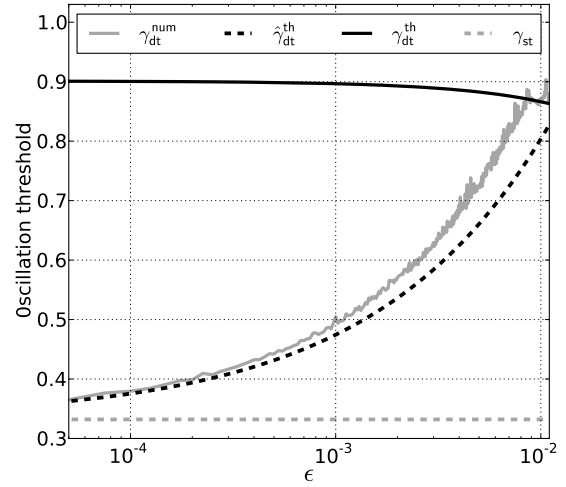


Figure 3. Graphical representation of γ_{dt}^{num} with respect to the slope ϵ , for $\gamma_0 = 0$ and $\zeta = 0.5$. Results are compared to analytic static and dynamic thresholds: γ_{st} , γ_{dt}^{th} and $\hat{\gamma}_{dt}^{th}$.

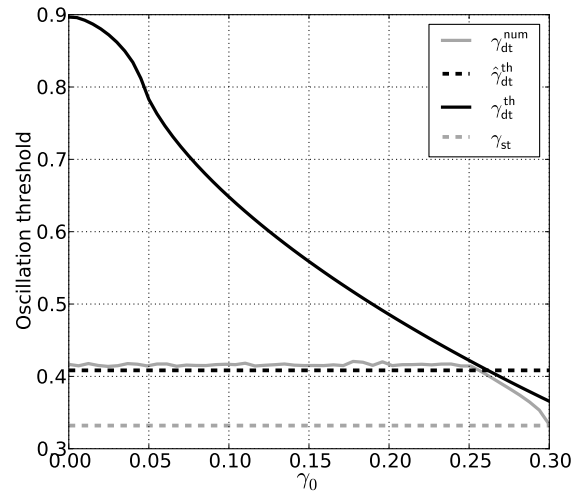


Figure 4. Comparison between theoretical prediction of oscillation thresholds (dynamic without noise: γ_{dt}^{th} and with noise: $\hat{\gamma}_{dt}^{th}$, and static γ_{st}) and the dynamic threshold γ_{dt}^{num} . Various thresholds are plotted with respect to the initial condition γ_0 with $\epsilon = 3 \cdot 10^{-4}$ and $\zeta = 0.5$.

In figures 3 and 4, the various thresholds are plotted with respect to ϵ and to γ_0 respectively. Unlike γ_{dt}^{th} (or γ_{dt}^{num} calculated on simulations with very high precision, see [1]), here the main properties of γ_{dt}^{num} are:

- γ_{dt}^{num} depends on the slope ϵ
- γ_{dt}^{num} does not depend on γ_0 over a wide range of γ_0 .

In both figures 3 and 4 we observe a good agreement between $\hat{\gamma}_{dt}^{th}$ and γ_{dt}^{num} . However, for large ϵ , in figure 3, and for γ_0 close to the static threshold γ_{st} , in figure 4, the theoretical threshold γ_{dt}^{th} for infinite precision is a better prediction of the dynamic threshold. Therefore, in this case, the round-off errors of the computer can be ignored.

In figure 5, thresholds are plotted with respect to the embouchure parameter ζ , showing that γ_{dt}^{num} decreases with ζ . In figure 5(a), two increase rates ϵ are used (10^{-4} and 10^{-3}). These slopes are sufficiently small so that the curves for γ_{dt}^{th} overlap⁴ (except for small values of ζ). In these situations, the estimation with noise $\hat{\gamma}_{dt}^{th}$ predicts correctly the observed dynamic thresholds γ_{dt}^{num} and, as expected, the prediction is better for the slower increase rate ϵ .

The behavior of the system changes for larger ϵ (cf. figure 5(b) where $\epsilon = 10^{-2}$). First of all, for this value of the slope the dependence of γ_{dt}^{th} on ϵ appears. Moreover, as in figure 4, beyond the intersection between $\hat{\gamma}_{dt}^{th}$ and γ_{dt}^{th} the theoretical estimation for infinite precision, γ_{dt}^{th} , becomes a better prediction of the bifurcation delay.

4. CONCLUSIONS

In a simplified model of the clarinet, the threshold in mouth pressure above which the oscillations occur can be obtained using an iterated map approach. This threshold corresponds to 1/3 of the reed beating pressure, but when the mouth pressure is increased with time, the oscillations start at a much higher value than this static threshold. The dynamic threshold calculated with infinite precision is independent on the rate of increase, depending only on the starting value of the mouth pressure.

Numerical simulations performed using finite precision show very different results in that the dynamic threshold depends on the increase rate and not on the starting value of the mouth pressure. A modified dynamic bifurcation theory including the effect of a stochastic variation in mouth pressure can be derived to correctly approximate the dynamic threshold when the precision is limited.

With a precision of 10^{-15} , this theory is seen to match the simulations performed with double-precision, showing that the threshold is situated between the static threshold $\gamma = 1/3$ and the dynamic threshold (close to $\gamma \simeq 0.9$ for typical playing conditions). This threshold increases with the rate of increase of mouth pressure ϵ , and does not depend on the initial condition γ_0 throughout most of the range below the static threshold. Although not shown here, the model can be applied to other values of precision. Because it is based on a stochastic analysis of the dynamic system, the model is expected to describe the behavior of the system subject for instance to turbulence noise.

Acknowledgments

This work is part of the project SDNS-AIMV “Systèmes Dynamiques Non-Stationnaires - Application aux Instruments à Vent”, sponsored by Agence Nationale de la Recherche (ANR).

⁴ cf. properties of γ_{dt}^{th} in section 2.2.2.

5. REFERENCES

- [1] B. Bergeot, C. Vergez, A. Almeida, and B. Gazengel, “Prediction of the dynamic oscillation threshold in a clarinet model with a linearly increasing blowing pressure,” *Nonlinear Dynamics*, pp. 1–14, 2013. [Online]. Available: <http://dx.doi.org/10.1007/s11071-013-0806-y>
- [2] J. Kergomard, S. Ollivier, and J. Gilbert, “Calculation of the spectrum of self-sustained oscillators using a variable truncation method,” *Acta. Acust. Acust.*, vol. 86, pp. 665–703, 2000.
- [3] J. Dalmont, J. Gilbert, J. Kergomard, and S. Ollivier, “An analytical prediction of the oscillation and extinction thresholds of a clarinet,” *J. Acoust. Soc. Am.*, vol. 118, no. 5, pp. 3294–3305, 2005.
- [4] A. Hirschberg, R. W. A. V. de Laar, J. P. Maurires, A. P. J. Wijnands, H. J. Dane, S. G. Kruijswijk, and A. J. M. Houtsma, “A quasi-stationary model of air flow in the reed channel of single-reed woodwind instruments,” *Acustica*, vol. 70, pp. 146–154, 1990.
- [5] A. Hirschberg, “Aero-acoustics of wind instruments,” in *Mechanics of musical instruments by A. Hirschberg/ J. Kergomard/ G. Weinreich*. Springer-Verlag, 1995, vol. 335 of *CISM Courses and lectures*, ch. 7, pp. 291–361.
- [6] C. Maganza, R. Caussé, and F. Laloë, “Bifurcations, period doublings and chaos in clarinet-like systems,” *EPL (Europhysics Letters)*, vol. 1, no. 6, p. 295, 1986.
- [7] J. Kergomard, “Elementary considerations on reed-instrument oscillations,” in *Mechanics of musical instruments by A. Hirschberg/ J. Kergomard/ G. Weinreich*. Springer-Verlag, 1995, vol. 335 of *CISM Courses and lectures*, ch. 6, pp. 229–290.
- [8] J. Kergomard, J. P. Dalmont, J. Gilbert, and P. Guillemin, “Period doubling on cylindrical reed instruments,” in *Proceeding of the Joint congress CFA/DAGA 04*. Société Française d’Acoustique - Deutsche Gesellschaft für Akustik, 22nd-24th March 2004, Strasbourg, France, pp. 113–114.
- [9] S. Ollivier, J. P. Dalmont, and J. Kergomard, “Idealized models of reed woodwinds. part 2 : On the stability of two-step oscillations,” *Acta. Acust. united Ac.*, vol. 91, pp. 166–179, 2005.
- [10] A. Chaigne and J. Kergomard, “Instruments à anche,” in *Acoustique des instruments de musique*. Belin, 2008, ch. 9, pp. 400–468.
- [11] P. Taillard, J. Kergomard, and F. Laloë, “Iterated maps for clarinet-like systems,” *Nonlinear Dyn.*, vol. 62, pp. 253–271, 2010.
- [12] M. J. Feigenbaum, “Quantitative universality for a class of nonlinear transformations,” *J. Stat. Phys.*, vol. 19(1), pp. 25–52, 1978.

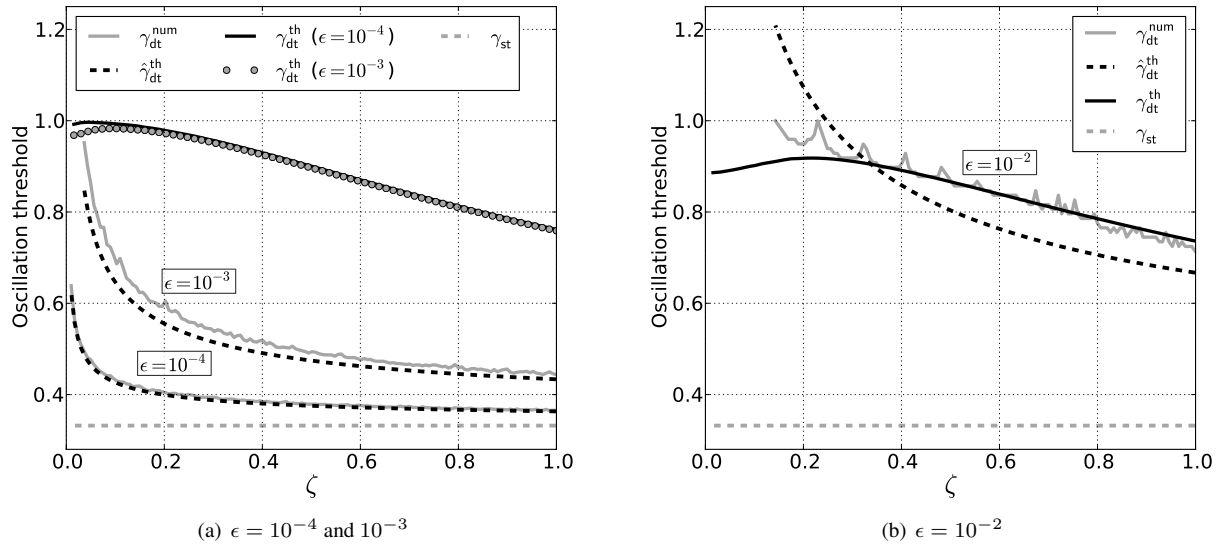


Figure 5. Comparison between theoretical prediction of oscillation thresholds (dynamic without noise: γ_{dt}^{th} and with noise: $\hat{\gamma}_{dt}^{th}$, and static γ_{st}) and the dynamic threshold γ_{dt}^{num} . Various thresholds are plotted with respect to the embouchure parameter ζ with $\gamma_0 = 0$ and (a) $\epsilon = 10^{-4}$ and 10^{-3} and (b) $\epsilon = 10^{-2}$.

- [13] M. J. Feigenbaum, "The universal metric properties of nonlinear transformations," *J. Stat. Phys.*, vol. 21(6), pp. 669–706, 1979.
- [14] C. Baesens, "Slow sweep through a period-doubling cascade: Delayed bifurcations and renormalisation," *Physica D*, vol. 53, pp. 319–375, 1991.
- [15] A. Fruchard and R. Schäfke, "Sur le retard à la bifurcation," in *International conference in honor of claude Lobry*, 2007. [Online]. Available: <http://intranet.inria.fr/international/arima/009/pdf/arima00925.pdf>
- [16] A. Fruchard and R. Schäfke, "Bifurcation delay and difference equations," *Nonlinearity*, vol. 16, pp. 2199–2220, 2003.
- [17] A. V. Oppenheim, R. W. Schaffer, J. R. Buck *et al.*, *Discrete-time signal processing*. Prentice-hall Englewood Cliffs, 1989, vol. 2.
- [18] S. M. Ross, *Introduction to Probability Models*, 9ème édition. Academic Press, 2006, ch. 2 "Random variables". [Online]. Available: <http://www.worldcat.org/isbn/0125980558>

An attempt at predicting the variation in playing frequencies for clarinets

Whitney Coyle

The Pennsylvania State University
Graduate Program in Acoustics
wlc5061@psu.edu
coyle@lma.cnrs-mrs.fr

J. Kergomard, P. Guillemain

C. Vergez, A. Guilloteau
LMA-CNRS Marseille, France
kergomard@lma.cnrs-mrs.fr

ABSTRACT

The input impedance measurement is today a standard method used by several wind instrument makers in designing modifications. For small modifications, knowing this quantity, and especially the resonance frequencies, is often sufficient. However, for a complete design, it is much better to know the playing frequencies themselves, which depend on several control parameters, such as the blowing pressure and reed opening. Using from values of these parameters, numerical computation (either in time or frequency domains allows us to determine the playing frequencies). This paper presents an attempt at deducing, analytically, these frequencies from the different control parameters, using of the input impedance curve. Three effects are known to alter the playing frequency and are examined separately: the flow rate due to the reed motion, the reed dynamics, and the inharmonicity of the resonator. An example of results is given for a clarinet.

1. INTRODUCTION

Musical instruments are complicated tools which continue to evolve as technology does.

. More and more alterations and improvements are being made on these instruments to give the best possible sound and functioning. Despite this

work there are still many complicated physical aspects of the instrument that are not yet understood. Although the shape and structure of a clarinet has been formed for quite some time there are modifications discovered often, to give a more clarinet-like timbre. The clarinet was not discovered and formed based necessarily on the physics of the instrument, which is complicated even with the help of today's modeling and simulation technology. The body of the French system clarinets and its key mechanisms are widely used and accepted as the benchmark and perhaps most ideal clarinet. There are still many unknown characteristics of the physics of the clarinet, the understanding of which could help a musician to play better or to choose a better instrument. For a musician one of the hardest characteristics to describe is what makes an instrument ideal, or even easy to play and this aspect of the clarinet, which has yet to be completely and thoroughly classified, is a key point in this research. An instrument which plays in tune, with very little help of the musician would be the preferred instrument. If all musicians were able to play in tune, with each other, without adjustment there would be room for more expressivity and the fatigue, which comes from adjusting the embouchure for tuning, would vanish. We begin the paper in Section 2 by discussing characteristic equations which are widely cited in literature which describes the functioning of the clarinet. Along with this we will describe the different playing parameters that are used in the models, blowing pressure and the reed opening. Next, in Section 3 the playing frequency is deduced for the ideal case, and in Section 4 we list and describe the analytic formulas which give

Copyright: ©2013 Whitney Coyle et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

the different frequency corrections based on the effects of reed induced flow, reed dynamics and inharmonicity. These three effects are assumed to be small and therefore they are studied independently. There are three examples of clarinet characteristics which can cause a variation in playing frequency and with these we can study carefully that difference between natural resonance frequencies of the instrument and the actual playing frequencies. Note that we are not taking into account, as of yet, the effects of a temperature gradient in the clarinet, which would also create a frequency variation, nor the impedance of the vocal tract. Section 5 will offer the results for the first register of a Buffet Bb clarinet and the paper concludes in Section 5 with some discussion, conclusions and a list of the planned future work.

2. BASIC EQUATIONS

2.1 Main equations

We use the classical model of 3 equations for the 3 unknowns p , u , and y , which are the pressure in the mouthpiece, the flow rate entering the instrument, and the displacement of the reed. The variables are dimensionless (see [1]) the pressures are divided by the closing pressure p_M ,

$$p_M = K_r H, \quad (1)$$

where K_r is the reed stiffness per unit area and H is the height of the reed channel at rest. The flow rates are divided by the ratio p_M/Z_c , where $Z_c = \rho c/S$ is the characteristic impedance at the input of the tube (ρ is the air density, c the speed of sound, and S is the cross section area at the tube input). The reed displacement y is divided by H : $y = -1$ is chosen for the beating reed, and $y = 0$ at rest.

The main dimensionless control parameters are the following:

- The mouth pressure is p_m , and its dimensionless expression is $\gamma = p_m/p_M$;
- The composite parameter ζ , is proportional to the maximum flow rate that can enter the tube, if w is the reed width:

$$\zeta = Z_c w H \sqrt{\frac{2}{\rho p_M}} \quad (2)$$

With these notations, the 3 equations can be written. Using the Bernoulli equations, and some hypotheses (see [2]), the flow rate is given by:

$$u = \zeta(1 - \gamma + x)\sqrt{\gamma - p} \text{ if } 1 + x - \gamma \geq 0 \quad (3)$$

$$u = 0 \text{ if } 1 + x - \gamma \leq 0$$

where $x = y + \gamma$. The second case corresponds to the point at which the reed begins to beat against the reed table, called the beating reed regime. Negative flow rate will not occur, in the permanent regime, for clarinet-like instruments (see for explanation [3]).

The other equations, which are linear are written in the frequency domain (notated with capital letters). The resonator is described by its input admittance $Y(\omega)$:

$$U(\omega) = Y(\omega)P(\omega) \quad (4)$$

The modal expansion of the dimensionless impedance $Z(\omega)$ can be written in a simplified way as follows:

$$Z(\omega) = j\omega \sum_n \frac{F_n}{\omega_n^2 - \omega^2 + j\omega\omega_n/Q_n} \quad (5)$$

where the ω_n are the resonance frequencies, and Q_n is the quality factor. F_n is the modal factor for the source and receiver placed at the input. For a perfect cylinder of length ℓ , it is equal to $2c/\ell$, and is independent of the rank of the resonance frequency. The reed motion is governed by the following equation:

$$[-\theta^2 + jq_r\theta + 1]x = p \quad (6)$$

where $\theta = \omega/\omega_r$ is the dimensionless frequency, and ω_r is the angular frequency of the reed resonance.

2.2 Supplementary equation: Flow rate due to the reed movement

The previous equations ignore the flow rate due to the reed movement, which is proportional to the velocity: $u_r = \lambda dx/dt$. Within the non-beating

reed conditions it is proportional to the reed compliance (see [4]) and can be introduced as a modification of the resonator equation:

$$U(\omega) = \left[Y(\omega) + j \frac{\omega V_{eq}}{c S_r} \right] P(\omega) \quad \text{with} \quad V_{eq} = \frac{\rho c^2}{K_r} S_r \quad (7)$$

where S_r is the reed area contributing to the flow rate. The effect is that of a compliance in parallel with the input admittance. This compliance is this of an added air volume at the entry of the instrument, V_{eq} . Moreover this equation is limited to the non-beating reed regime, and it ignores the reed dynamic (both $q_r = 0$ and $\omega_r = \infty$)¹. For convenience the notation $V_{eq} = S \Delta l$ is often employed, but this does not mean that the effect of the reed movement is simply an additional length correction, Δl , to the resonator, as will be discussed in the next section.

For a beating reed, the reed displacement is limited by the mouthpiece lay, therefore the flow rate is limited as well. Here we summarize the work done by Dalmont et al [5], who published a satisfactory comparison between a simple model and experiment. For a beating reed, the signal of the mouthpiece pressure is not far from a square signal, with the following values of the pressure: $-\gamma$ (in the case of a beating reed), and $+\gamma$ (in the case where the reed is open). These values are obtained in the approximation of a lossless resonator and a reed without dynamics. The reed displacement is supposed to vary between $-H$ and 0: also its first harmonic has an amplitude of H/π , and the amplitude of the reed velocity $v_r(t)$ is $j\omega H/\pi$. The amplitude of the first harmonic of the mouthpiece pressure is $2p/\pi$. Now, if we restrict the square signal to its first harmonic, we get (with dimensions):

$$V_r(\omega) = j\omega H \frac{P(\omega)}{2p_m}. \quad (8)$$

Therefore, the effect on the reed movement is equivalent to the effect of an additional air volume at the entry of the instrument, equal to

$$V_{eq} = \rho c^2 \frac{H S_r}{2p_m} = \rho c^2 \frac{S_r}{2\gamma K_r}. \quad (9)$$

¹ Reed dynamics could be taken into account, by dividing V_{eq} by $1 - \theta^2$, but the influence is very small, and in accordance with the assumption that the three effects mentioned in the introduction are independent, we ignore this correction.

Comparing to Eq. 7 shows that this result is coherent with the case of non-beating reed: for the simplest theory, a lossless resonator and the reed without dynamics, the beating threshold is given by $\gamma = 1/2$. Finally, we consider both cases of non-beating and beating reed by using the equation:

$$U(\omega) = \left[Y(\omega) + j \frac{\omega \Delta l}{c} \right] P(\omega) \quad (10)$$

$$\text{with } \Delta l = \frac{\rho c^2 S_r}{K_r} \frac{1}{S F(\gamma)}$$

$$F(\gamma) = 1 \quad \text{if } \gamma < 0.5$$

$$F(\gamma) = 2\gamma \quad \text{if } \gamma > 0.5$$

3. PLAYING FREQUENCY IN THE IDEAL CASE

When the reed flow rate and the reed dynamics as well as the inharmonicity are ignored, the playing frequency at the oscillation threshold is given by

$$A = Y(\omega), \quad \text{where } A = \zeta(3\gamma - 1)/(2\sqrt{\gamma}) \quad (11)$$

is the coefficient of the linear term of the nonlinear characteristic. This implies that $Im[Y(\omega)] = 0$, thus for the first regime $\omega = \omega_1$, and $A_{th} = Re[Y(\omega_1)]$.

$$\gamma_{th} \simeq \frac{1}{3} + \frac{2Re[Y(\omega_1)]}{3\sqrt{3}\zeta}. \quad (12)$$

However, what will happen above the oscillation threshold? When the reed dynamic is ignored, there is a static nonlinear characteristic which links the two variables pressure p and flow rate u . Therefore it is possible to use the “reactive power rule” found by Boutillon for bowed instruments [6]:

$$\sum_n |P_n|^2 n Im[Y(n\omega)] = 0. \quad (13)$$

This equation, where $P_n = P(n\omega)$ is the amplitude of the n th harmonic of the pressure, is one of the harmonic balance system of equations. If all the resonance frequencies are harmonically related to the first one, this equation is satisfied for $\omega = \omega_1$ regardless of the spectrum (or, equivalently, the excitation conditions). Thus the playing frequency does not change with the excitation level. The previous explanation seems to be trivial, but this will be useful when studying the non-ideal case, treated as perturbation of the ideal one.

4. APPROXIMATIONS FOR THE PLAYING FREQUENCY IN THE NON-IDEAL CASE

With Equations 1- 10 from the model, we are now able to deduce approximations for the difference between natural frequencies and the playing frequencies f_p of the clarinet. We are searching for:

$$\Delta f = f_p - f_1, \text{ with } f_1 = \omega_1/(2\pi).$$

If small enough, the relative difference can be expressed in cents, as follows:

$$\frac{\Delta f}{f_1} = 0.0006 N_{cents},$$

because one semi-tone (100 cents) corresponds to a ratio of 1.06. For a cylinder of length ℓ , we can define a length correction as:

$$\frac{\Delta \ell}{\ell} = -\frac{\Delta f}{f_1}.$$

In what follows, we consider the three effects separately, assuming that the frequency shifts (or the length corrections) can be simply added.

4.1 Flow rate due to the reed movement

If both the reed dynamics and the influence of higher order harmonics are ignored, the playing frequency is given by its value at the oscillation threshold (see [7]), i.e. given by

$$\text{Im} \left[Y(\omega) + j \frac{\omega \Delta \ell}{c} \right] = 0. \quad (14)$$

If the quality factor of a given impedance peak is high enough, only this peak will be kept in the modal decomposition around the resonance frequency ω_n , then the following approximation is valid¹:

$$\text{Im} [Y(\omega)] = \frac{2}{F_n} (\omega - \omega_n), \quad (15)$$

Therefore, for the first register, the solution of Eq. 14 is:

$$\omega_p = \frac{\omega_1}{1 + \frac{\Delta \ell F_1}{2c}}. \quad (16)$$

¹ This approximation supposes that ω_n is sufficiently far from the other natural frequencies, including the (negative) $-\omega_n$. This can be easily checked by expanding the impedance around ω_n , at the first order.

For a perfect cylinder, if we write $\omega \Delta \ell / c \approx \tan(\omega \Delta \ell / c)$, we find immediately the same result

$$\omega_p = \frac{\omega_1}{1 + \frac{\Delta \ell}{\ell}}, \text{ with } \omega_1 = \frac{\pi c}{2\ell},$$

as well as its condition of validity: $\omega_1 \Delta \ell / c \ll 1$. For this particular case, the effect of the flow rate can be viewed as a simple length correction. Formulas 10 and 16 are those proposed for this first effect, for both a beating reed and a non-beating reed.

4.2 Reed dynamics

When reed dynamics is considered, Eq. $A = Y(\omega)$ is no longer valid. The study of the effect of the reed dynamics on the oscillation threshold (frequency and mouth pressure) has been done by Wilson and Beavers [8], and completed by Ref. [9] who added the effect of the reed flow rate. The results are valid for the case of strong reeds (e.g. organ reeds, with small reed damping), and weak reeds (e.g. woodwind reeds, with high damping by the lips). This method involves the linearization of Eq. 3, and solving the characteristic equation. Here we consider the case of high damping, i.e. large q_r . The result is given by Chaigne and Kergomard [10]:

$$\omega_p = \omega_1 \left[1 - \frac{\zeta F_1}{2\sqrt{3}} \frac{q_r}{\omega_r} \right]. \quad (17)$$

For a perfect cylinder, this is again equivalent to a length correction, $\Delta \ell = q_r \zeta / (k_r \sqrt{3})$, where $k_r = \omega_r / c$. Notice that Nederveen tried to find this effect, but he considered a reed with high damping, with an infinite natural frequency [4].

This formula is valid at the threshold, therefore for a non-beating reed and a very small excitation pressure γ . Nevertheless, Kergomard and Gilbert [11], using the harmonic balance method analytically (limited to the first harmonic), found the following dependence on the excitation pressure:

$$\omega_p = \omega_1 \left[1 - \frac{\zeta F_1}{2\sqrt{3}} \frac{q_r}{\omega_r} \left[1 + \frac{3}{4} (\gamma - \gamma_{th}) \right] \right]. \quad (18)$$

This is the beginning of a series expansion above the oscillation threshold γ_{th} , therefore the validity should not be trusted at high mouth pressures.

This formula is obtained for non-beating reeds only. We do not have, for the moment, a formula for the beating reed case. The value of the oscillation threshold γ_{th} is given by Silva et al. [9], and we simplify it here by considering small values of $\theta = \omega/\omega_r$, and therefore use Eq. (12).

4.3 Effect of the inharmonicity of the resonator

It is well known that if the impedance peaks are not “aligned”, i.e. if the resonance frequencies are not exactly harmonic, the playing frequency changes with the level of excitation, because there is a weighting of the impedance peaks by the spectrum of the mouthpiece pressure (see e.g. Benade [12]). We define the inharmonicity by the following expression:

$$\omega_n = n\omega_1(1 + \eta_n), \text{ with } \eta_1 = 0.$$

It is valid for any given shape of the nonlinear characteristic $u = F(p)$, therefore for both non-beating and beating reed. Because we consider here no reed dynamics, this (static) characteristic exists ($x = p$ in Eq. (3), see also Eq. (6)). For clarinet-like instruments, we can limit the summation in Eq. 13 to odd harmonics. It is possible to use Eq. 15 near every resonance frequency, and we seek the playing frequency in the form:

$$\omega_p = \omega_1(1 + \epsilon).$$

From Eq. (13), we get, at the first order in η_n and ϵ :

$$\sum_n n^2 |P_n|^2 \frac{\epsilon - \eta_n}{F_n} = 0.$$

An interesting simplification can be done for nearly cylindrical instruments, because the modal factor F_n is almost independent of n (and equal to $2c/l$) and because $\eta_1 = 0$ the final result for the playing frequency is:

$$\epsilon = \frac{\sum_{n \geq 3} \eta_n d_n}{1 + \sum_{n \geq 3} d_n} \text{ with } d_n = n^2 \left| \frac{P_n}{P_1} \right|^2. \quad (19)$$

(notice, that for a square signal $d_n = 1$ for every odd n). Now, if the dependence of the spectrum with respect to the excitation pressure is known, it

is possible to deduce the variation of the playing frequency. Approximate formulas for clarinet-like instruments were given by Kergomard et al [2]. The decreasing of the higher harmonics is always faster than this of the square signal, therefore it is reasonable to search for an approximate formula by limiting the series to the third harmonic only. The ratio of the amplitude P_3/P_1 was found to be:

$$\frac{P_3}{P_1} = -\frac{1}{3} \frac{1}{1+z}; z = \frac{Y_3 - Y_1}{A - Y_1}; Y_n = Y(\omega_n). \quad (20)$$

Experimentally, there have been measurements of P_3/P_1 smaller than 1/3, but this formula is a good approximation for both nonbeating and beating reeds. Finally a first order approximation is found to be:

$$\epsilon = \frac{\eta_3}{1 + |1+z|^2}. \quad (21)$$

At the threshold $A = Y_1$, $\epsilon = 0$ (the signal is sinusoidal), and Y_1 is real. For large excitation pressure, z tends to zero, and ϵ to $\eta_3/2$. Notice that because the reasoning is based upon a perturbation at the first order, the values of Y_3 and Y_1 can be determined without inharmonicity, i.e. they are real and are the inverse of the values of the impedance peaks.

Refinements of this formula should be very intricate (as an example the formula for the 5th harmonic is very complicated, see [2]). Nevertheless, this formula exhibits the sense of variation of the effect on inharmonicity of the second peak.

5. TOTAL VARIATION: RESULTS FOR A BUFFET B♭ CLARINET

The language used in this section refers to notes of the first register of the B♭ soprano clarinet with the numbers 1 - 19. Note 1 represents the lowest note on the B♭ clarinet, the fingered E whereas Note 19 represents the highest note in the first register, the fingered B♭.

We utilized the same MKS values of the following for each note on the clarinet presented in this section: vibrating surface area of the reed, $S_r = w \cdot l = 6.5 \cdot 10^{-5}$, where l is the vibrating length of the reed and w is the width of the reed. We used the quality factor of the reed, $q_r = 0.2$, resonance frequency of the reed, $f_r =$

1200, speed of sound, $c = 340$, density, $\rho = 1.4$ and $\zeta = 0.2$ as reasonable choices for the playing parameters of the reed, and environment based on previous work by Wilson and Beavers [8] and Dalmont [13]. Many of these parameters are extremely difficult to measure in an experimental setting so it is best to choose well accepted values at this time. Concerning the effect of the reed movement, it is very difficult to estimate it from the knowledge of the reed area contributing to the flow rate, S_r . Because of this we have empirically used a value for the parameter Δl of 7 mm: this allows us to find a total length correction around 9 mm (see discussion below), in accordance with experimental results obtained by Dalmont et al. Each of these formulas discussed in the previous sections led to a specific frequency correction which will add to the expected resonance frequency: this becomes the calculated playing frequency of the clarinet. For reference, Figure 1 represents the extracted modal frequencies of the Buffet B \flat clarinet and Figure 2 represents the extracted modal factors for the the clarinet. The method of extraction is based upon a local optimization procedure (nonlinear least-square algorithm), and is not discussed here. Notice that the modal factor being inversely proportional to the length for a cylinder, it is an increasing function of the fundamental frequency of the notes, which are inversely proportional to the length of the equivalent cylinder.

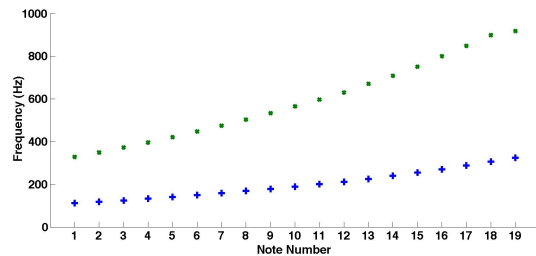


Figure 1. Values for the extracted modal frequencies, first register. The (blue) '+'s represent the first modal frequency and the (green) 'o's represent the second modal frequency.

Figures 3, 4 and 5 represent the frequency correction, in cents, for three different notes in the

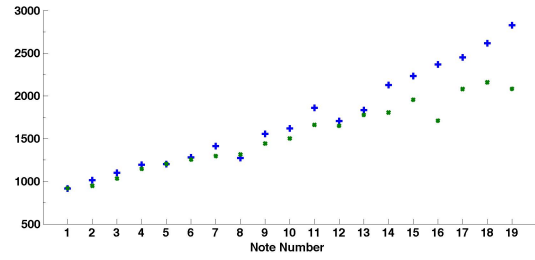


Figure 2. Values for F_n , the extracted modal factor, first register. The (blue) '+'s represent the factor for the first mode and the (green) 'o's represent that of the second mode.

first register of the B \flat clarinet (fingered E, E \flat and A \flat). The figures compare the calculated playing frequency to the extracted modal frequencies shown in Figure 1 for each particular choice of parameters. Notice that the frequency corrections are listed separately (reed induced flow, reed dynamics and inharmonicity) as well as totaled. Recall that, although the figures give a maximum value of $\gamma = 0.7$, the reed dynamics formula (18) is only accurate until $\gamma \approx 0.5$ (this value corresponds to the simplest expression of the beating reed threshold). Each note of the clarinet exhibits a negative inharmonicity though the two notes, 12 and 17 were chosen since they offer the maximum and minimum inharmonicities possible for the first register of the instrument in question. Note 1 was chosen for its different behavior and to show the lower extreme of the instrument. Each of the figures shows the expected behavior: the reed induced flow will have the greatest effect on the playing frequency and the effect of the reed dynamics does not depend very heavily on the value of the blowing pressure, γ .

Figure 6 shows the length corrections, in mm, as a function of note in the first register of the clarinet. It is calculated by the following formula:

$$\Delta \ell = -\frac{\Delta f}{f} \ell, \text{ where } \ell = \frac{F_1}{2c}.$$

Instead of using note numbers, the fingered note on the clarinet is shown. The value for ζ remained 0.2 and the value for γ was chosen to be γ_{th} , the threshold of oscillation for each given note. In this

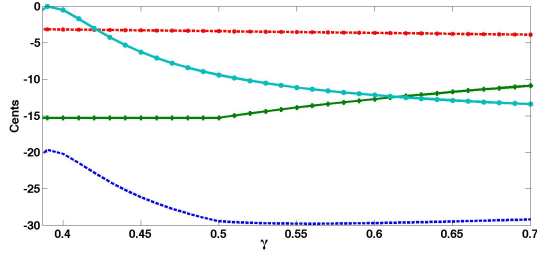


Figure 3. The frequency difference N_{cents} for Note 1 (fingering for E) of the clarinet. Note 1 values: $\zeta = 0.2$, $\eta_3 = -0.0201$, $F_1 = 915$ and $f_1 = 112$ Hz. The (green) solid line with '+'s represents the frequency shift due to reed induced flow, the (red) dotted line with 'x's is the effect of reed dynamics on the frequency and the (teal) solid line with 'o's shows the effect of the resonator's inharmonicity. Finally, the dotted (blue) line shows the combination of the three effects.

figure, notice that the curve for the inharmonicity is absent, this is because the formula for the inharmonicity of the resonator has no effect at the threshold of oscillation. For the reed motion effect, we find, as expected, a value close to 7mm since for a perfect cylinder this value would be exactly 7mm. As required, the total magnitude of the length corrections for these notes corresponds well to the work done by Dalmont et al in [13] which showed length corrections totaling 7mm to 10mm.

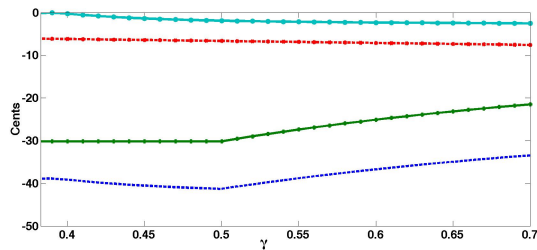


Figure 4. The frequency difference N_{cents} for Note 12 (fingering for Eb) of the clarinet. Note 12 values: $\zeta = 0.2$, $\eta_3 = -0.0036$, $F_1 = 1706$, $f_1 = 211$ Hz.

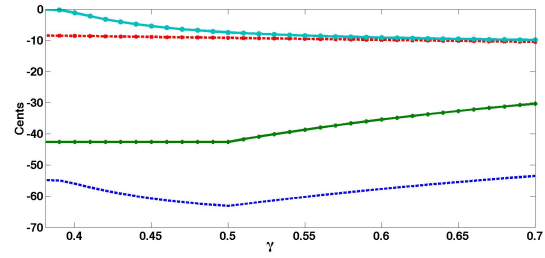


Figure 5. The frequency difference N_{cents} for Note 17 (fingering for Ab) of the clarinet. Note 17 values: $\zeta = 0.2$, $\eta_3 = -0.0142$, $F_1 = 2451$, $f_1 = 288$ Hz.

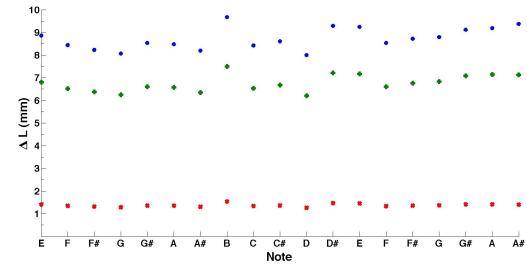


Figure 6. This figure represents the length corrections ($\Delta\ell$) values, in mm, as a function of note fingered on the clarinet, in the first register at the threshold of oscillation. Realize that the plot for the inharmonicity effect is absent since it is zero at the threshold of oscillation.

6. CONCLUSIONS

There are possibilities of improving the values of some parameters by attempting to measure them carefully, in the laboratory. Some of them can be deduced using an artificial mouth and perfect cylinders. Nevertheless the temperature gradient is not taken into account, nor the effect of the vocal tract (but it is known that it is small for the first register). The important point is that these results can help us to deduce the inhomogeneity of tuning from the inhomogeneity of the modal parameters. On this point of view, the aim of the present paper joins that of the paper by Almeida et al [14], who used an experimental approach. This is an attempt at predicting the values of the playing frequencies from the natural frequencies, with re-

spect to the variations of the main control parameters. The values below the beating reed regime are reasonable and there is still much that can be improved. To realize that there can be up to a 60 cent difference between the expected or measured impedance peak and the actual playing frequency is an important point in which to place focus. Instrumentalists know, intrinsically, that there are these tuning variations through the instrument but for the manufacturers, knowing where they are and their intensities is information that could help better the instruments overall and note-by-note tuning in turn creating a better overall instrument. The next step in this work is to improve the predictions above the beating reed regime, i.e. above approximately $\gamma = 0.5$. To do this we plan to compare the curves after this point to simulations of the playing frequency obtained by ab initio computation and fit, empirically, each individual frequency correction for each note for the total range of the clarinet and the varying values of γ and ζ . This will also be compared to measurements of the playing frequencies using an artificial mouth apparatus as well as using actual musicians to verify the validity of the analytical formulas over the playing parameters given.

Acknowledgments

Thanks to The Pennsylvania State Graduate Program in Acoustics and Dr. Dan Russell, funding from the NSF-GRFP and The ARCS Foundation, to LMA-CNRS, Marseille, and finally the ANR-Cagima project.

7. REFERENCES

- [1] A. Hirschberg, J. Kergomard, and G. Weinreich, *Mechanics of Musical Instruments*. Springer, 1995.
- [2] J. Kergomard, S. Ollivier, and J. Gilbert, "Calculation of the spectrum of self-sustained oscillators using a variable truncation method: Application to cylindrical reed instruments," *ACUSTICA - acta acustica*, vol. 86, pp. 685–703, 2000.
- [3] P.-A. Taillard, J. Kergomard, and F. Laloë, "Iterated maps for clarinet-like systems," *Non-linear Dynamics*, vol. 62, pp. 253–271, 2010.
- [4] C. J. Nederveen, *Acoustical aspects of woodwind instruments*. Northern Illinois University Press, 1998.
- [5] J.-P. Dalmont, P. Guillemain, and P.-A. Taillard, "Influence of the reed flow on the intonation of the clarinet," *Acoustics 2012 Nantes*, 2012.
- [6] X. Boutillon, "Analytical investigation of the flattening effect - the reactive power balance rule," *J. Acoust. Soc. Am.*, vol. 90, pp. 724–763, 1991.
- [7] N. Grand, J. Gilbert, and F. Laloë, "Oscillation threshold of woodwind instruments," *Acustica*, vol. 83, pp. 137–151, 1997.
- [8] T. Wilson and G. Beavers, "Operating modes of the clarinet," *J. Acoust. Soc. Am.*, vol. 56, pp. 653–658, 1974.
- [9] F. Silva, J. Kergomard, and C. Vergez, "Interaction of reed and acoustic resonator in clarinet-like systems," *J. Acoust. Soc. Am.*, vol. 124, pp. 3284–3295, 2009.
- [10] J. Kergomard and A. Chaigne, *Acoustique des instruments de musique*. Paris: Belin, 2008.
- [11] J. Kergomard and J. Gilbert, "Influence of the reed flow on the intonation of the clarinet," In *Actes du 5ème congrès Français d'acoustique*, 2000, pp. 294–297.
- [12] A. Benade, *Fundamental of musical acoustics*. Oxford University Press, 1976.
- [13] J.-P. Dalmont, B. Gazengel, J. Gilbert, and J. Kergomard, "Aspects of tuning and clean intonation in reed instruments," *Applied Acoustics*, vol. 46, pp. 19–60, 1995.
- [14] A. Almeida, "Clarinet parameter cartography: automatic mapping of the sound produced as a function of blowing pressure and reed force," *Proceedings of the International Symposium on Music Acoustics*, 2010.

ON REEDS AND RESONATORS: POSSIBLE EXPLANATIONS FOR CYCLIC SPECTRAL ENVELOPES IN THE CASE OF DOUBLE REED INSTRUMENTS

Sandra Carral

Institut für Wiener Klangstil (Musical Acoustics)
University of Music and performing Arts, Vienna
carral@mdw.ac.at

Christoph Reuter

Institut für Musikwissenschaft
University of Vienna
christoph.reuter@univie.ac.at

ABSTRACT

Proponents of the Pulse Forming Theory claim that the reed closing time of wind instruments remains approximately constant over their playing range, causing constant spectral gaps and therefore formants in the spectra of wind instruments. One of the latest measurements of the oboe reed closing time of 0.4 to 0.5 ms dates back to the 1970's, and is often quoted thereafter in other articles. They also claim that the reed closing time depends only on the characteristics of the reed, and disregard the influence of the geometry of the instrument. In this article a modern, minimally intrusive and accurate method for measuring the closing time of the oboe reed is presented. Other possible theories that could explain the approximately constant closing time of the reed in woodwind instruments are also discussed.

1. INTRODUCTION

Wind instruments are often shortened or lengthened in order to change the pitch. This results in a completely different resonance curve for every note. It would be logical to deduce that every note has therefore its own unique sound colour. Nevertheless, the perceived sound colour of wind instruments seems to remain constant across much of its playing range [1]. The reason for this seems to lie both in the sound production mechanism (pulse forming process) and in the resonance and radiation characteristics of the resonator itself.

1.1 Pulse forming theory

Several researchers including Frans Fransson [2] [3], Wolfgang Voigt [4] and Jobst Fricke [5] found that the spectra of wind instruments have constant spectral gaps and constant formant areas in between these gaps. Fransson [6] used an ionophon instead of a reed to excite the resonator, and changed the pulse length per period until the radiated sound was as close as possible to that of the instrument played with the reed. Voigt [4] used a high speed camera to record the movement of a bassoon reed. Both concluded

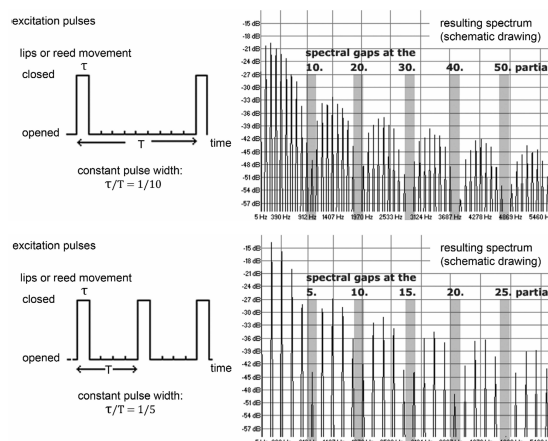


Figure 1. Lips or reed movement and the resulting source spectrum calculated from a closing time of $1/10$ of the period (above) and calculated from a closing time of $1/5$ of the period (below) [8]

that the reason why there are gaps in the spectrum is that the closing time τ of the reed remains approximately constant independently on the playing frequency. The ratio between the reed closing time τ and the signal period T determines the position of the spectral minima: When the reed is closed e.g. $1/10$ of the complete period, the spectrum will show minima on the 10th, 20th, 30th, etc. harmonic, in other words, the spectral minima lie on the $\frac{T}{\tau}$ th harmonic and its integer multiples (see Figure 1).

With his ionophon experiment, Fransson found a closing time of 0.4 ms for the oboe. Likewise, Heptner [7] quotes a closing time of 0.45 ms.

1.2 Interference Theory

Eugen Skudrzyk [9] and Nicolass Franssen [10] claim that the characteristic sound colour of wind instruments comes from the fact that the resonances of the instrument deviate from the harmonic structure of the reed or lip vibration. They claim that the inharmonicity of the resonances are due to the varying end corrections of every note. This theory would point again to the geometry of the instrument (see Figure 2).

This is in agreement with what Arthur Benade [11] [12] described: depending on which tone holes are open, the harmonics from the source will be reinforced when their

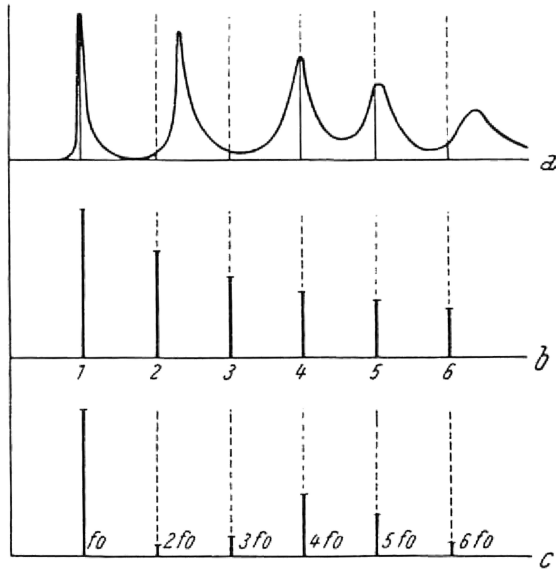


Figure 2. (a) Impedance peaks of a wind instrument, (b) harmonic spectrum of the source (reed or lip vibration), and (c) harmonic spectrum of the radiated sound [9]

frequency falls into a resonance, and attenuated when they fall in between the resonances of the instrument.

1.3 Analogy of a cone with a bowed string

It is well known (see for instance [13]) that the spectrum of a bowed or plucked string strongly depends on the position at which the string is excited. Plucking or bowing near to an antinode for a particular mode will excite it, whereas plucking or bowing near a node for a particular node will suppress it. Thus, if the string is excited at $1/2$ of its length, the spectrum will have missing harmonics: 2nd, 4th, 6th, etc. If it is excited at $1/5$ of its length, the missing harmonics will be the 5th, 10th, 15th, etc, and so on.

According to Ollivier et al. [14], changing the position of the string excitation is equivalent to changing the length of the truncation for conical woodwinds, such as the oboe (see Figure 3). A crucial difference between string and woodwind players though, is that woodwind players cannot control this parameter, since the length of the truncation is fixed. Let $T = T_o + T_c$ be the signal period, T_o and T_c the opening and closing times respectively, L_b the length of the truncated cone, and L_a the length of the truncation. If $\frac{T_o}{T_c}$ or $\frac{T_c}{T_o} = \frac{L_b}{L_a}$, the oscillation is called “Helmholz motion”, in which case the ratio of the durations of the two parts of the signal is determined by the resonator [15].

The constant closing time (and therefore the missing harmonics) found in the Pulse Forming Theory in Section 1.1 can be explained by this analogy as follows: Since at every note the length of the resonator changes, the ratio $\frac{L_b}{L_a}$ changes as well, however, the length L_a remains constant independently of the playing frequency. Assuming a Helmholz motion scenario, the ratio $N = \frac{L_b}{L_a}$ is equal to the ratio $N = \frac{T_o}{T_c}$, and since the signal period $T = T_c + T_o$ is directly proportional to the total length

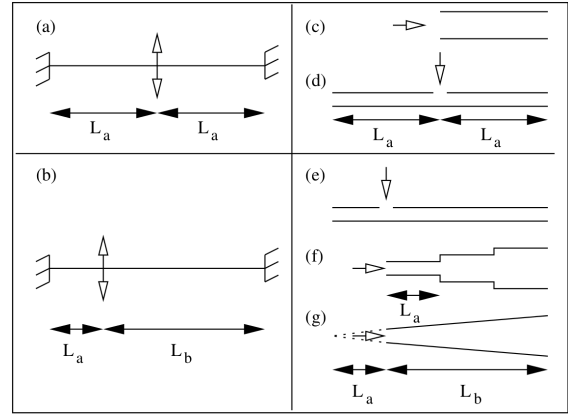


Figure 3. Summary of the formal analogy between the bowed string and woodwind resonators. White arrows indicate the location of the mouthpiece or the bow. A string bowed in the middle (a) is analogous to a clarinet (c) or (d). A string bowed at a location such as $L_a \neq L_b$ (b) is analogous to a cylindrical saxophone (e). If $\frac{L_b}{L_a} = N$ is an integer, it is equivalent to a stepped cone with N cylinders (f). It is approximately analogous to a truncated cone (g) of length L_b , with L_a the length of the missing part of the cone (dashed lines) [14]

of the cone $L = L_a + L_b$, then it follows that T_c remains constant independently of playing frequency (as does L_a) and is directly dependant on L_a , the length of the truncation (see Appendix). From this theory, the closing time T_c depends only on the geometry of the cone, specifically on the length of the truncation L_a .

1.4 Objectives

It is conceivable that both the reed characteristics and the instrument impedance are responsible for the characteristic sound colour of each instrument. The objectives of this paper are, on the one hand, to present a modern, minimally intrusive and accurate method for measuring the closing time of the oboe reed, and on the other hand to compare our results with what is found in the literature.

2. METHODOLOGY

According to [14], the reed opening and the mouthpiece pressure have the same phase, so one can measure the reed closing time by looking at the mouthpiece pressure, which can be measured with a microphone. Since the oboe reed is so small, the place at which the microphone is placed becomes problematic. A compromise was found between the closeness to the reed and the practicality of placing a microphone, by opening the first hole of the oboe body from the top, which corresponds to the second octave key hole, where a $1/8$ " microphone can be inserted. Special consideration had to be made on the choice of microphone due to the high sound pressure levels inside the oboe while being played.

2.1 Recordings

A 1/8" G.R.A.S. pressure microphone model 40DP was inserted tightly inside the second octave key hole of the oboe. A professional oboist was asked to play a full Conservatory professional automatic oboe (Stencil oboe marked Reisser Musik, made by Hans Kreul in Tübingen, Germany in the early 1980's, model 9111VA) and a Viennese Yamaha student semiautomatic oboe with two different reeds each. She played a C major scale over two octaves in three different dynamic levels (piano, mezzoforte and forte) with no vibrato, and in a mezzoforte dynamic level with strong vibrato. Each note was played separate from every other note, and lasting between 1 and 3 seconds. The microphone was inserted in the second octave key hole of the oboe, and was connected to a G.R.A.S. preamplifier model 26AS, then to a BSWA microphone conditioning unit model MC702. The output of the signal conditioner was connected to a Phantom MPA 2017 preamplifier. The level of the signal from the preamplifier was adjusted to be as high as possible (when playing forte) without saturating. This adjustment was made once, at the beginning of the recording session. The output from the preamplifier was then connected to an ADAT HD25 digital recorder, which converted the signals from analog to digital, and into an optical interface, before being connected to an RME computer sound card model DIGI96/8 PST. The sounds were recorded with the aid of the program Sony Sound Forge v 7.0. The sampling frequency f_s was set to 44100 Hz.

2.2 Signal Analysis

From the mouthpiece pressure signal it is possible to find out the opening and closing times of the reed. This is done as follows:

The mean value of the pressure over a period must be zero [14]. The pressure $P_{ref} = P_{max} - P_{min}$ is calculated. The time in which the pressure is below P_{ref} is then the closing time [16].

Each period of the steady state of the time domain signal was found by looking at the pressure maxima. Then P_{ref} was calculated, and the number of samples that fell above and below P_{ref} were counted and saved as opening N_o and closing N_c samples respectively (see Figure 4). The opening T_o and closing T_c times are calculated from the number of samples as follows: $T_o = \frac{N_o}{f_s}$ and $T_c = \frac{N_c}{f_s}$.

The mean and standard deviation of N_o and N_c throughout the duration of the steady state of the recorded note was calculated for each of the four recording conditions (piano, mezzoforte, forte without vibrato, and mezzoforte with strong vibrato).

3. RESULTS

Figures 5 (a) and 6 (a) show the measured mean \pm standard deviation of T_c for the French and Viennese oboes recorded with two different reeds each, at the three dynamic levels piano, mezzoforte and forte without vibrato, and mezzoforte with strong vibrato. These plots show that T_c is slightly affected by dynamic level changes, and that

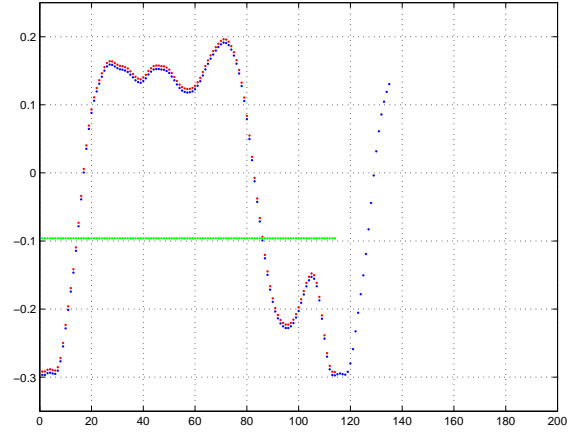


Figure 4. Method to extract T_o and T_c from the time domain signal: The blue dots show the time domain samples of a segment of the signal, the red dots show the time domain samples of one period with mean pressure value of zero taken from maximum to maximum amplitude, the green dots show the threshold pressure $P_{ref} = P_{max} - P_{min}$. Samples that are above P_{ref} are counted to calculate T_o , samples below P_{ref} are counted to calculate T_c .

it does not remain constant from note to note: The mean of T_c varies from about 0.65 to 1.15 ms in the notes measured.

Figures 5 (b) and 6 (b) show both T_o and T_c . They show how T_o dramatically shortens as the playing frequency increases, whereas T_c always remains somewhere around 1 ms.

4. DISCUSSION

As mentioned in Section 1.1, the closing time of the excitation signal of wind instruments has been measured in the past by several researchers, who claim that it remains approximately constant along the playing range of the instrument. The measurements presented in this paper support that. The measured closing time presented here of approximately 1 ms differs from the closing time of approximately 0.45 that Fransoon, Voigt and Heptner have published.

One method that has been used in the past to estimate the closing time has been to look for gaps that are equally spaced in the spectrum of the radiated sound of the instrument. In its simplest form, the reed movement can be assumed to be a square wave, in which case, the closing time is the inverse of the frequency where the first gap occurs. However, the estimation of the closing time from the position of the gaps depends not only on the frequency position of the gaps, but also on the exact form of the reed movement, as pictured in Figure 7 [4]. One or both could be the reason for the discrepancy.

Assuming a perfectly squared reed movement, the gaps in the spectrum occur at frequencies that are integer multiples of M , where $M = T/T_c$. For a harmonic spectrum, it is possible to find a harmonic that falls at exactly a "gap frequency" only when $M \in \mathbb{N}$. This can be problematic if for instance $M = 2.5$. In the spectrum, there will be gaps at the 5th, 10th, 15th, etc. harmonics, just as there

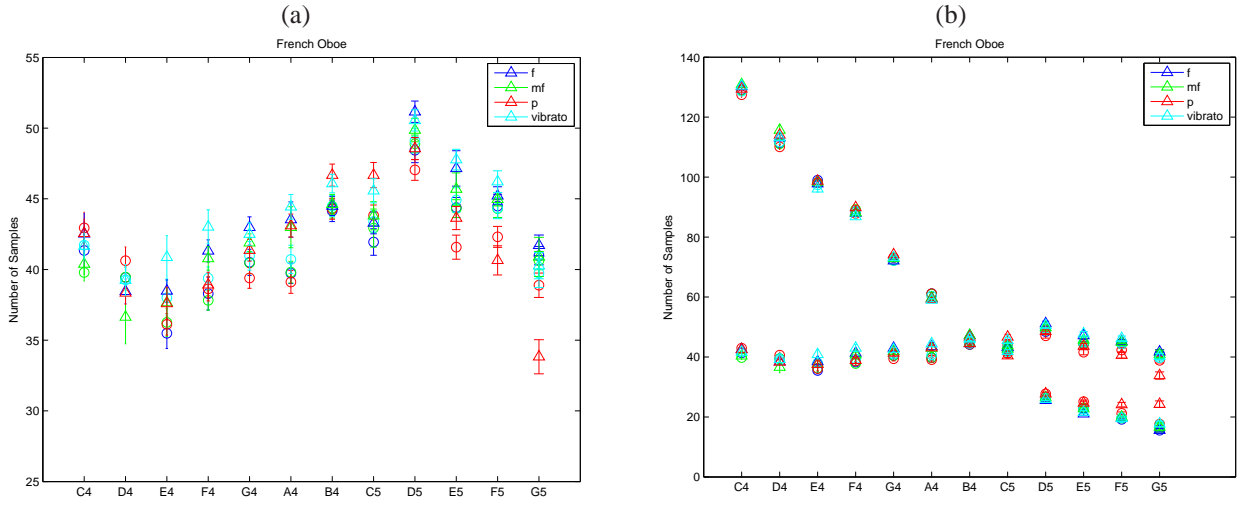


Figure 5. (a) Closing times $T_c = \frac{N_c}{f_s}$ and (b) Opening $T_o = \frac{N_o}{f_s}$ and closing T_c times of the notes of a C major scale played on a French oboe with two different reeds (one marked with circles and the other one with triangles) with three different dynamic levels without vibrato and on a mezzoforte dynamic level with strong vibrato

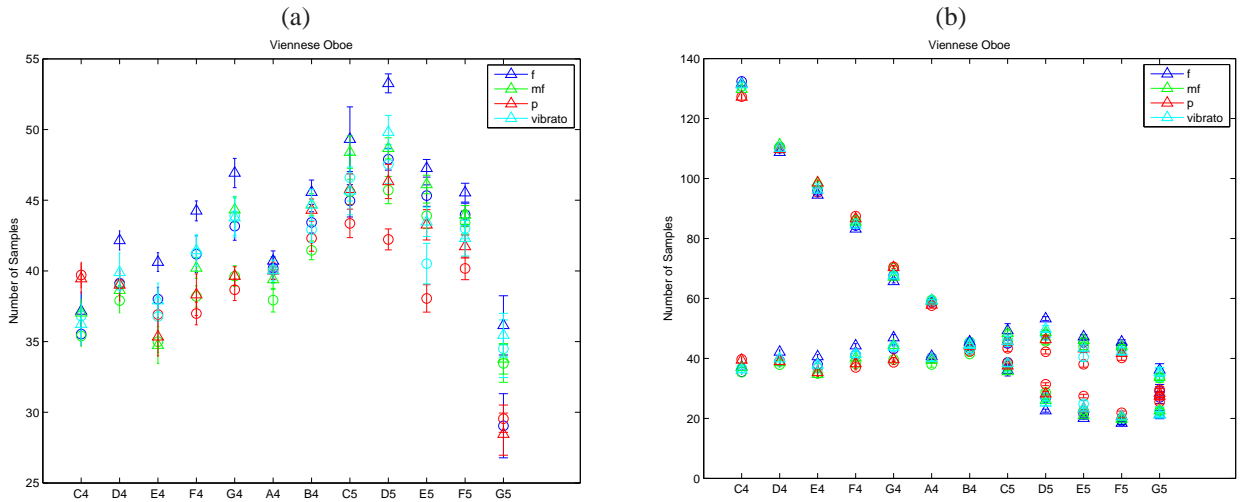


Figure 6. (a) Closing times $T_c = \frac{N_c}{f_s}$ and (b) Opening $T_o = \frac{N_o}{f_s}$ and closing T_c times of the notes of a C major scale played on a Viennese oboe with two different reeds (one marked with circles and the other one with triangles) with three different dynamic levels without vibrato and on a mezzoforte dynamic level with strong vibrato

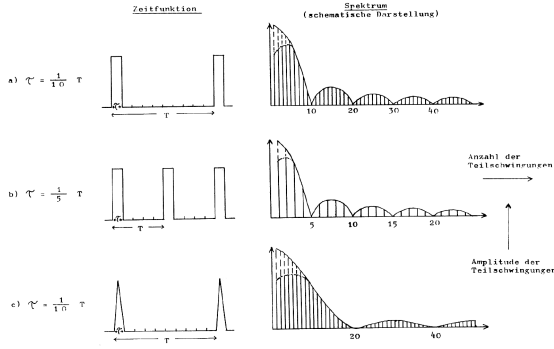


Figure 7. The position of the spectral gaps depends on the form of the reed movement: a) and c) have got the same period T and the same closing time T_c (marked as τ in the picture), however, the first gap is found on the 10th and 20th harmonic respectively, due to the fact that the area under the curve on c) is half of that on a) [4].

would be if $M = 5$, but since the other gaps (that is at 2.5, 7.5, 12.5, etc. times the fundamental frequency) fall in between the harmonics, it can be difficult to see them in the spectrum.

It is conceivable that the discrepancy found between the closing time found in the literature and the closing time presented here comes from the fact that, on the one hand, the measured signals like the one shown in Figure 4 are not square, and on the other hand, finding gaps and therefore estimating the gap frequency in the spectrum where $M \notin \mathbb{N}$ can be difficult.

They have also concluded that the closing time T_c depends mainly on the characteristics of the reed. Fransson made this conclusion following an experiment [2], whereby he played a cylindrical tube of 2.1 cm diameter and 1 m length with a clarinet reed and mouthpiece and with a bassoon reed¹. Since the measured closing times differed, he concluded that the resonator had little influence in them, therefore they must be set by the reed and its characteristics.

We have performed a similar experiment, whereby a fully Conservatory oboe (used for the recording described in Section 2.1) was played with two clarinet reeds (attached to the instrument via a special mouthpiece developed by one of the authors [17]). The closing times as well as the opening and closing times are shown in Figure 8 (a) and (b) respectively, and are approximately like the ones measured with the standard double reed, contrary to Fransson's conclusion.

The influence of the resonator on the reed closing time was also disregarded by other researchers such as Heptner [7], with the argument that the resonances of the body change with every note, therefore making gaps in the spectrum at the same places in frequency impossible².

As presented in Section 1.3 (see also the Appendix), the analogy of the conical resonator with the bowed string [14]

¹ It is not clear from this document how the bassoon reed was attached to the cylinder. It is possible that this was done with the aid of its conical bocal

² "The resonance curve of the conical resonator is not the reason for the formation of the formants in the spectrum"

would explain the missing harmonics in the spectrum, that lead to the so called formants described in Section 1.1.

It is possible that the inharmonicity presented by the interference theory (see Section 1.2) also causes areas in the spectrum where the impedance is low independently of the played note, therefore causing what is regarded as formants.

5. CONCLUSIONS AND FUTURE WORK

Several researchers have pointed out the presence of constant spectral gaps and constant formant areas in the spectra of wind instruments. They have observed that these gaps and therefore formants seem to be caused by an approximately constant reed closing time T_c . In the case of the oboe, the reed closing time T_c cited in the literature is somewhere between 0.4 and 0.5 ms. They have concluded that the closing time is determined by the characteristics of the reed, and at the same time they have disregarded the geometry of the resonator as a possible cause for it.

In this paper, a modern, accurate and minimally intrusive method for measuring the reed closing time T_c is presented. The reed closing time of a French and Viennese oboes, both played with two different double reeds, and the French oboe also with two single reeds was found to be consistent, and always around 1 ms. Two possible explanations for the discrepancy between these measurements and those found in the literature have been presented and discussed. In contrast to T_c , the opening time of the reed is dramatically reduced as the playing frequency increases. The fact that the closing time T_c on the French oboe remains in the same range when played with two clarinet reeds provides strong evidence that the influence of the reed characteristics on the closing time has been overestimated. The possibility exists as well that the influence of the resonator has been likewise underestimated.

Two other theories have been presented (the interference theory and the analogy of the conical resonator with the bowed string) that could explain how through the geometry of the resonator the spectral gaps and therefore formants can be formed in the spectrum.

It is certainly possible that the reed and its characteristics has some influence in the reed closing time, but at the same time it seems that the geometry of the resonator plays an important role in determining it. The next step will be to investigate the influence of both the reed and resonator on the origin and position of spectral gaps and formant areas, and their influence on the sound colour of double reed instruments.

6. APPENDIX

According to [14] and [15], the ratio N of the cone and its truncation is the same as the ratio of the two signal episodes:

$$N = \frac{L_b}{L_a} = \frac{T_o}{T_c} \quad (1)$$

$$T_o = NT_c \quad (2)$$

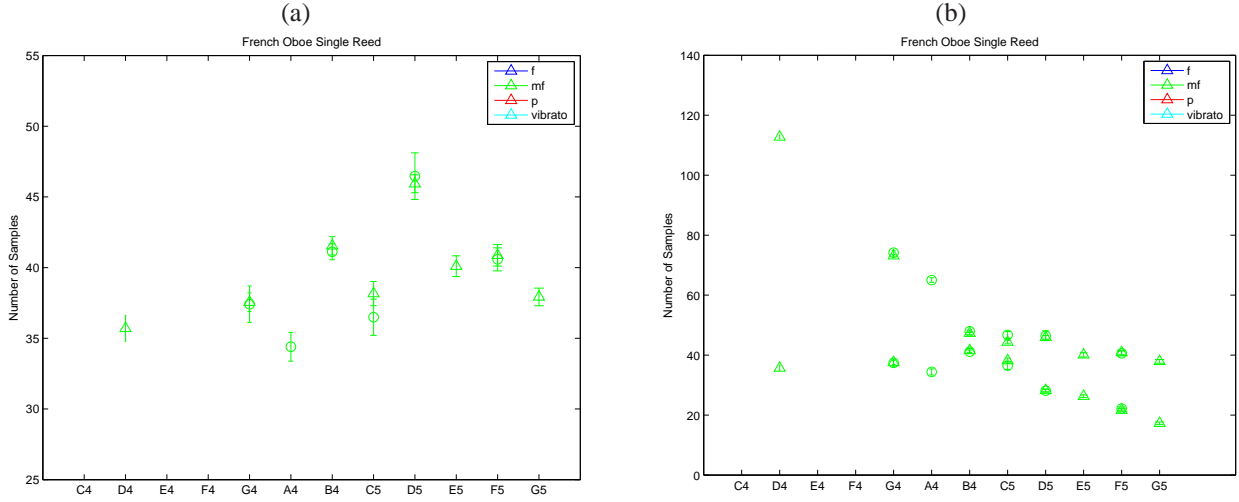


Figure 8. (a) Closing times $T_c = \frac{N_c}{f_s}$ and (b) Opening $T_o = \frac{N_o}{f_s}$ and closing T_c times of the notes of a C major scale played on a French oboe with two different single clarinet reeds (one marked with circles and the other one with triangles) without vibrato at a mezzoforte dynamic level

$$L_b = N L_a \quad (3)$$

$$(4)$$

The signal period T is related to the total length of the cone (including truncation) by:

$$L_a + L_b = \frac{\lambda}{2} \quad (5)$$

$$\lambda = \frac{c}{f} \quad (6)$$

$$2(L_a + L_b) = c(T_c + T_o) \quad (7)$$

$$T_c + T_o = \frac{2}{c} \cdot (L_a + L_b) \quad (8)$$

$$T_c + N T_c = \frac{2}{c} \cdot (L_a + N L_a) \quad (9)$$

$$T_c(N + 1) = \frac{2}{c} \cdot L_a(N + 1) \quad (10)$$

$$T_c = \frac{2}{c} \cdot L_a \quad (11)$$

Therefore, T_c is directly proportional to L_a , and given that L_a is constant, so is T_c

7. REFERENCES

- [1] C. Reuter, *Klangfarbe und Instrumentation*. Frankfurt: Lang, 2002, p. 111f.
- [2] F. Fransson, "The source spectrum of double-reed wood-wind instruments I," Department for Speech, Music and Hearing, KTH Computer Science and Communication, Tech. Rep., 1966.
- [3] —, "The source spectrum of double-reed wood-wind instruments II," Department for Speech, Music and Hearing, KTH Computer Science and Communication, Tech. Rep., 1967.
- [4] W. Voigt, "Untersuchungen zur Formantbildung in Klängen von Fagott und Dulzianen," Ph.D. dissertation, Regensburg: Bosse, 1975.
- [5] J. Fricke, "Formantbildende Impulsfolgen bei Blasinstrumenten," in *Proceedings of the DAGA 1975, 13. Jahrestagung für Akustik*, 1975, pp. 407–411.
- [6] F. Fransson, "The STL ionophone sound source," Department for Speech, Music and Hearing, KTH Computer Science and Communication, Tech. Rep., 1965.
- [7] T. Heptner, "Zur Akustik der Oboe: Theoretische Erörterungen und experimentelle Ergebnisse," *TIBIA*, vol. 12, no. 1, pp. 325–339, 1987.
- [8] M. Oehler and C. Reuter, "Dynamic excitation impulse modification as foundation of a synthesis and analysis system for wind instrument sounds," in *Communications in Computer and Information Science*, T. K. . T. Noll, Ed. Berlin: Springer, 2009, vol. 3, pp. 189–197.
- [9] E. Skudrzyk, *Die Grundlagen der Akustik*. Wien: Springer, 1954.
- [10] N. V. Franssen, "The mechanism of the human voice and wind instruments," in *Proceedings of the 4th International congress on Acoustics (ICA)*, Copenhagen, Denmark, 1962, pp. 21–28.
- [11] A. H. Benade, *Fundamentals of musical acoustics*. Oxford University Press, 1976.
- [12] —, "On woodwind instrument bores," *Journal of the Acoustical Society of America*, vol. 31, no. 2, pp. 137–146, 1959.
- [13] M. Campbell and C. Greated, *The musician's guide to acoustics*. Schirmer Books, 1987.
- [14] S. Ollivier and J.-P. Dalmont, "Idealized models of reed woodwinds. part I: Analogy with the bowed string," *Acta Acustica united with Acustica*, vol. 90, pp. 1192–1203, 2004.

- [15] J.-P. Dalmont, J. Gilbert, and J. Kergomard, “Reed instruments, from small to large amplitude periodic oscillations and the Helmholtz motion analogy,” *Acta Acustica united with Acustica*, vol. 86, pp. 671–685, 2000.
- [16] J.-P. Dalmont and J. Kergomard, “Elementary model and experiments for the Helmholtz motion of single reed wind instruments,” in *Proceedings of the International Symposium on Musical Acoustics*. Dourdan, France: Société Française d’Acoustique, 1995, pp. 115–120.
- [17] S. Carral, C. Vergez, and C. J. Nederveen, “Toward a single reed mouthpiece for the oboe,” *Archives of Acoustics*, vol. 36, no. 2, pp. 267–282, 2011.

EXTERNAL PIPE RESONATORS AND HARMONICA ACOUSTICS

James P. Cottingham

Physics Dept., Coe College
Cedar Rapids, Iowa 52402, USA
jcotting@coe.edu

Casey N. Brock

School Mech. Eng., Vanderbilt University
Nashville, Tennessee 37235, USA
casey.brock@vanderbilt.edu

ABSTRACT

Measurements of reed motion and sound field have been made on a diatonic harmonica mounted on a fixed volume wind chamber. These measurements include variation of sounding frequency with blowing pressure, and the degree to which the sounding frequency and sound spectrum can be altered by attaching external pipe resonators. The current results are compared with the results of measurements made in earlier studies. Differences were observed between the behavior of blow and draw reeds, as well as the dependence of the results on whether the secondary reed in the reed chamber is allowed to vibrate. For frequencies bent down with the pipe resonators, the sound and reed motion spectra displayed interesting changes in timbre and increased high frequency activity. Some of the high frequency activity may be related to the second transverse mode of reed vibration possibly involving coupling to the reed chamber resonance.

1. INTRODUCTION

In the diatonic harmonica, pitch bending has become essential to any skilled player. In recent years there have been a number of experimental and theoretical investigations of the acoustics of the diatonic harmonica in general and pitch bending in particular. The pioneering paper of Johnston [1], the work of Antaki and his associates [2], the work of Millot [3] and the paper of Ricot and Causse [4] are of particular importance.

In the absence of pipe resonators, the harmonica is a wind instrument in which the resonances of the vocal tract of the player play an essential role. These resonances, along with the presence of two differently pitched reeds sharing each wind chamber, are of prime significance in pitch bending.

Studies have also been done on the effects of free reeds coupled to pipe resonators which can result in pitch bending but also, in some cases, include dramatic changes in timbre and other unusual effects. [5]

In this paper results are presented of experiments in which a standard diatonic harmonica (Hohner Marine Band in G) is fitted with external pipe resonators. Although some other reeds were studied, the results reported here are all from experiments involving the reeds from

Hole 1 (blow reed C_4 /draw reed D_4) and Hole 8 (blow reed B_5 /draw reed A_5). This paper expands on an earlier report of some of these results. [6]

2. JOHNSTON'S EXPERIMENTS

Johnston studied a diatonic harmonica with an artificial vocal tract consisting of a pipe with an adjustable length L to simulate the effect of changes in vocal tract resonances. His results showed a correlation between the sounding pitch of a harmonica and the resonant frequencies of the player's vocal tract. [1] Johnston's data was replotted, converting the pitches to frequencies. By plotting his data together with theoretical resonance curves of the "vocal tract" pipe, the correlation between vocal tract resonance and sounding frequency can be easily seen in Figure 1.

Each reed chamber in the harmonica has two reeds, one designed to sound for each direction of air flow. The one normally sounding for a given air flow direction is referred to as the primary reed, and the other as the secondary reed. The secondary reed can have an effect on the sounding frequency, and is important in pitch bending. Johnston repeated his experiment with one of the two reeds blocked as well as with both reeds unconstrained. The figure shows his results for all three of these trials.

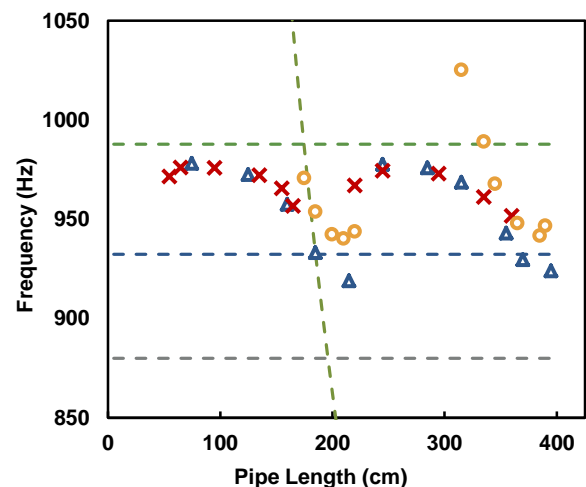


Figure 1. Johnston's Figure 4 redrawn with resonance frequency approximations. Triangles represent data for both reeds unconstrained; for crosses the secondary reed is blocked; for circles the primary reed is constrained. The horizontal dashed lines represent frequencies for the pitches A_5 , $A\#_5$, and B_5 , and the green dashed curves are the first two pipe resonances.

3. FREQUENCY VS. PRESSURE EXPERIMENTS

It is well known that the sounding frequency of a free reed, as in a harmonica, typically decreases as the blowing pressure is increased. This effect was investigated for the harmonica, with attention to any effect of the secondary reed on this expected pitch bending.

3.1 Experimental Setup

A laboratory blower was connected to a wind chest, consisting of a rectangular wooden box with an approximate volume of 11.3 liters. The harmonica was mounted above the box, with the desired playing hole of the harmonica over an opening in the wind chest. A manometer attached to the side of the box was used to determine the playing pressure, and a microphone was mounted near the harmonica. Figure 2 shows the experimental setup.

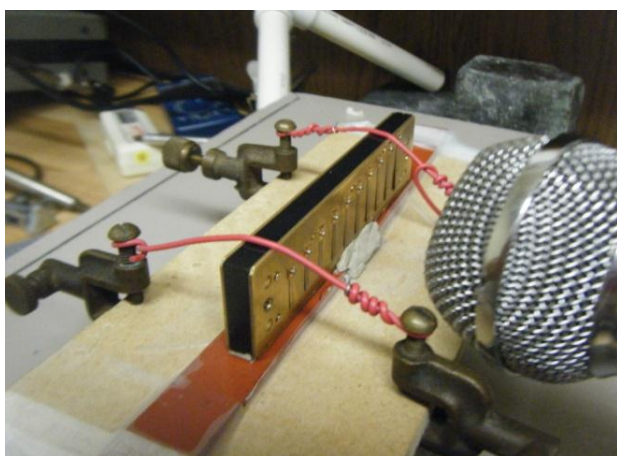


Figure 2. Harmonica mounted on wind chamber.

3.2 Procedure

The following experiment was done for seven reeds: both reeds in Hole 1 (blow reed G_3 -196 Hz and draw reed A_3 -220 Hz); both reeds in Hole 8 (blow B_5 -988 Hz/draw A_5 -880 Hz); both reeds in Hole 3 (blow D_4 -247 Hz/draw $F_4^\#$ -370 Hz); and the draw reed only for hole 5 (C_5 -523 Hz). The frequency was measured as the pressure in the wind chest was increased. The experiment was repeated with the secondary reed taped over preventing its oscillation and not allowing air to escape. It was repeated again with the secondary reed held in place with a small piece of putty, which constrained the reed but allowed air to escape. Figure 3 shows results for the blow reed in hole 1.

In all cases the pitch tended to decrease as blowing pressure increased. The secondary reed has an effect on the bending properties, but it is difficult to generalize the effect since the effect was different for each reed. There was a slight but systematic difference between taping the reed and putting the reed for high pressures. The lowest frequencies are found for the blocked reed channel. It appears that the vibration of the secondary reed plays a part in the harmonica's pitch bending characteristics.

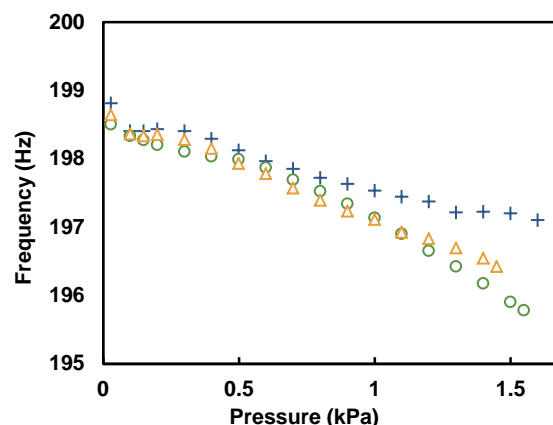


Figure 3. Sounding frequency as a function of pressure for the G_4 Blow 1 reed (196 Hz). Crosses represent data for both reeds unconstrained; for circles the secondary reed is blocked; for triangles the secondary reed is constrained, but air can escape.

4. PITCH BENDING WITH PIPES

4.1 Experimental Setup

A system of PVC pipes and connectors was created allowing any length of pipe between 2 cm and 70 cm in length to be couple to the harmonica. These pipes were connected to the exterior of the harmonica as shown in Figure 4, the setup otherwise the same as for the frequency vs. pressure experiments.



Figure 4. Harmonica mounted on wind chamber with attachment of variable length PVC pipes.

For each of the two reeds in Hole 8, a pressure was chosen near the threshold pressure. Pipes of different lengths were couple to the harmonica and the sounding frequency was determined. The experiment was repeated at higher pressures. The threshold pressure and frequency at threshold pressure were determined for each pipe length.

4.2 Pitch Bending Results

Figure 5 shows the results for the Blow 8 reed at the relatively low blowing pressure 0.2 kPa. There is a clear relationship between the sounding frequency trends and the resonance curves of the pipe. This is not surprising considering Johnston's results with the pipes used as an artificial vocal tract [1]. It is of interest that stopping the secondary reed does not destroy the bends due to pipe resonators, either in this case or the pitch bending shown in Figure 6. For typical harmonica bends executed by players, the secondary reed can become more active than the primary reed, and stopping it is expected to destroy the bend [1, 3].

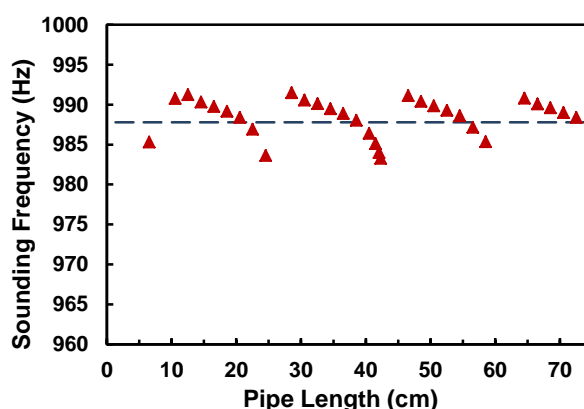


Figure 5. Sounding frequency vs. pipe length for the Blow 8 reed at 0.2 kPa. The dashed horizontal line marks the nominal pitch (B_5) of the reed.

The four curves in Figure 5 that appear as almost vertical lines represent the first four pipe resonances. It can be seen that there are gaps in pipe length where the reed pipe combination will not sound. The amplitude of reed vibration was also measured, but not shown, for each case in Figure 5. The maximum amplitudes consistently occur where the sounding frequency is close to 988 Hz, the frequency of the nominal pitch B_5 .

If the pressure is increased sufficiently, the reed pipe can be sounded for almost any pipe length. Figure 6 shows the sounding frequencies for the Blow 8 reed at threshold pressure, that is, the lowest pressure at which the reed can be made to sound. The graph shows the sounding frequency for threshold pressures ranging from 0.19 to 0.97 kPa for pipe lengths ranging from 2.5 to 11.0 cm, along with the corresponding values of the threshold pressure. What the graph does not show is the change in tone quality as the reed vibration is pulled farther from its natural frequency by the pipe resonance. The sound spectrum becomes much simpler, with far lower amplitudes of the high harmonics that characterize typical “free reed” sound.

There is a clear relationship between the sounding frequency trends and the resonance curves of the pipe. The following section discusses some interesting effects that occur when there is a significant mismatch between the pipe resonance and the natural frequency of the reed.

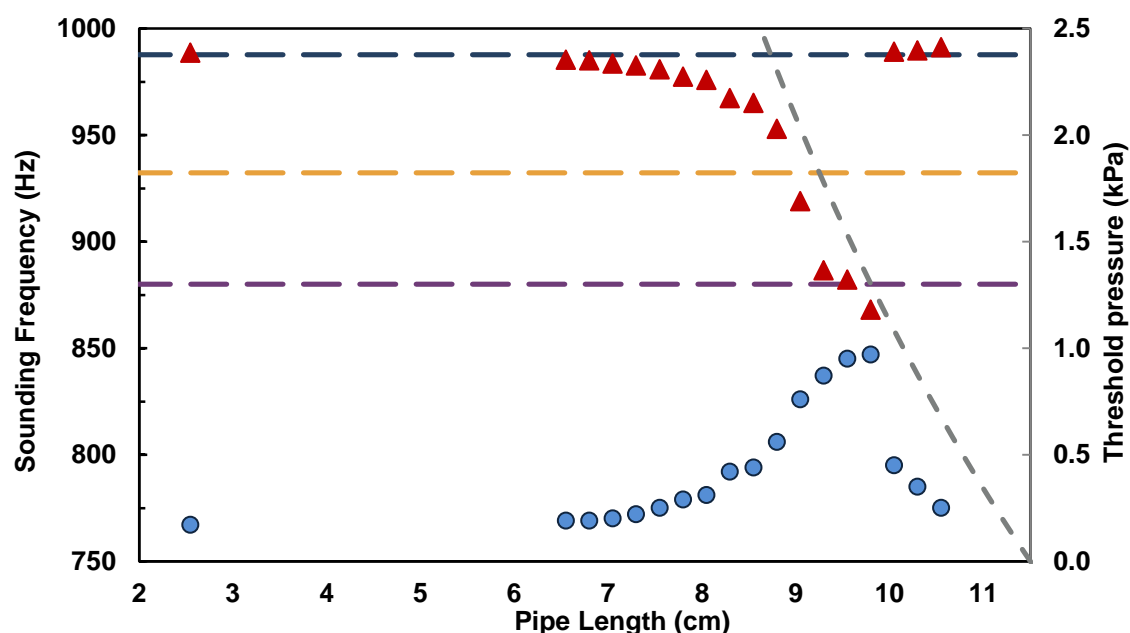


Figure 6. Sounding frequency and threshold pressure as a function of pipe length for blow 8 reed at threshold pressure. Triangles indicate the sounding frequency, circles show the threshold pressure. The pressure values observed (secondary axis) vary from 0.19 to 0.97 kPa. The horizontal dashed lines identify A_5 , $A\#_5$, and B_5 .

4.3 Some High Frequency Effects

Significant high frequency activity in the sound waveform of the blow reed was observed when the notes were bent down using the pipe resonators. The sound spectrum in Figure 7 shows a strong fundamental and groups of peaks in the 8.5 kHz and 17 kHz areas. The fundamental has been bent to 875 Hz, about a semitone below the nominal B₅ pitch of the Blow 8 reed.

On close inspection the high frequency group around 8.5 kHz consists of a strong peak at 8313 Hz (approximately 9.5 times the fundamental) surrounded by peaks with frequencies differing by multiples of 875 Hz. The grouping around 17 kHz includes frequencies which can be interpreted as harmonics of the fundamental, but further analysis below shows that difference frequencies are also involved there.

The high frequencies remain if the secondary reed is stopped, which eliminates the secondary reed as the source of this phenomenon. The high frequency content of the spectrum is not present if the pipes are removed.

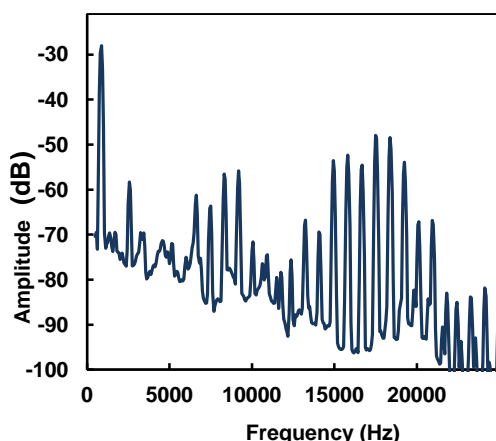


Figure 7. Spectrum for B₅ Blow 8 reed with 9.35 cm pipe, blown at 1.3 kPa

5. THE SECOND TRANSVERSE MODE

5.1 Studies of Reed Motion Using Laser Vibrometry

As noted above, the spectrum shown in Figure 7 at first appears to have formant-like features around 8.5 kHz and 17 kHz. A possible explanation for this would be the enhancement of harmonics in these regions by the resonator. In this case it would not be the simple cylindrical pipe, but the small reed chamber, 1-2 cm in length, which forms part of the complete resonator. This would be reminiscent of the high frequency content in the spectra of some organ reed pipes due to resonances of the shallot, as discussed by Miklos and coworkers. [7]

Further investigation, however, verified that the overtones were being produced by a higher transverse mode of vibration in the reed as shown in Figure 8. Using a laser vibrometer, velocity waveforms and spectra were obtained at various points on the reed tongue of the Blow 8 reed with the pipe attached as in the previous section. It was found that there were two strong peaks, one at the

fundamental at 875 Hz and one at 8133 Hz with two small difference frequency peaks at 8313 ± 875 Hz.

Motion of both the reeds in this wind chamber was studied using a laser vibrometer. Although the frequency sounded in this case (875 Hz) is close to the pitch of the secondary draw reed (A₅), there was little vibration of the secondary reed observed, and stopping the draw reed did not appreciably affect the observed sound spectrum.

It was observed that the vibration of the blow reed involved the second transverse mode of the reed tongue with significant amplitude, as shown in Figure 8.

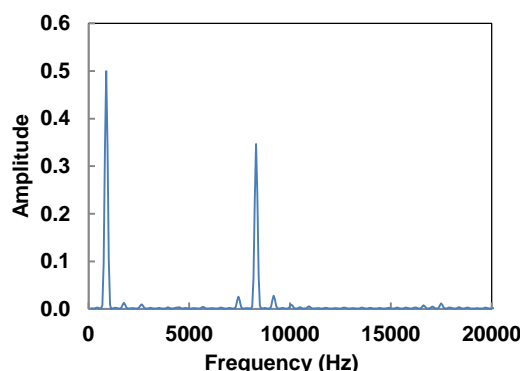


Figure 8. Spectrum of reed motion of the Blow 8 reed with the 9.35 cm pipe attached as observed with a laser vibrometer. The large peaks represent the first and second transverse modes of vibration.

5.2 Investigation of the Difference Frequencies

It was discovered that the note could be bent further by partially covering the open end of the pipe, lowering the pipe resonance frequency. When the pipe is covered, the fundamental is lowered and the difference frequencies shift towards their respective overtone.

A sound file was recorded as the end of the pipe was successively covered and uncovered. A spectrogram of this file is shown in Figure 9. There appear to be two main high frequency overtones, each surrounded by difference frequencies, differing in increments of the fundamental (around 875 Hz).

The graph in Figure 9 is in nine sections. In Section 1 there is no pipe. In Sections 2, 4, 6, and 8 the pipe is attached with the end uncovered. In Sections 3, 5, and 7 the pipe is attached with the end covered. In Section 9 the pressure is decreasing. Section 3 is highlighted with arrows to emphasize the shift in the difference frequencies. Notice that the two main high frequency overtones remain constant throughout. This is noticeable especially in Section 9 when even the fundamental has almost disappeared from the spectrum.

The waterfall plot in Figure 10 verifies that the second transverse mode is significant in the reed vibration with the 9.35 cm pipe attached. The mode at 8.31 kHz is clearly shown with the nodal pattern. Although it is less clear, the nodal pattern for the third mode is discernible around 17.5 kHz. It is interesting to note that the frequencies in the sound spectrum of Figure 9 related to this mode remain constant as the pipe resonance is changed, while the fundamental frequency, corresponding to the first transverse mode, is altered.

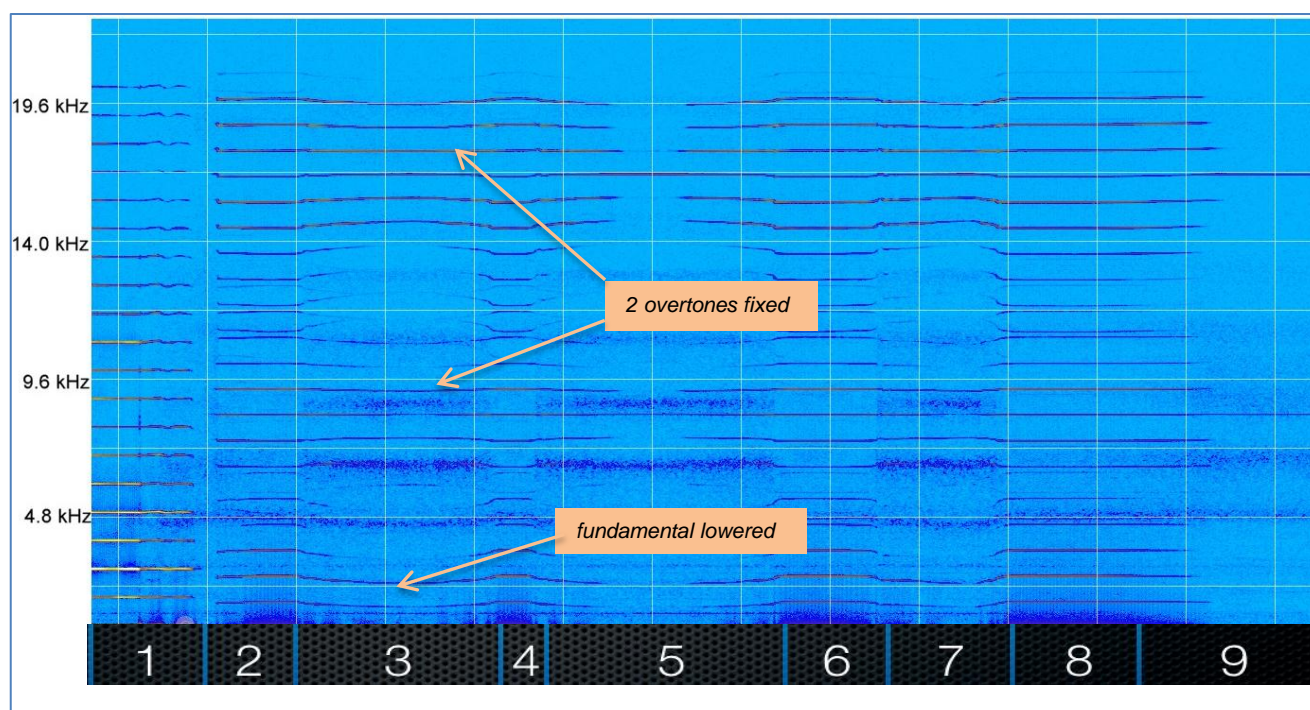


Figure 9. Spectrogram of the Blow 8 note bending at 1.3 kPa with the 9.35 cm pipe successively covered and uncovered. In regions 3 and 8 the two overtones at 8.31 kHz and 17.5 kHz stay fixed as the pitch is bent, while the difference frequencies shift.

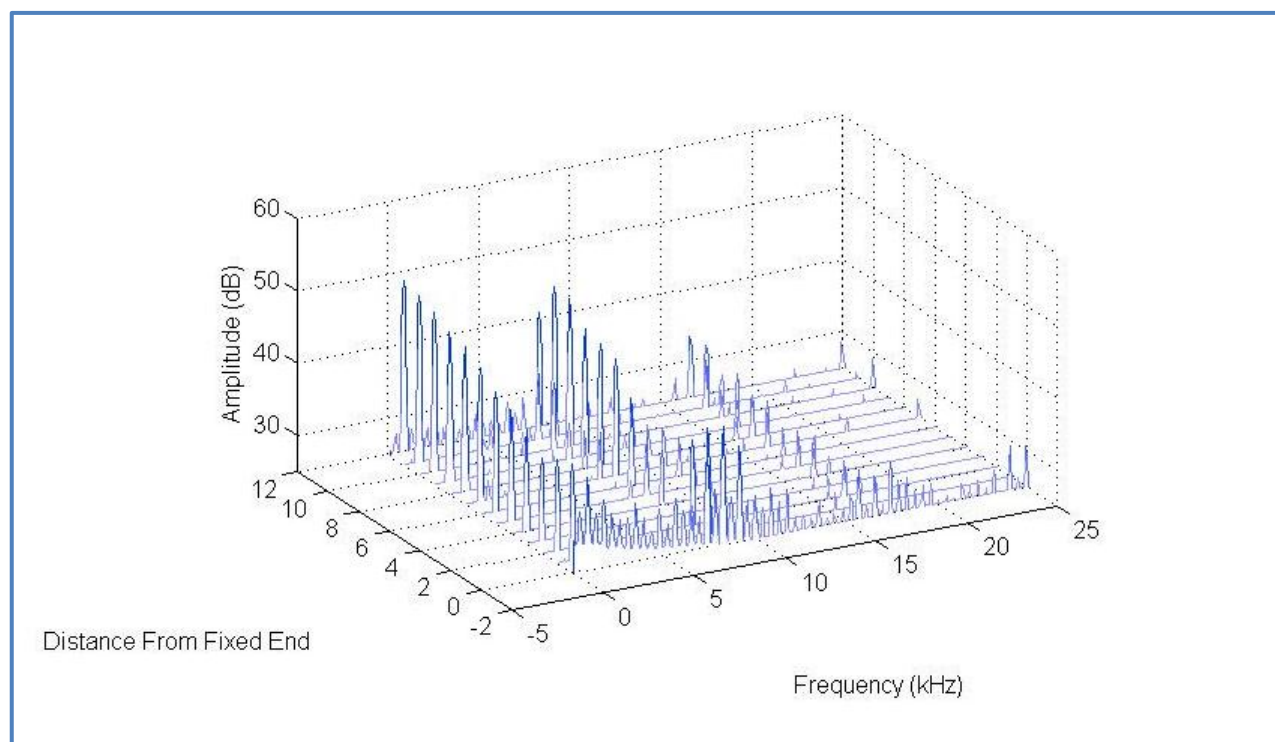


Figure 10. Waterfall plot of Blow 8 reed motion with a 9.35 cm pipe attached, observed with a laser vibrometer. Note the nodal pattern of the second mode at 8.31 kHz as well as the presence of the third mode around 17.5 kHz. (The transverse mode frequencies vary from those expected for the uniform beam because of the non-uniform cross section of the reed tongue.)

6. DISCUSSION

The results presented here show that pitch bending in the harmonica is possible with external pipe resonators, and the results regarding change in pitch correspond generally with other experiments done with free reeds mounted in pipes. The kind of pitch bending with pipes exemplified in Figures 5 and 6 is very similar to results reported earlier by Cottingham for reed organ reeds mounted in plastic pipes [5] and by Braasch for free reed organ pipes. [8] These results are also consistent with Johnston's results, although the pipes in this investigation are inserted between the harmonica and the outside air rather than between the air supply and the harmonica.

The results in Figures 5 and 6 are also quite similar to those published much earlier by Wilson and Beavers [9] for the clarinet. This is not surprising, since the type of free reed used in the harmonica, accordion and reed organ behaves as a blown-closed reed under normal circumstances. The Asian free reed mouth organs use free reeds in bamboo pipes, but these are reeds of symmetric construction, often sounding on both directions of airflow. They behave as blown-open reeds, sounding above the resonance frequency of the pipe. The sounding frequencies of both types of free-reed pipe can be calculated well using the general model of Fletcher [10], and this has been done for the Asian instruments. [11]

Some interesting and unexpected effects occurred when there was a strong mismatch between the pipe resonances and the natural frequency of the reed as described in Section 5. In this case the reed would not sound at normal blowing pressure and, when it did sound at higher pressure, the high frequency components described above appeared. Instances of the sounding of the second transverse mode in a free-reed pipe, either alone or together with the fundamental, have been reported previously, but without the presence of the prominent difference frequencies. [5]

The explanation for the prominence of the second mode frequency in this case is not completely clear, but it is possible that a resonance in the reed chamber supports the second mode vibration at 8313 Hz, and that this mode becomes relatively more prominent as the fundamental mode receives less support from the pipe resonances. This phenomenon may be similar in origin to the high frequency effects reported by Miklos, et al, due to resonances of the shallot in organ reed pipes. Further investigation of this is necessary.

7. CONCLUSIONS

Because free reed instruments such as the harmonica, accordion, reed organ, and harmonium generally do not use pipe resonators, it is often thought that coupling between the free reed and resonators is of only secondary importance. In harmonica playing, of course, it is well established that proper use of resonances of the player's vocal tract are essential in many styles of playing. Continuing study of reed-resonator interaction in the harmonica and other free reed instruments is necessary for full

understanding of these instruments. These investigations are also useful in the understanding of the operation of wind instruments in general.

Acknowledgments

This work received partial financial support from the Coe College Acoustics Research Fund and National Science Foundation REU Grant PHY-1004860. Miles Faaborg helped with this project in many ways, most notably as the first person to observe the high frequency effects discussed in Section 4.3 and Section 5. Boone Mathems and Quinn Wilson assisted in preparation of the graphics.

REFERENCES

- [1] R. Johnston. "Pitch Control in Harmonica Playing." *Acoustics Australia* vol. 15, pp. 69-75, 1987.
- [2] H. Bahnsen, J. Antaki, Q. Beery, "Acoustical and physical dynamics of the diatonic harmonica." *J. Acoustical Society of America*, vol. 103, pp. 2134-2144, 1998.
- [3] L. Millot, "A Proposal for a Minimal Model of Free Reeds." *Acta Acustica/Acustica*, vol. 93, pp. 122-144, 2007.
- [4] D. Ricot, R. Caussé, N. Misdariis, "Aerodynamic excitation and sound production of blown-closed free reeds without acoustic coupling: The example of the accordion reed." *J. Acoustical Society of America*, vol. 117, pp. 2279-2290, 2005.
- [5] J. Cottingham, "Pitch Bending and Anomalous Behavior in a Free reed Coupled to a Pipe Resonator." in *Proceedings of the International Symposium on Musical Acoustics*, Barcelona, 2007.
- [6] C. Brock, J. Cottingham. "The diatonic harmonica, pipe resonators, and the siren." *J. Acoustical Society of America*, vol. 130, p. 2343, 2011.
- [7] A. Miklos, J. Angster, S. Pitsch, T. Rossing, "Interaction of reed and resonator by sound generation in a reed organ pipe." *J. Acoustical Society of America*, vol. 119, pp. 3121-3129, 2006.
- [8] J. Braasch, C. Ahrens, "Attack Transients of Free Reed Pipes in Comparison to Striking Reed Pipes and Diapason Pipes." *Acustica/Acta Acustica*, vol. 86, pp. 662-670, 2000.
- [9] T. Wilson, G. Beavers, "Operating modes of the clarinet." *J. Acoustical Society of America*, vol. 56, pp. 653, 658, 1974.
- [10] N. Fletcher, "Excitation Mechanisms in woodwind and Brass Instruments." *Acustica*, vol.43, pp. 63-72, 1979.
- [11] J. Cottingham, E. Dieckman, "Measured and calculated sounding frequencies of pipes coupled with free reeds," *Proc. Meetings on Acoustics*, vol. 16, 035003, 2009.

Influence of pad “resonators” on a saxophone

Pauline Eveno & René Caussé

Ircam
(UMR CNRS 9912)
1 place Igor Stranvinsky
75004 Paris, France
pauline.eveno@ircam.fr

Marthe Curtit

Itemm
Pôle d’Innovation
des Métiers de la Musique
71 avenue Olivier Messiaen
72000 Le Mans, France
marthe.curtit@itemm.fr

Jean-Pierre Dalmont

Laboratoire d’Acoustique
de l’Université du Maine
(UMR CNRS 6613)
72085 Le Mans, France
jean-pierre.dalmont@univ-lemans.fr

ABSTRACT

Toneholes have an important role in the acoustics of woodwind instruments. In the saxophone, the toneholes are surmounted by a key provided with a pad and what musicians and artisans refer as a “resonator”. They are flat disks made of metal or plastic fixed in the middle of the pad and are called with the acoustically neutral term “cover” in the article. In order to understand their role, measurements of the input impedance of a cylinder topped by a key with interchangeable pads (with and without covers) are performed. For closed holes, pads with covers have a low absorption coefficient; for open holes, effects of the covers on the radiation are highlighted when the key is at a small height. An analysis of the vibrations of the pads shows that these effects can be explained by the high mobility of the pads without a cover, which seems to act as a stiffener. Then, the input impedance measurement of an entire saxophone confirms that, when the holes are closed, the effect of a pad without a cover is to increase the damping. The effect on open holes is close to negligible. Finally, measurements in playing situations show that saxophones without covers have higher harmonic spectral centroids and require greater mouth pressure.

1. INTRODUCTION

Toneholes have an important role in the acoustics of woodwind instruments. Their opening and closing allows the musician to play different notes. Changing their position or their geometric features provides a way to modify the playing frequencies and the timbre of the instrument.

Toneholes can have a complex geometry (e.g. conical shape, undercutting) and involve several elements such as chimneys, key, pad, and finger. In his thesis, Lefebvre [1] gives a large overview of what is already known about toneholes and what remains to be investigated. The simple unflanged tonehole (i.e. a tonehole with a chimney which can be found in modern metal instruments such as concert flutes or saxophones) is now well described [1–4]. This is also the case for the tonehole directly drilled in

the wall (i.e. a tonehole without a chimney which can be found in many instruments made of wood such as classical flutes or recorders) [1, 5]. Moreover, Dickens [6] provides fit-formulae that match his experimental results for closed drilled holes. Dalmont et al. [5] give an analytical formula for keys positioned above a hole with a chimney that is valid for a range of key heights excluding very small values [1, 7, p. 80-85]. Some studies have been carried out on undercut holes [8, 9, p. 321] but no models are given.

An aspect that players find important, which is not addressed by the studies cited above, is the influence of the material properties of the pad¹. Indeed pads of different materials are used, some with flat disks made of metal or plastic affixed in the middle. These are called pad “resonators” by makers and musicians. To our knowledge, the acoustic role of these “resonators” remains unknown, as there is no scientific literature on the subject. This paper aims at investigating the effective role of pad “resonators” and since there is no obvious reason to call them “resonators” we will use in the rest of the article the more acoustically neutral name “pad covers”. The study focuses on the saxophone, since it is the first instrument on which pad covers were introduced (likely due to its large holes).

In order to evaluate the influence of pad covers, measurements of the input impedance of a cylinder topped by a key with interchangeable pads (with and without resonators) are performed. Open and closed situations are investigated. In a second step, the vibration of these pads when submitted to sound pressure is measured. Impedance measurements of an entire saxophone are then made, and finally measurements in playing situation are taken in order to evaluate the perceptibility of the observed impedance differences.

2. INFLUENCE OF THE PAD ON THE IMPEDANCE OF THE HOLE

In order to investigate the role of a single pad, the input impedance of a tube of diameter equal to that of a side hole and surmounted by a key is measured. The tube is 100 mm long, with an inner radius of 12.8 mm and an external radius of 15 mm. Three identical keys of diameter

Copyright: ©2013 Pauline Eveno et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ <http://forum.saxontheweb.net/showthread.php?180766-Resonators-why>
<http://cafesaxophone.com/showthread.php?2671-Time-to-install-resonators%85but-which-ones>

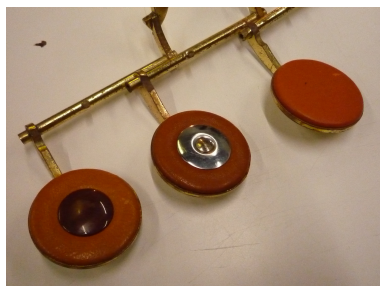


Figure 1. Three kinds of pads: with a plastic cover on the left, with a metal cover on the middle, and without cover on the right.

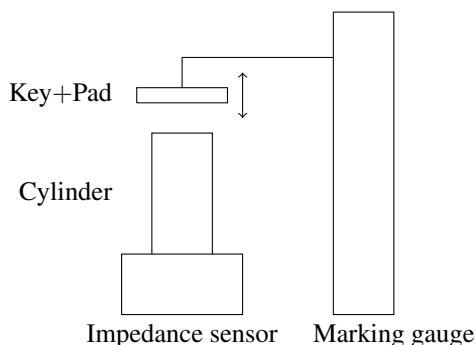


Figure 2. Technical sketch of the experimental protocol used for the measurements of the cylinder input impedance.

36 mm, provided with three different pads (see Figure 1) are measured: one with a plastic cover, one with a metal cover and one without cover. The pads consist of a cardboard covered with leather, provided or not with a cover fixed at the middle with a rivet. A marking gauge is used to move the key, whose position is measured with a tenth of a millimetre precision. Closed and open situations with different heights are investigated. A sketch of the experiment is provided in Figure 2.

2.1 Closed cylinder

The input impedance of the cylinder closed with the three types of pads is measured, from which the absorption coefficient of each type of pad can be deduced. Figure 3 shows that the pads with cover have a low absorption coefficient, circa 0.1, and that the pad without cover has a significantly higher absorption coefficient, circa 0.4. Moreover, Figure 3 shows some pad resonances. For the pad without cover a first resonance with a low Q-factor appears around 1700 Hz and a second resonance with a larger Q-factor around 2500 Hz. Pads with covers also present resonances, but they are shifted to higher frequencies. These results suggest that the input impedance of a saxophone might be significantly influenced by the presence or absence of a pad cover. This item is investigated in section 4.

2.2 Open cylinder

When the hole is open, the pad might also have an influence on the radiation impedance of the hole. Therefore,

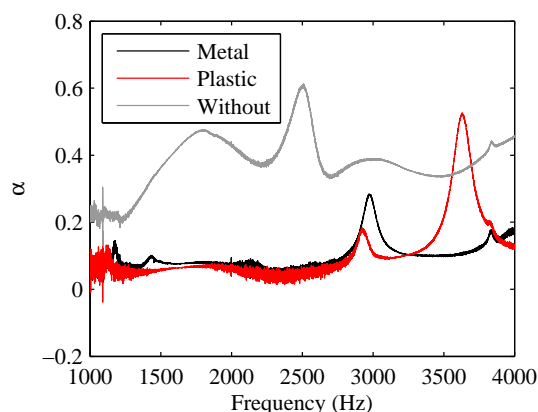


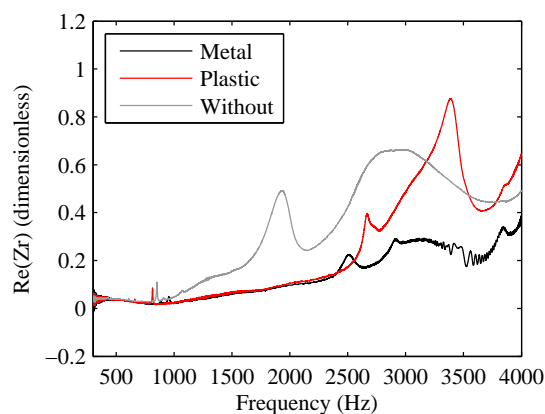
Figure 3. Absorption coefficient of the pads as function of the frequency. Pad with metal cover (in black), with plastic cover (in red) and without cover (in grey).

the input impedance of the tube surmounted by the pad is measured for different values of the distance between the tube and the key. Figure 4 shows the real part of the radiation impedance for the different pads and for two key heights: 1 and 5 mm. The effect of the pad is clearly visible when the key is close to the cylinder output. Indeed, Figure 4 (a) shows that the real part of the impedance is much higher for the pad without cover between 1000 and 3000 Hz, which means that the pad absorbs a lot of energy. In practice, a distance of 1 mm might only occur in a transitory state. The 5 mm case corresponds more where an open key is at rest. In that case, the difference in the radiation impedance is much more reduced but it still might be detectable. The question whether or not the effect is visible on a whole saxophone is investigated in section 4. Moreover, some peaks appear in the curves which seem to indicate a resonant behaviour of the pads. Vibration measurements are thus done to highlight this behaviour.

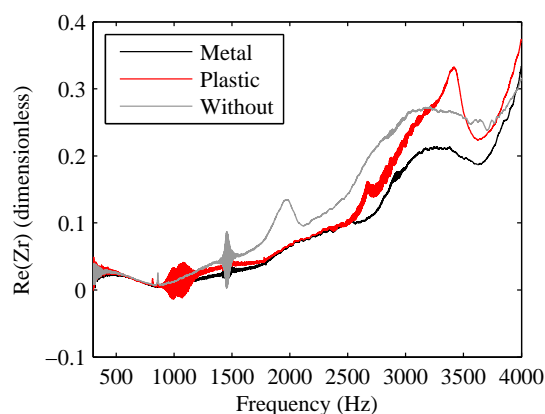
3. VIBRATION ANALYSIS

In order to understand why the damping induced by a pad without cover is so large, vibration analysis is carried out. For these measurements, as shown on the sketch in Figure 5, the key provided with the studied pad is fixed with a vice and excited by a loudspeaker with a sweep from 200 to 6000 Hz. The pressure is measured by a microphone placed near the pad and the pad vibrations are measured with a laser velocimeter in the middle and at the edge. The vibration of the key is measured by pointing the velocimeter at the reverse side of the pad, on the metallic part.

Figure 6 shows the mobility $H = \text{Velocity}/\text{Pressure}$ measured at the edge and in the middle of all the pads and the key. There are more differences between the pads in the middle than at the edge. This is obviously due to the presence of the cover in the middle of the pad. The pads with a cover have a mobility about 30 dB lower than the pad without: the cover is in fact a “stiffener”. Indeed, in Figure 6 the pad without cover has a resonance around 2000 Hz. A resonance around 3500 Hz is also visible on all curves, which seems to be related to a mechanical resonance of the



(a) 1 mm



(b) 5 mm

Figure 4. Radiation impedance of the cylinder topped by the three pads (Pad with metal cover in black, with plastic cover in red and without cover in grey) at a height of (a) 1mm and (b) 5mm.

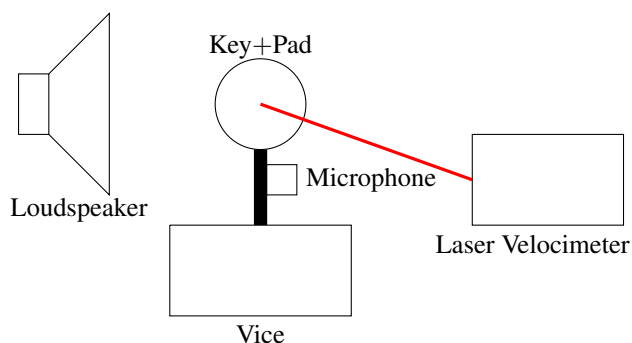
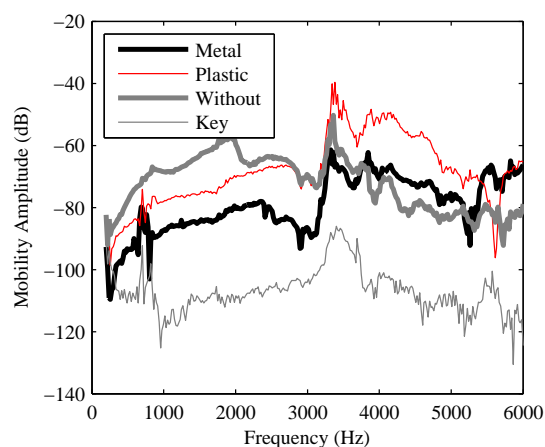


Figure 5. Technical sketch of the experimental protocol used for the vibration measurements.

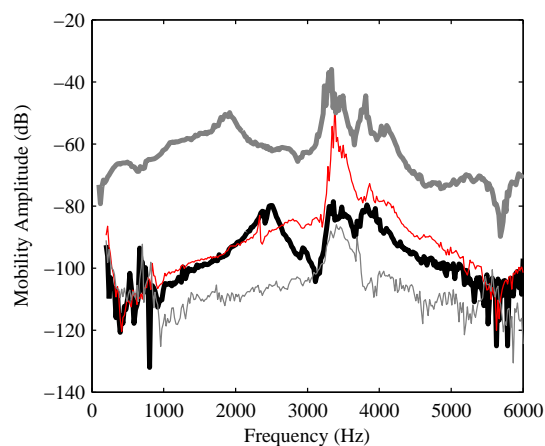
key. From the curves, it appears that the cover made with metal seems to be the most efficient “stiffener”.

4. INPUT IMPEDANCE MEASUREMENT OF THE WHOLE SAXOPHONE

The main effects of the pad covers have been highlighted, which leads to the conclusion that they are in fact stiffeners which reduce significantly the absorption induced by



(a) At the edge



(b) In the middle

Figure 6. Mobility measured (a) at the edge and (b) in the middle of the pad. Key mobility is given as a reference.

the pads. It is now interesting to evaluate the influence of such pad covers on a whole instrument, both for open and closed holes. The saxophone used in this study is a tenor Yamaha YTS 275. Its input impedance is measured for several fingerings of the low register: B \flat , C, D, E \flat , E, F \sharp (played with the side key and not with the fork fingering)². The F \sharp fingering is particularly interesting since it requires the opening of the hole whose key has a relatively large diameter (around 30 mm) and a relatively small lift height (around 5 mm). Considering the results of the previous section, vibrations of the pad above an open hole might have a measurable influence on this fingering. The experimental set-up [10] developed jointly by the LAUM³ and the CTTM⁴ is used to measure the input impedance. This impedance sensor provides measurements with a relative error of ± 3 cents [11].

The input impedances are measured in four configurations, summarized in Figure 7: (a) the original saxophone

² Note names are written, so sounding notes are one tone higher.

³ Laboratoire d’Acoustique de l’Université du Maine, Avenue Olivier Messiaen, 72085 LE MANS Cedex 9

⁴ Centre de Transfert de Technologie du Mans, 20 rue Thalès de Milet, 72000 Le Mans

(fully provided with plastic covers), (b) the saxophone with pads without covers (all the pads from the original saxophone have been removed and replaced by pads without cover), (c) the “hybrid” saxophone (the smallest pads up to the last pad of the F fingering have been replaced with the original plastic covers, the larger pads remaining the one without cover) and (d) the saxophone fully repadded with its original plastic pads. The instrument has been meticulously tuned for each configuration by a skilled worker, using in particular a leaking light method.

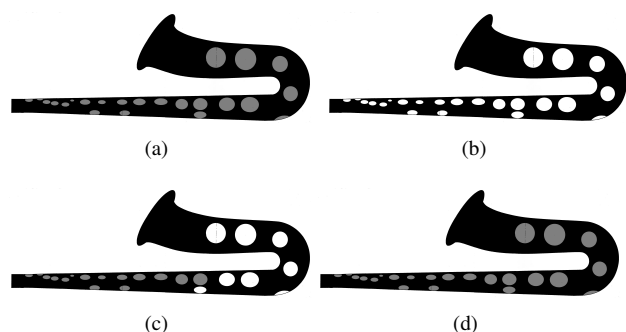


Figure 7. State of the saxophone during each series of measurements, in chronological order from (a) to (d). The pads with plastic cover are represented in grey and those without cover are given in white.

These measurements may confirm several hypothesis and effects seen on the previous sections:

- Pads without cover are more absorbent than pads with cover for closed toneholes.
- A cumulative effect on the toneholes should be observed: the absorbent effect should be more visible for the B \flat fingering, where all the toneholes are closed, and should decrease according to the number of open toneholes.
- The resonance of the pad without cover may be detected on the impedance of the F \sharp since the first raised key after the closed toneholes has a small height. That is why the “hybrid” configuration is used. It allows isolating the effect of the pad without cover located on this key, by making a comparison with the “all plastic” configurations.

The saxophone input impedance is measured at the crook input, without the mouthpiece (see Figure 8). A study [7, p. 159-162] was carried out on how the pressing force on the keys can modify the input impedance of the instrument. Differences up to 8 dB were obtained on the amplitude (for the input impedance of the same saxophone but whose fingerings are performed by different persons). Consequently, in order to have repeatable measurements, pliers made of piano wire are used to press the keys with the same strength throughout the study. Each measurement is performed three times, by taking off the saxophone from the impedance sensor (and putting it back on) and removing all the pliers (and replacing them). The repeatability

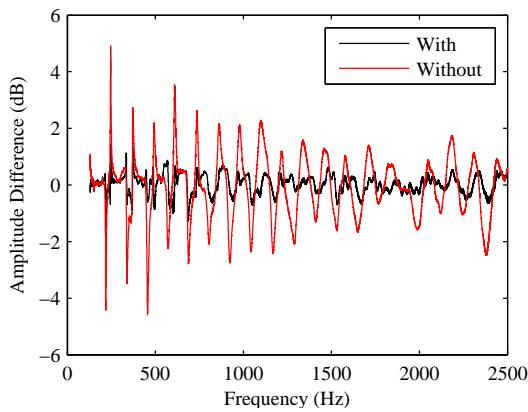


Figure 8. Photo of the saxophone input impedance measurements.

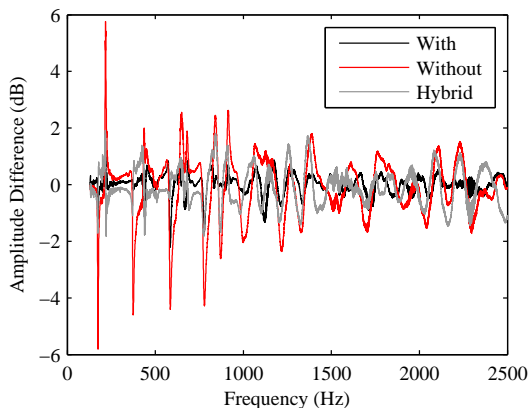
error is then estimated to be 4 cents for the resonance frequencies and 1 dB for the amplitude. Since measurements are performed over several days, a frequency correction is applied to take the speed of sound changes due to the temperature into account.

The input impedance of the saxophone with plastic covers has been measured twice: first, with the original saxophone and second, with the saxophone repadded with its original pads. These two measurements make it possible to quantify the repeatability error of the measurement combined with the effect of tuning and preparation. In the rest of the study the input impedance of the original saxophone is chosen as a reference. In Figure 9, amplitude differences between the input impedance of that reference and the input impedances of the saxophone in the other configurations are plotted for two fingerings. The black curves show that the repeatability error of the measurements is less than 1 dB and that both pad substitution and tuning do not lead to major modifications of the saxophone input impedance. Indeed, the difference between the reference and the input impedance of the saxophone without cover is much larger. This means that the pads have a visible effect on the input impedance of the whole saxophone. As predicted, this effect is cumulative, so that the difference is higher for the B \flat fingering (3 dB in average), where all the toneholes are closed, than for the F \sharp fingering (2 dB in average).

Figure 10 shows that the effect of the pads without cover, when toneholes are closed, is to lower the amplitude of the impedance peaks, without changing the resonance frequency. Clearly, when it is closed, the effect of a pad without cover is to increase the damping. On the other hand, Figure 9 (b) shows that the input impedance of the “hybrid” saxophone is closer to the input impedance of the saxophone with covers. Indeed, the radiation mostly takes place at the first open tonehole, the following toneholes have practically no influence on the input impedance. So, despite the fact that the influence of the open hole without pad cover is significant, this effect is in practice negligible compared to that of the closed holes.



(a) B \flat



(b) F \sharp

Figure 9. Difference between the input impedance (in dB) of the original saxophone and the input impedance (in dB) of the saxophone without cover (in red), of the hybrid saxophone (in grey) and of the saxophone repadded with its original pads (in black) for two fingerings: (a) B \flat and (b) F \sharp .

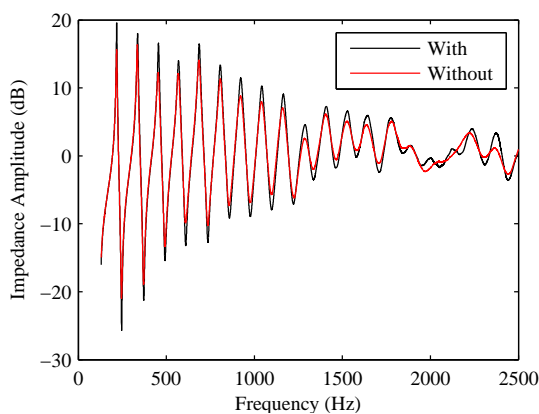


Figure 10. Comparison between the input impedance of the saxophone with plastic covers (in black) and the impedance of the saxophone without cover (in red) for the B \flat fingering.

5. PAD EFFECT IN PLAYING SITUATION

After having characterised the effect of the pads on the saxophone with an objective criterion, the input impedance, it is interesting to determine their influence in playing situation. For these measurements, two “identical” saxophones are needed so that one is kept in its original padding throughout the whole study, as a reference.

5.1 Choice of the saxophones

The choice of the two closest saxophones is first done by the musician among three saxophones of same design: tenor Buffet Crampon Evette provided with metal covers. These three saxophones are named by the last two numbers of their serial number: 65, 72 and 81. The choice is made by doing a blind pair-wise comparison: the musician plays two saxophones in a row without knowing their number and has to give a rating between 0 and 4 to quantify their difference (0 for very similar and 4 for very different). This exercise is repeated several times for all the possible pairs and the results are given by the matrix of dissimilarity on Table 1. The saxophone 65 stands out from the two others as some of its notes are difficult to play.

	65	72	81
65	0/1	3/4	3/4
72		0/1	1/2
81			1/2

Table 1. Matrix of the dissimilarity between the three saxophones.

Input impedance measurements on these saxophones confirm the musician’s choice. Indeed, Figure 11 presents the difference between the input impedance amplitude of each pair of saxophone for the B \flat fingering. Differences are much bigger for the pairs involving saxophone 65 than between saxophones 72 and 81 (only 1 dB in average), which confirms the musician’s choice. It was therefore decided to keep saxophones 72 and 81 for the rest of the study.

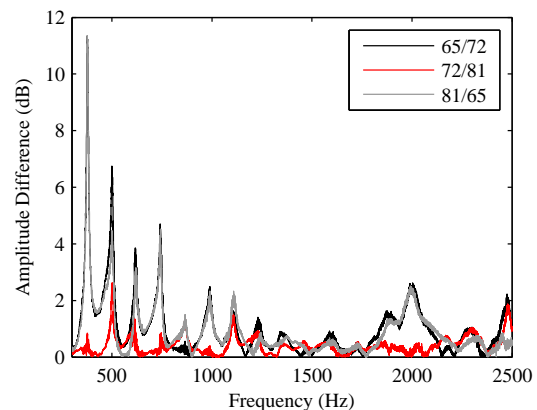


Figure 11. Differences (in absolute value) between the impedances (in dB) of each saxophone pair: 65 and 72 (in black), 72 and 81 (in red), 81 and 65 (in grey) for the B \flat fingering.

Descriptor	Description	ANOVA p-values
Attack Time (AT)	Time of the transient regime	0,37: n.s
Harmonic Spectral Centroid (HSC)	Frequency of the centroid of the harmonic spectrum	< 0,0001: sig
Odd/Even Ratio (OER)	Energy ratio of odd harmonics to even harmonics	0,79: n.s
Tristimulus, 1st coefficient (TR1)	Energy ratio of the fundamental component to the total	0,72: n.s
Tristimulus, 2nd coefficient (TR2)	Energy ratio of harmonics 2, 3 and 4 to the total	0,18: n.s
Tristimulus, 3rd coefficient (TR3)	Energy ratio of higher-order harmonics to the total	0,31: n.s
Threshold pressure (TP)	Overpressure in the mouth that makes oscillate the reed	0,53: n.s
Mouth pressure (MP)	Mean pressure in the mouth of the musician	0,004: sig
Pressure level (PL)	Level of pressure at the bell	0,40: n.s
Efficiency (E)	Ratio of the Pressure level to the Mouth pressure	0,70: n.s

Table 2. List of the descriptors and their ANOVA p-values for the effect of the pads (significant threshold: 5%, n.s : non significant, sig : significant)

5.2 In vivo measurements

The purpose of these measurements is to characterise the effect of the pads on the musician's playing parameters and on the radiated sound. The pressure in the musician's mouth and the radiated pressure at the bell are measured while 9 notes are being played (high register C \sharp , A, F \sharp , E \flat and low register C \sharp , A, F \sharp , E \flat and B \flat) on the two saxophones. These recordings are realised five times in order to get a meaningful average. Then, the saxophone 72 is kept in its original shape and the 81 is provided with pads without cover for another series of recordings.

The pressure in the mouth is measured with a differential pressure transducer Endevco 8507-C2. The radiated pressure at the bell is recorded with a microphone located 10 cm from the bell in the axis of the instrument, as shown in Figure 12 (see also [12] for more photos of the experimental protocole). The data are collected using a National Instruments BNC-2110 acquisition board, with a 50 kHz sampling frequency.



Figure 12. Photo of the in vivo measurements.

Several physical descriptors characterising the timbre, the radiated sound and the musician's way of playing, have been considered. These are listed in Table 2. All these descriptors, except the attack time and the threshold pressure, are averaged on the stationary part of the signal. Then, the values of the descriptors are averaged on the five trials.

The influence of each descriptor on the distinction between the saxophones can be modeled using the analysis

of variance method (ANOVA) [13]. Table 2 shows that only two descriptors can distinguish the saxophones with or without covers: the HSC and the mouth pressure. Saxophone without cover has globally a higher HSC⁵ (HSC = 9.81) than a saxophone with covers (HSC = 8.63). The mouth pressure is in average 3640 Pa for a saxophone without cover and 3280 Pa for a saxophone with covers for an average pressure level at the bell around 10 Pa for both. This result is consistent with the input impedance analysis: indeed, the pads without covers tend to lower the amplitude of resonance peaks; more energy is thus needed to obtain the same output pressure provided by a saxophone with covers⁶. Nevertheless, these measurements are not sufficient to draw a conclusion regarding the influence of the pads on the radiated sound and on the playing parameters. It is only a preliminary study that need to be verified with more measurements. Moreover, since the pressure level remains the same with the two types of saxophone, it is possible that the musician compensate by adjusting his vocal tract. Therefore, measurements should be done with an artificial mouth to avoid this effect.

6. CONCLUSION

The covers have a measurable effect on the acoustical characteristics of the saxophone. Their main role is to stiffen the pad, so that the cover should in fact be called "stiffener". The reflection coefficient is increased by the presence of a cover when the tonehole is closed and the amplitude of the saxophone input impedance peaks is consequently increased of several dB. The effect appears to be greater with more closed tone holes. Measurements with a musician seem to confirm these results. Indeed, the musician needs, for obtaining a given output pressure, to produce a higher mouth pressure when the saxophone pads do not have a cover. It has been observed that pad vibrations can influence the acoustic radiation coming out of open toneholes. Nevertheless, this effect is small and is significant for small key heights only. Even if this effect has

⁵ The HSC is dimensionless, divided by the fundamental frequency.

⁶ $20 \log 3640/3280 = 0.9$ dB, this is a little less than the 2 or 3 dB differences found in section 4. The mouth pressure is indeed averaged for different fingerings, including particularly the C \sharp fingering where all the holes are open and where the influence of the covers is therefore quite negligible.

been effectively measured in some configurations, the impact of pad vibration is negligible compared to that of pad adjustment influencing leaking and tuning. The influence of pad resonators on saxophone timbre remains questionable. A higher HSC has been measured on the radiated sound of a saxophone without cover and there is a bigger pressure in the mouth of a musician playing a saxophone without cover. In vivo measurements need to be carried on in order to confirm or deny these results. Several musicians and more saxophones are required in order to make a more reliable study. Similar preparation of the saxophones are mandatory. Measurements should also be carried out with an artificial mouth in order to avoid the effect of the musician. The future work should include a perceptive study on the timbre and the ease of playing. This might be completed by a listening test on the recorded sounds.

Acknowledgments

This research was funded by the French National Research Agency ANR within the PAFI project (Plateforme d'Aide à la Fabrication Instrumentale in French). The authors would like to thank all participants for all discussions and interesting work done together. Many thanks to M. Ferreira and D. Beuvry from ITEM for all the re-padding work, to B. Gazengel from LAUM for the in vivo measurements, to J.-F. Petiot from IRCCyN for the statistical analysis and to V. Yaremchuk and B. Lopatin for proofreading.

7. REFERENCES

- [1] A. Lefebvre, "Computational acoustic methods for the design of woodwind instruments," Ph.D. dissertation, Computational Acoustic Modeling Laboratory, McGill University, Montreal, Quebec, Canada, 2010.
- [2] D. Keefe, "Theory of the single woodwind tonehole," *J. Acoust. Soc. Am.*, vol. 72, no. 3, pp. 676–687, 1982.
- [3] C. J. Nederveen, J. K. M. Jansen, and R. van Hassel, "Corrections for woodwind tone-hole calculations," *Acustica*, vol. 84, pp. 957–966, 1998.
- [4] V. Dubos, J. Kergomard, A. Khettabi, J.-P. Dalmont, D. H. Keefe, and C. J. Nederveen, "Theory of sound propagation in a duct with a branched tube using modal decomposition," *Acustica*, vol. 85, pp. 153–169, 1999.
- [5] J.-P. Dalmont, C. J. Nederveen, and N. Joly, "Radiation impedance of tubes with different flanges : numerical and experimental investigation," *J. Sound Vib.*, vol. 244, no. 3, pp. 505–534, 2001.
- [6] P. Dickens, "Flute acoustics: measurement, modelling and design," Ph.D. dissertation, School of Physics, University of New South Wales, 2007.
- [7] P. Eveno, "L'impédance d'entrée pour l'aide à la facture des instruments de musique à vent : mesures, modèles et lien avec les fréquences de jeu (the input impedance for the support of the musical instruments making: measurements, models and link with the playing frequencies)," Ph.D. dissertation, Univ. Pierre et Marie Curie (Paris VI), 2012. [Online]. Available: <http://goo.gl/RvujY>
- [8] M. Curtit, P. Bolton, and F. Masson, "Accord d'un instrument à vent : quelques stratégies du facteur illustrées par une analyse de l'impédance d'entrée (tuning of a wind instrument: some makers' strategies illustrated by an input impedance analysis)," in *Proceedings of the 10th Congrès Français d'Acoustique*, Lyon, 2010.
- [9] A. Chaigne and J. Kergomard, *Acoustique des instruments de musique*, Belin, Ed., 2008.
- [10] J.-C. L. Roux, J.-P. Dalmont, and B. Gazengel, "A new impedance tube for large frequency band measurement of absorbing materials," *Proceedings of Acoustics'08*, 2008.
- [11] C. A. Macaluso and J.-P. Dalmont, "Trumpet with near-perfect harmonicity: design and acoustic results," *J. Acoust. Soc. Am.*, vol. 129, no. 1, pp. 404–414, 2011.
- [12] B. Gazengel and J.-P. Dalmont, "Mechanical response characterization of saxophone reeds," in *Proceedings of Forum Acusticum*, Aalborg, Denmark, 2011.
- [13] H. Scheffé, *The Analysis of Variance*. New York: Wiley, 1959.

A Study of Sound Characteristics of a new Bassoon as compared to modern German Bassoon

Timo Grothe

Hochschule für Musik Detmold,
Erich-Thienhaus-Institut
grothe@hfm-detmold.de

Peter Wolf

Guntram Wolf Holzblasinstrumente GmbH,
Kronach, Germany
peter.wolf@guntramwolf.de

ABSTRACT

An interdisciplinary team of instrument makers, engineers and musicians has recently developed a new bassoon-like instrument. Projected fields of use for this instrument are ensembles, in which bassoonists play together with brass players. In order to overcome limitations in the dynamics of the modern bassoon and to achieve a brighter tone-color, a drastically re-designed air column is introduced. Compared to the bassoon, the new instrument has a larger taper which is constant along the main bore, as well as wider and shorter tone holes, which are entirely operated by keys. A chromatically playable prototype with full keywork has been built. The new instrument has powerful bass, an enriched overtone spectrum and plays significantly louder than the modern bassoon, so we call it *Basso forte*.

In this comparative study we analyze sound recordings of the modern German bassoon and *Basso forte* in a reverberation chamber and in a recording studio, and provide some considerations on its acoustical design based on input impedance curves. Our results indicate, that the cutoff-frequency of the air column is correlated with formant frequencies in the sound spectrum. This suggests an optimization target for reconsiderations of the tone hole concept, to overcome an observed unevenness in timbre.

1. INTRODUCTION

The generic design of the bassoon has barely changed in the past 80 years. Several construction peculiarities of modern bassoons are more or less remains of its historical predecessors. Among these are long and narrow finger-holes, obliquely to the main axis, as well as many local variations in taper of the bore. Being a fully developed orchestral instrument, today's modern German bassoon has a limited dynamic range due to its specific design.

1.1 Idea and requirements

It was the idea of the woodwind-instrument maker Guntram Wolf and one of the authors (P.W.) to develop a "wind

band bassoon" which in volume and tone color would be able to compete with the brass in symphonic wind bands. Embarking on such a new design also meant that acoustic limitations could be addressed from the outset, in order to reduce some of the bassoon's familiar weaknesses of intonation. Abandoning the Almenräder/Heckel system offered the opportunity of replacing it with a new and coherent approach. Therefore a team of scientists and instrument makers revived a principle of design which Theobald Böhm had also worked on adapting for the bassoon. This had inspired some quite promising designs by other instrument makers in the late 19th century but these were seen as being in competition with the bassoon, and therefore had plainly not been found convincing in terms of sound. The following profile of requirements should be met by the new double reed instrument suitable for wind bands, which was to design. Compared with the modern bassoon it should

- be capable to cover a larger dynamic range,
- have even and stable intonation,
- have a handling familiar to bassoonists, and
- be comparable in sound color.

At the heart of this development is a new body constructed according to acoustic rules.

1.2 A new bore design

Greater dynamic range and tonal volume should be achieved by widening the conical bore, enlarging its taper and keeping it constant along the main axis. Additionally, the range of the instrument was extended in the bass by a semitone down to A1 ($f_0 = 55$ Hz). That obviously necessitated a completely new arrangement of the tone holes.

Various tone hole configurations have been calculated by means of a one dimensional air column model [1,2]. Transmission line elements for conical duct segments and tone holes have been implemented in MATLAB. For a given main bore, the tone hole positions were obtained in a nested optimization procedure, with the aim of matching the fundamental air column resonance with a nominal frequency for each fingering. A second design goal concerned the harmonicity of higher peaks.

Prediction of sounding frequencies of reed instruments from the geometry of the resonator is possible to some extent,

if the sound generating reed mouthpiece is taken into account by means of an equivalent volume and further corrections [3]. This approach is based on a simple physical model [4] and may be source of crucial errors, especially if the reed design is not known a priori, as in our case. We tackled this problem by analyzing the acoustics of modern German bassoons through impedance measurements and extensive investigations of the bassoon reed and embouchure [5]. As many aspects are not simple to confer to a new air column, we acknowledge a major contribution in the tone-hole design process by the experience of one co-operating instrument maker¹.

1.3 User-centered development

Tone-hole concepts were tested for two main bores of different taper by means of functional playable models, manufactured as wooden half-shells with a cnc-router. These were evaluated in playing experiments with the help of a professional bassoonist and reed maker². As the generic sound color is believed to depend mostly upon the resonator, blowing tests have been carried out with a tentative bocal. The preferable design for the wooden corpus was chosen by rating its sound and dynamical characteristics. Subsequently, the length and inner bore profile of the bocal was determined by means of the transmission line optimization routine. On a playable functional model with a simplified keywork, the bocal design could be confirmed and the reed design was refined.

From the scientific point of view, this evolution may seem upside down, however it is owed to the wish of involving the experience of musicians at a most early stage in the design process: For musicians, it is very difficult to judge inherent musical characteristics like intonation, dynamics and timbre by performing on a merely technical device like the functional models. The assessment of performance characteristics with artificial mouths provides valuable generic insights, but many details, which may be important to musicians in a multi-modal perception context, are probably not yet accessible by measurements.

We consider this user centered design-approach to be reasonable and decided to built a fully playable prototype with a complete keywork, being aware that the evaluation may require further modifications of the resonator. Figure 1 (a) shows the mark I instrument³ aside to a usual bassoon. It is called *Bassofoorte*, and we hope to refine woodwind design principles by learning from its acoustical properties as observed in the musical performance.

To adress the sound characteristics in real playing conditions, standard acoustic measurement techniques and signal processing algorithms have been used and will be presented here. The aim of the present comparative study is to report differences between the modern German bassoon and the *Bassofoorte*, and to discuss these and some informal impressions reported by testers with respect to the characteristics of the resonator, as represented by acoustic



Figure 1: (a) Modern German Bassoon (left) and the first prototype of the *Bassofoorte* (right)
(b) corresponding double-reed mouthpieces

impedance curves.

2. MATERIALS AND METHODS

The investigation comprises calculations of acoustic input impedance, sound pressure measurements in a reverberation chamber, and studio recordings. In the latter two measurements a student (reverberation chamber) and a professional musician (studio) played both a modern German bassoon and the *Bassofoorte*. Both players used their own reed and instruments, which were an old Püchner, with a Heckel bocal C1, and a Wolf S2000 plus E⁴, with a Wolf bocal T8 F9-3, respectively.

2.1 Acoustic Impedance Calculations

The linear acoustic properties of an air-column are described by its acoustic input impedance $Z = P/U$ where P and U are the pressure and acoustic volume flow, respectively, in the frequency domain.

The Transmission-Line Method (TLM) has been proven to be appropriate to model the quasi one-dimensional wave acoustic propagation in ducts with sideholes and thus to calculate the acoustic impedance at any cross section of the main duct [6,7]. Impedance curves have been measured on both bassoon and *Bassofoorte* with a commercially available impedance measurement system "BIAS"⁵, designed for brass instruments measurements. Using a self-made

¹ Benedikt Eppelsheim

² Stefan Pantzier

³ built at Guntram Wolf Holzblasinstrumente, Kronach, Germany

⁴ plus E: manufactured from yew (*taxus baccata*)

⁵ Fa. Artim, Wien, Austria

connection adapter for bassoon bocal, we achieved an acceptable agreement with impedance curves obtained from TLM calculations, namely a good agreement in peak (and dip) frequencies, but a significant discrepancy in impedance magnitudes. The most probable reason for the experimentally observed increase in damping is an imperfection in our measurement, although it should be noticed that the same effect has been obtained earlier with a different measurement device specifically designed for bassoons [8].

For a discussion of harmonicity of woodwind resonators a great precision of the impedance curve is crucial. Due to the possible imperfections in our measurement setup, we rely on careful geometric measurements of the two instruments⁶ and base the comparison of bassoon and *Basso-forte* in terms of harmonicity upon numerical impedance curves.

To quantify the alignment of air column resonances with the perfectly harmonic partials in the sound, a *harmonicity descriptor* HD is introduced. This scalar bases on a weighting function $w(n)$ of Gaussian type to rate the frequency deviation between the n^{th} air column resonance f_n and the closest harmonic $k f_0$ in Cent, according to

$$w(n) = e^{-\frac{1}{2} \left(\frac{2 \log_2 \left(\frac{f_n}{k f_0} \right)}{\sigma} \right)^2}. \quad (1)$$

Note that k may be different from n . The steepness of the weighting function is defined by the parameter σ , which is set to $\sigma = 0.14$, i.e. a frequency shift of an air column resonance by ± 100 Cent (1 semitone) with respect to an integer multiple of f_0 will be weighted by $w = 0.5$.

The harmonicity descriptor is defined as

$$HD_N = \frac{1}{N} \sum_{n=1}^N w(n), \quad (2)$$

where N is the number of air column resonances taken into account.

2.2 Reverberation Chamber Measurements

Under playing conditions the bassoon has a complex radiation pattern, due to the distribution of tone holes across the long and slim corpus. The same applies to the *Basso-forte*, but the radiation pattern may be different in detail, since the tone holes are shifted and pads are hanging above all of them. To avoid variance in acoustic measurement data due to these effects, blowing experiments with a player were carried out in a reverberation chamber. Assuming an ideally diffuse sound field, the measurement result is independent of the relative position of source and sensors. Under the same assumption, the measured sound pressure level L_p is proportional to the radiated sound power P_{ac} .

The reverberation chamber at the Erich-Thienhaus-Institut has a volume of $V_R = 127 \text{ m}^3$, a surface area of $A_R = 178 \text{ m}^2$, and an equivalent absorptive area of $\alpha A_R = 2 \text{ m}^2$. The reverberation time T_{60} is approximately 10 seconds within the frequency of 150 to 600 Hz. In order to meet

the assumption of a diffuse sound field, the position of the source (the player⁷) and the sensors (Brüel&Kjær 4190, 4938) have been changed randomly, and each note has been repeated several times, to achieve ten independent measurements per note. To detect the upper limits of the dynamic range, the player was instructed to play as loud as possible.

2.3 Studio Recordings

Finally, the sound of both instruments has been recorded in a professional studio, by a professional sound engineer⁸, with a 5 microphone setup typical for bassoon recordings. It included an ORTF stereomicrophone (Schoeps MSTC 64 U) at $+45$ degrees at a height of 3 m (M1), a cardioid microphone (Neumann U87) at 0 degrees (in front of the player) at a height of 1.3 m (M2), 2 boundary layer microphones (Neumann GFM132) approximately ± 30 degrees (left(-) and right(+)) of vertical main axis of the players head, at a distance of 0.6 m and 1.5 m (M3 and M4), and a subcardioid microphone (Schoeps MK21) at -60 degrees at a height of 2.5 m pointing on the bell of the instrument (M5). The instruments were played by a professional bassoonist⁹, who was seated and advised to avoid position changes during the recordings. Two short pieces in a slow tempo have been recorded using a metronome to keep the timing constant. One piece was a scale practice and the other was a theme of Mussorgsky's *Pictures at an Exhibition*. Sound snippets of constant pitch with a minimum length of 125 ms have been extracted from the audio data. The microphones were not calibrated.

3. RESULTS

3.1 Air Column Resonances

The comparison of input impedance spectra shown in Fig. 2 reveal for all fingerings a slight shift of the *Basso-forte* curve to lower magnitudes, which corresponds to the larger input cross section of the *Basso-forte* bocal. In other words, when excited with a volume-flow excitation of constant amplitude, a reduced pressure amplitude acts at the input end of the *Basso-forte*, where the reed is connected to the resonator. The new instrument is played with considerably larger reed¹⁰, as shown in Fig. 1 (b). To model the compliant influence of the reed, an equivalent volume of 1.8 cm^3 and 3.4 cm^3 , respectively has been numerically "attached" to the resonators of bassoon and *Basso-forte* for the impedance calculations in Fig. 2.

In terms of harmonicity we find a better alignment of the first four air-column resonances with the partials in the sound for the *Basso-forte* compared to the modern German Bassoon. Fig. 3 shows, that for first mode fingerings the harmonicity descriptor HD_4 (Eq. 2) is larger on the new instrument.

⁷ Hannes Fritsch, Hochschule für Musik, Detmold, Germany

⁸ Maximilian Pauls, Hochschule für Musik Dresden, Germany

⁹ Stefan Pantzner, MDR Sinfonieorchester, Leipzig, Germany

¹⁰ Reed designs comparable to that of historical bassoons from the classical period have been found to be suited.

⁶ The *Basso-forte* is compared to a bassoon Wolf S2000

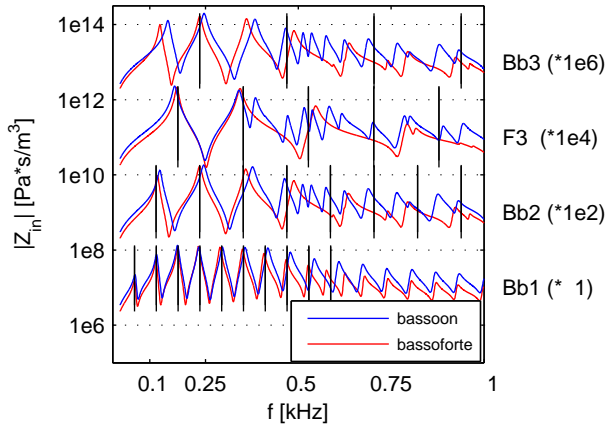


Figure 2: Acoustic impedance curves of 4 notes Bb1, ($f_0 = 59$ Hz); Bb2, ($f_0 = 117$ Hz); F3, ($f_0 = 175$ Hz); Bb3, ($f_0 = 235$ Hz) for bassoon and *Bassoforte*. TLM calculations with a reed equivalent volume of 1.8 cm^3 (bassoon) and 3.4 cm^3 (*Bassoforte*). The curves are shifted vertically for better readability of the plot. Vertical black lines mark the partials of the sounded note.

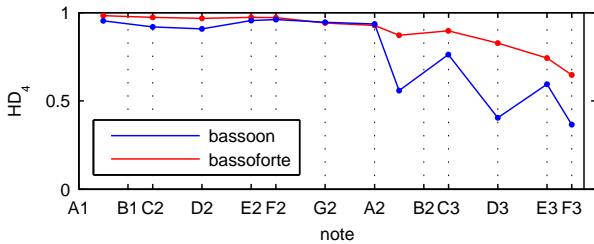


Figure 3: Harmonicity of several first mode fingerings of bassoon and *Bassoforte* calculated from the first 4 air column resonances according to Eq. 2.

3.2 Dynamics

The reverberation chamber measurements have been analyzed by means of standard algorithms to calculate the power spectral density (Fig. 4), and loudness spectra (Fig. 5). From the integral results shown in Table 1 it can be seen, that the sound pressure levels do not differ largely, which indicates that the radiated sound power is approximately the same for both instruments.

From the subjective sound impression this is a bit surprising; further investigations are needed to quantifying the power feed of the player into the instrument. The loudness quantifies the perceived dynamic level. These results mirror the intuitive impression, that the dynamic level is significantly increased for the *Bassoforte* sounds. Furthermore, the measurements for both sound level and loudness indicate, that the maximum sound level increases with pitch.

3.3 Sound Spectra

For the sound comparison of bassoon and *Bassoforte*, formants have been detected. The un-calibrated audio data from the studio recordings were analyzed with the pho-

Table 1: Sound pressure level and loudness of bassoon and *Bassoforte* notes played in a reverberation chamber

	note	Bb1	Bb2	F3	Bb3
L_p [SPL dB(A)]	bassoon	88.9	93.7	95.1	96.1
	bassoforte	92.6	92.6	95.2	97.7
N [Sone]	bassoon	68.7	78.2	87.3	92.3
	bassoforte	75.7	82.0	97.7	106.1

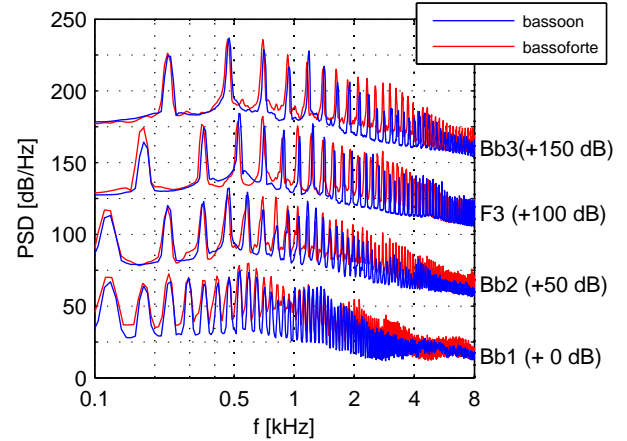


Figure 4: Sound spectra of 4 notes Bb1, ($f_0 = 59$ Hz); Bb2, ($f_0 = 117$ Hz); F3, ($f_0 = 175$ Hz); Bb3, ($f_0 = 235$ Hz) played on bassoon and *Bassoforte*. The curves are shifted vertically for better readability of the plot

netics software "praat"¹¹. Formants are broad elevations in the spectral envelope whose positions on the frequency axis are independent of the pitch of the stimulus. Formants can be obtained by linear predictive coding or from long-term averaged spectra. The absolute and relative center frequencies of formants and their bandwidths can be used for vowel detection in human speech, so they can be regarded as attributes to the timbre of a sound.

It is well-known that pronounced formants are present in the sound spectra of bassoon tones [9, 10]. In our recordings, three formants $F1$, $F2$ and $F3$ have been detected in the frequency range below 3 kHz, their center frequencies are located at 500 Hz, 1.2 and 2 kHz, which is in agreement with the literature [10]. While the measured $F1$, $F2$ and $F3$ frequencies on the bassoon are merely constant for all measured tones, the formants in the *Bassoforte* vary with the pitch (Fig. 6). In particular, $F1$ is increasingly shifted to higher frequencies for the so-called "short notes" at the upper end of the first register (B2, C3, D3, E3). At the register border, $F1$ almost reaches 700 Hz, falling again for the overblown notes above $F3$ ($f_0 = 175$ Hz) in the second register. Despite a much larger spread of values, the same trend can be observed for $F2$ and $F3$.

The spectral centroid is another spectral parameter which correlates with the timbre impression of brightness. The MATLAB algorithms implemented for the MPEG-7 standard¹² have been used to extract timbre descriptors from

¹¹ <http://www.fon.hum.uva.nl/praat/>

¹² ISO/IEC 15938

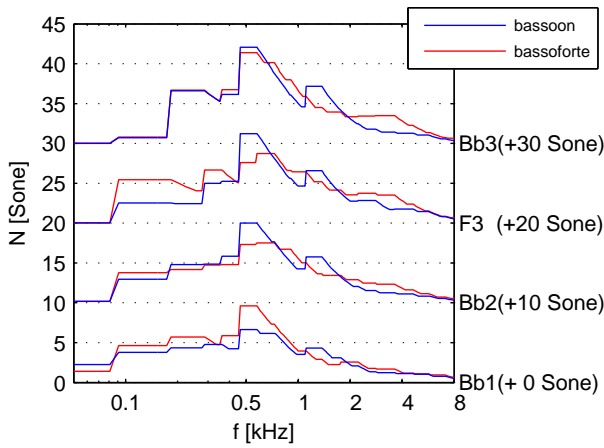


Figure 5: Loudness spectra of 4 notes Bb1, ($f_0 = 59$ Hz); Bb2, ($f_0 = 117$ Hz); F3, ($f_0 = 175$ Hz); Bb3, ($f_0 = 235$ Hz) played on bassoon and *Bassoforte*. The curves are shifted vertically for better readability of the plot

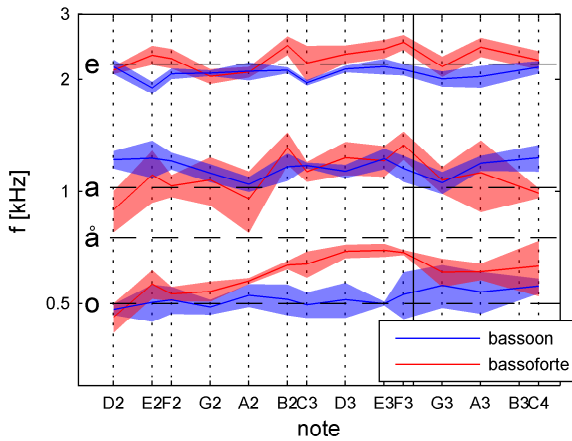


Figure 6: Formant center frequencies $F1$, $F2$, and $F3$ in the sounds of bassoon and *Bassoforte*. The shaded regions mark the scattering of observed center frequencies at 5 different microphone positions. The dashed horizontal lines mark the center frequencies of the vowels *o*, *a*, *e* [10].

the audiodata. For both bassoon and *Bassoforte*, the spectral centroid increases with pitch (Fig. 8). It is further observed, that the harmonic spectral centroid differs largely in the microphone signals. For the same note, the differences in spectral centroid reach up to 200 Hz when comparing between microphones, which is comparable to the largest differences found between the two instruments comparing within microphones. Notice, that the spread of values for the harmonic spectral centroid is much larger than for the detected $F1$ formant frequencies (Fig. 7). This means that the formant frequencies are relatively independent on the recording situation, which simplifies timbre investigations.

Comparing bassoon and *Bassoforte*, the mean values of the spectral centroid between do not show significant differences, expect for the "short-notes". Possibly, the $F1$ shift observed for this notes is the reason; the result is an in-

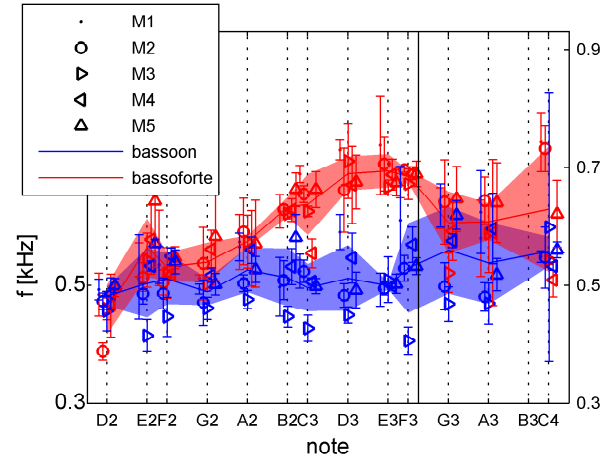


Figure 7: Center frequency of the first formant in the sounds of bassoon and *Bassoforte*. The shaded regions mark the scattering of observed center frequencies at 5 different microphone positions (M1-M5). The vertical bars mark the spread due to dynamical and temporal variations in the sound recordings.

creased brightness of these sounds.

For all notes, the sound spectra of the *Bassoforte* reveal a dedicated energy boost in the frequency band between 2 and 4 kHz, which includes both formants of the vowels *e* and *i*. This can be observed in the harmonic sound spectra of the single notes (Fig.4) and even more clearly in the long-term averaged spectra shown in Fig. 9. The latter graphs clearly shows, that this energy boost is not an issue of the recording situation or the radiation pattern, as it is observed to an equal extent in all 5 microphone positions. The same holds for an energy boost in the low frequency range indicating a reinforcement of the fundamentals which is known to produce a characteristic keyed sound color.

4. DISCUSSION

4.1 Intonation

The Bassoforte has a geometrically regular grid of wide tone holes with short chimneys, which assures a regular alignment of air column resonances. It was one of the design goals, as it is intuitive that stable oscillations are built up, if the harmonics to the sounding fundamental frequency are supported by well-aligned higher air column modes. While the measurement of the passive linear behavior of an acoustic wind instrument resonator is as well-defined as the pitch of a sound, an objective assessment of the intonation quality of a reed wind instrument is not easy. The reed and embouchure have to be taken into account. This is still a difficult task to tackle both in modeling and in the experiment. Theoretical physical models become very complex if the reed and player interaction has to be modeled. With experimental approaches using an artificial mouth, it is often unclear whether simplified technical embouchures are able to mimic the natural embouchure of a musician with the needed level of detail. Moreover, in-

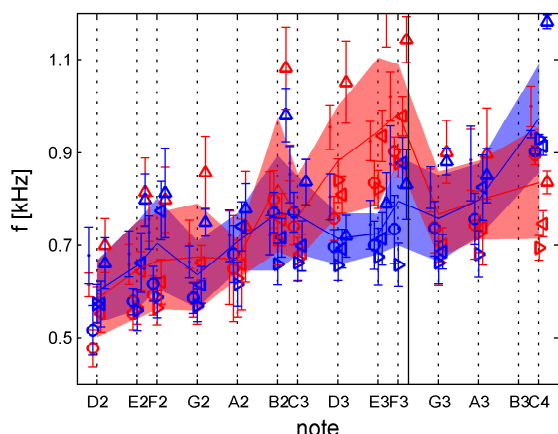


Figure 8: Harmonic spectral centroid of several sounds of bassoon and *Bassoforte*. The marker correspond to the 5 different microphone positions (M1-M5); the legend of Fig. 7 also applies to this plot.

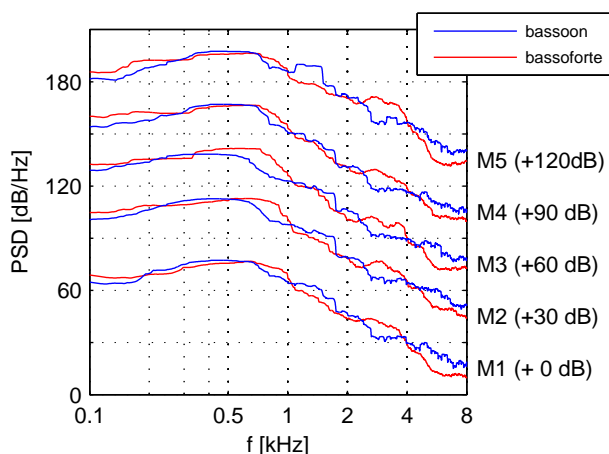


Figure 9: Long-term averaged spectra from bassoforte and bassoon sounds, recorded at 5 different microphone positions (M1-M5).

tonation quality cannot be rated by one scalar value, e.g. a frequency ratio of actually sounding and nominal pitch, because a reasonable bending range should be accessible for a reed-wind player musician. In a study on bassoon embouchures, pairs of values of blowing pressure and lip force have been measured, which were able to produce the nominal pitch on various bassoon fingerings using an artificial mouth [5]. It is obvious, that the desired dynamic level, and thus the blowing pressure, must be taken into account in the discussion of intonation quality.

No scientific attempts have been made yet, to investigate the tuning properties of the *Bassoforte* under playing conditions. However, informal playing tests have been carried out with various musicians: Players of the *Bassoforte* report, that the instrument is tuned very well and that its intonation is "stable" and "safe". This means, that all fingerings produce the expected note when blowing, and that the respective nominal sounding frequency lies approximately in the middle of the bending range available for this note.

Moreover, the bending ranges are small enough and scales can be played without the need of unexpectedly large and irregular embouchure corrections.

4.2 Ease of playing

The downward shift of the impedance curve explains the experience of the players who feel that they have to provide more air-flow when playing the *Bassoforte*. But still sound production is far being from exhaustive on this instrument. On the contrary, the impression of players is, that it is easy to blow, provides a small "resistance" and is much more "responsive" than the bassoon. We can possibly attribute these impressions to the increased harmonicity of the resonator which yield a stabilization of oscillation regimes. Besides the increased radiation efficiency at the enlarged bell diameter and tone-holes, this may also contribute to the increased loudness of the *Bassoforte*.

4.3 Sound Color

While the regular tone hole grid led to a very positive rating concerning the intonation quality, the sound color of the instrument is somewhat inhomogeneous. This can be immediately seen from the formant frequencies, shown in Fig. 6. Ascending the scale from the bass the impression of the sound color becomes somewhat "nasal" at the upper end of the first register. This is especially noticeable for the so-called "short notes" (B2,C3,D3,E3), for which F1 is largely increased.

For the bassoon, F1 is located at 500 Hz, which is comparable to the formant of the human voice when articulating the vowel *o*. Due to this analogy it can be well understood, that the bassoon sound often is described as a "warm", "dark" and "muffled". An interesting finding [10, 11] is, that for historical bassoons, the frequency of the main formant F1 is higher than in modern bassoons. In comparison, the sound of baroque bassoons is described as more "open", a bit "nasal" and with less sonority. The *Bassofortes* F1 frequency for the short notes approaches 700 Hz, which is the formant frequency of the vowel *ä*. Considering that the energy boost of the *Bassofortes* sounds between 2 and 4 kHz falls into the formant regions of the vowels *e* (2.2 Hz) and *i* (3.2 kHz) it is likely that the F1 shift leads to a "nasal" timbre for these notes. The energy boost at 2 to 4 kHz however may help the players of the *Bassoforte* to compete with the bright sound of the brasses.

Our results reveal an interesting correlation of the cut-off frequency of the air-column and the center frequency of the main formant. For the notes with a shifted F1 the air-column cutoff-frequency is also increased (Fig. 10). The fact that F1 immediately drops by 100 Hz when switching to the second register strongly supports our hypothesis, that the formant(s) in the sound depend upon the cut-off frequency(ies). The importance of the cutoff frequency on the timbre has already been emphasized by Benade [12], but was not studied in a systematic way. The present results are a clear indication that the wind instruments rear end (part of the main bore behind the first open tone hole) creates formants by functioning as a passive filter. Previous studies implemented impulse shaping for woodwind sound

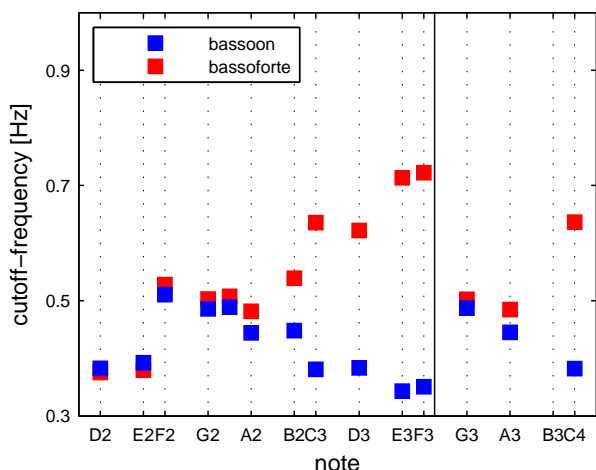


Figure 10: Cutoff frequencies of the air column of bassoon and *Bassoforte*.

synthesis and related the formants to the source spectrum [9]. If our hypothesis can be proven, it provides a link between the design of the resonator and sound color attributes. This enables a model-based design of wind instrument air columns with a timbre-related target function. One-dimensional transmission-line models, like the one used in this study, have recently been employed successfully to design a new tone-hole concepts for a clarinet [13], where the cutoff frequency of the normal clarinet has been preserved.

5. CONCLUSIONS

Today's wind instruments have undergone a long evolution process. With most measurement techniques, we can only expect small differences when comparing a hand crafted masterpiece of a musical instrument and an industrially manufactured counterpart model for beginners. Naturally, understanding relations between the make of the instrument and its performance characteristics is easier, if clear effects can be observed. However, because of their high degree of maturity, investigations on real musical instruments do not easily allow for modifications which are fundamental and reasonable at the same time. For minor modifications, effects may be difficult to detect with measurement techniques, even if the consequences appear to be major ones to a sensitive musician.

Working on functional models on the other hand, whether numerically or experimentally, it is possible to implement any modification. However, the model may be quite far from the actual musical application and the transfer of results to the real world is crucial but far from simple.

The interdisciplinary design approach from which the *Bassoforte* emerged seems to be an intermediate way with links to practice and theory. This was and is helpful to both instrument makers and scientists who are willing to learn from each other.

The reported measurement differences in the sound of bassoon and *Bassoforte* are clearly larger than the variance

introduced by the experimental and postprocessing methods. The study experimentally confirms the importance of the cut-off frequency for the sound color of wind instruments and points out an interesting relation to the formant frequencies in the sound. Both cut-off frequency and formant frequencies are straightforward to obtain by standard measurements. Formant center frequencies are much more robust with respect to the recording situation than the harmonic spectral centroid is.

We believe that our results will be helpful to define measurable sound design targets. We hope to understand better the link between the geometry and the sound characteristics by monitoring the acoustical changes during further modifications on the *Bassoforte* and possibly other wind instruments.

Acknowledgments

A major part of this work has been carried out at Technische Universität Dresden, under supervision of Prof. Roger Grundmann. We are indebted to Johannes Baumgart and Prof. Malte Kob for fruitful discussions and providing help in the measurements. We further acknowledge the bassoonists Stefan Pantzier (studio) and Hannes Fritsch (reverberation chamber), as well as Maximilian Pauls who carried out the studio recordings at Hochschule für Musik Dresden.

The construction of the *Bassoforte* was supervised by the instrument makers Guntram and Peter Wolf at Guntram Wolf Holzblasinstrumente GmbH, Kronach, and Benedikt Eppelsheim Blasinstrumente Munich, who also largely contributed to the tone-hole design of the present instrument. The work was supported by the German Federal Ministry of Economics and Technology (BMWi), in the projects ZIM KF 2229603 and KF2254301.

This work is dedicated to Guntram Wolf who with his cordial and open-minded character has greatly inspired the research cooperation in the past years.

6. REFERENCES

- [1] D. Keefe, "Experiments on the single woodwind tone hole," *Journal of the Acoustical Society of America*, vol. 72, no. 3, pp. 688–699, 1982.
- [2] A. H. Benade, "Equivalent circuits for conical waveguides," *The Journal of the Acoustical Society of America*, vol. 83, no. 5, pp. 1764–1769, 1988.
- [3] C. Nederveen, *Acoustical Aspects of Woodwind Instruments*. Northern Illinois University Press, 1998.
- [4] J. Kergomard, *Elementary Considerations on Reed-Instrument Oscillations*. Springer Verlag Wien - New York, 1995, ch. 6, pp. 230–290.
- [5] T. Grothe, "Investigation of bassoon embouchures with an artificial mouth," in *Program of the 11th Congrès Français d'Acoustique and 2012 IOA annual meeting, Acoustics2012, Nantes, 2012*.

- [6] D. H. Keefe, "Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximations for isothermal and nonisothermal boundary conditions," *The Journal of the Acoustical Society of America*, vol. 75, no. 1, pp. 58–62, 1984.
- [7] J.-P. Dalmont, C. J. Nederveen, V. Dubos, S. Ollivier, V. Meserette, and E. te Sligte, "Experimental determination of the equivalent circuit of an open side hole: Linear and non linear behaviour," *Acta Acustica united with Acustica*, vol. 88, pp. 567–575(9), July/August 2002.
- [8] R. H. Cronin and D. H. Keefe, "Understanding the operation of auxiliary fingerings on conical double-reed instruments." in *Program of The 131st Meeting of the Acoustical Society of America*, vol. 99, no. 4. ASA, 1996, pp. 2456–2457.
- [9] W. Voigt, *Untersuchungen zur Formantbildung in Klängen von Fagott und Dulzianen*. Gustav Bosse Verlag, 1975.
- [10] J. Meyer, *Acoustics and the Performance of Music – Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instruments Makers*. Springer, 2009.
- [11] T. Grothe, J. Baumgart, and R. Grundmann, "Klangfarbenuntersuchung an historischen Fagotten – Vergleich von Dulzian, Barockfagott, klassischem und modernem Fagott," *Rohrblatt - Magazin für Oboe, Klarinette, Fagott und Saxophon* 04/2006, 2006.
- [12] A. Benade, *Fundamentals of Musical Acoustics*. Dover Publications, Inc., New York, 1990.
- [13] D. Noreland, J. Kergomard, F. Laloe, C. Vergez, P. Guillemain, and A. Guilloteau, "The logical clarinet: numerical optimization of the geometry of woodwind instruments." *published online*, vol. arXiv:1209.3637v2 [physics.pop-ph], 2013.

NUMERICAL MODELING OF A RECORDER IN THREE DIMENSIONS

N. Giordano

Department of Physics
Purdue University
West Lafayette, IN 47907 USA
giordano@purdue.edu

ABSTRACT

A modeling study of the recorder using a direct numerical solution of the Navier-Stokes equations is described. Results for the dynamics of the air jet and the spectrum as a function of the blowing speed u are presented. The motion of the air jet is in good agreement with the general behavior found in experimental studies. An increase in the mode frequencies and switching between modes are observed as u is increased, in qualitative agreement with the behavior found for recorders and flue organ pipes.

1. INTRODUCTION

The modeling of musical instruments using the fundamental equations of mechanics has attracted much interest in recent years. It is now possible to calculate the sound produced by several instruments using Newton's laws to describe the motion of all parts of the instrument (i.e., strings, hammers, bows, and plates) along with the equations of linear acoustics to describe the sound produced by vibrating surfaces. This type of “first-principles” modeling, often referred to a physical modeling, has yielded many interesting and realistic results for string instruments such as pianos and guitars (see, e.g., [1,2] for representative examples of this work and many additional references), and for several percussion instruments (see, e.g., [3,4]). These studies have begun to yield new insights into tone production and have the potential to aid in designing modifications (and perhaps even improvements) to instruments to achieve specific tonal objectives.

Recent years have seen an increased interest in the application of first-principles modeling to wind instruments. Unlike string and percussion instruments, wind instruments require application of the Navier-Stokes equations, and dealing with those equations requires much greater computational speed than do the equations of linear acoustics. Even so, Navier-Stokes modeling of wind instruments is now becoming feasible with high performance computers and this paper describes a step in that direction.

We hasten to add that wind instruments have certainly been the object of many modeling studies for many years.

Copyright: © 2013 N. Giordano. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Much has been learned using linear acoustics to describe wave propagation inside the instrument tube and some sort of lumped model (often with parameters taken from experiment) to describe the nonlinear processes that take place in and near the mouthpiece. One objective of Navier-Stokes-based modeling is replace these lumped models with a first-principles description of such nonlinear processes. For example, Navier-Stokes modeling of a recorder should make it possible to predict how the tonal properties are affected quantitatively by changes in the geometry of the channel and labium, such as the addition of chamfers (a problem on which we have recently reported [5]).

Previous workers have reported Navier-Stokes-based modeling of the recorder, with several different numerical techniques. Some work has involved what is generally termed “direct numerical solution” of the Navier-Stokes equations [6-11]. That work has yielded interesting results for the spectrum and the velocity field. We should particularly note the recent work of Miyamoto and coworkers [11] in which the Navier-Stokes equations were solved using an algorithm similar to the one we have employed (see below). They report results that compare well with those reported below, although their studies in three dimensions were limited by restrictions set by computational speed. Other workers [6,12-15] have employed an alternative numerical approach, the lattice-Boltzmann method. The lattice-Boltzmann method has yielded results basically similar to those found with direct numerical solutions albeit with the restriction of requiring an unphysically large viscosity to maintain numerical stability. It seems fair to say that the previous work has been extremely interesting and has opened the way to quantitative studies of wind instruments, but that much remains to be done.

This paper describes such a study of the recorder in three dimensions using a direct numerical solution of the compressible Navier-Stokes equations. We present spatiotemporal maps showing the dynamics of the air jet near the labium and study how the spectrum depends in detail on the blowing speed, and compare with experimental results for recorders and flue organ pipes.

2. MODELING APPROACH

Many numerical methods have been developed for dealing with the compressible Navier-Stokes equations. We have chosen to use the MacCormack method [16,17], an explicit finite difference-time domain method. This is a

relatively simple method that is well tested for the range of Reynolds numbers found in the recorder. We have written a custom algorithm that combines the MacCormack method with artificial viscosity to suppress short wavelength instabilities that are not relevant at acoustic frequencies [18,19]. We model the surfaces of the recorder and the surfaces of a box surrounding the recorder as an acoustic impedance as described by Botteldooren [20], and with the parallel component of the velocity zero at all surfaces. This acoustic impedance relates the pressure and the normal component of the velocity at the surface, and results in both reflection and a small absorption of energy at the surface, as found at real surfaces. Further details of our algorithm are given elsewhere [21] where we reported a study of the recorder in two dimensions. In this paper we describe the extension of our work to a fully three dimensional instrument.

Figure 1 shows the model geometry. This “instrument” is smaller than a normal recorder, and corresponds approximately to a soprano recorder. However, the height of the channel and the distance from the labium to the exit of the channel were the same as found in normal (full size) recorders. The channel was straight without chamfers and the labium was aligned with the center of the channel. The recorder tube had a square $d \times d$ cross-section with $d = 10.5$ mm. The instrument was enclosed in a region of dimensions $0.2 \times 0.06 \times 0.06$ m³ with surfaces that reflected and absorbed acoustic energy as described by an acoustic impedance as noted above. A non-uniform Cartesian numerical grid was employed, with a grid spacing of typically 0.1 mm near the channel and labium, and in the tube, and greater outside. (Some calculations were carried out with smaller and larger grid spacings to check that this spacing was sufficiently small.) This grid design made the calculation much faster than with a uniform grid, but even with the nonuniform grid the total number of grid points was typically 3×10^7 . An air velocity parallel to the channel was imposed on the left side of the channel in Figure 1, with the spatial dependence in the perpendicular directions set to that expected for Poiseuille flow. The magnitude of this velocity, which we will refer to as the blowing speed u , was increased linearly from zero at $t = 0$ to a final magnitude at $t = 5$ ms.

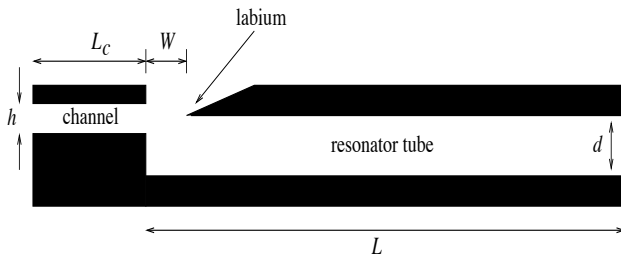


Figure 1. Recorder geometry showing the plane that cuts through the center of the instrument. The resonator tube had a square cross-section with $d = 10.5$ mm. The other dimensions were $L_c = 8.0$ mm, $W = 4.0$ mm, $h = 1.0$ mm, and $L = 105$ mm. The instrument was enclosed in a region with dimensions $0.2 \times 0.06 \times 0.06$ m³.

The algorithm was implemented on a parallel computer and achieved an increase in computational speed greater

than $0.9 \times N$ due parallelization, where N was the number of CPUs employed in the calculation. Typical values of N were 16–48. In terms of real time, 1 ms of sound could be calculated in about 2 hours of real time using a computer with a rated speed of 0.3 Tflop.

3. RESULTS

Figure 2 shows images of the air jet as it exits the channel and strikes the labium. These images were recorded at intervals of 1 ms, and the images in Figure 2(a)-(h) show approximately one complete period of the oscillation. The air jet is seen to oscillate between positions above and below the labium, as found in experimental studies [22]. The behavior seen here differs qualitatively from the behavior found in two dimensional simulations where the oscillations of the air jet never reached below the labium [21]. This difference was not unexpected due to the effect of dimensionality on the dissipation of vortices and eddies (see the discussion, e.g., in [11]). There is a wealth of information in the results that produced these images, and we will report a quantitative analysis elsewhere. In the rest of this paper we will consider the behavior of the sound spectrum as a function of blowing speed.

Figure 3 shows typical results for the acoustic pressure p as a function of time at a location outside the recorder tube. As noted above, the flow speed at the center of the channel was zero at $t = 0$ and increased linearly to a final value at $t = 5$ ms (here that value was $u = 30$ m/s). Any shock waves or similar kinds of discontinuities in p generated as the jet was initiated were very small on the scale of Figure 3. Between 5 ms and about 15 ms, the tone is complex, containing substantial weight at several frequencies. The details of the spectrum during this attack portion of the tone are discussed in [5]. In the remainder of this paper we concentrate on the behavior when steady state is reached; this occurs at about 20 ms, after which the tone has a much simpler character. Although it is not shown here, the behavior seen for $t > 20$ ms in Figure 3 seems to be maintained for as long as the air jet is maintained (for the longest simulations we have studied, which extended to $t = 150$ ms). The sound pressure level observed in Figure 3 at steady state is about 110 dB, a typical result found in real recorder playing. For this value of u the resulting tone had a fundamental frequency of $f_1 \approx 1380$ Hz, a value consistent with the length of the recorder and the expected end corrections.

The spectrum of the pressure signal in Figure 3 is shown in Figure 4. This spectrum was calculated after the pressure signal reached steady state, and one sees that the power at the fundamental frequency is greater than that at the second harmonic by more than an order of magnitude. Several spectral components are also visible at higher frequencies.

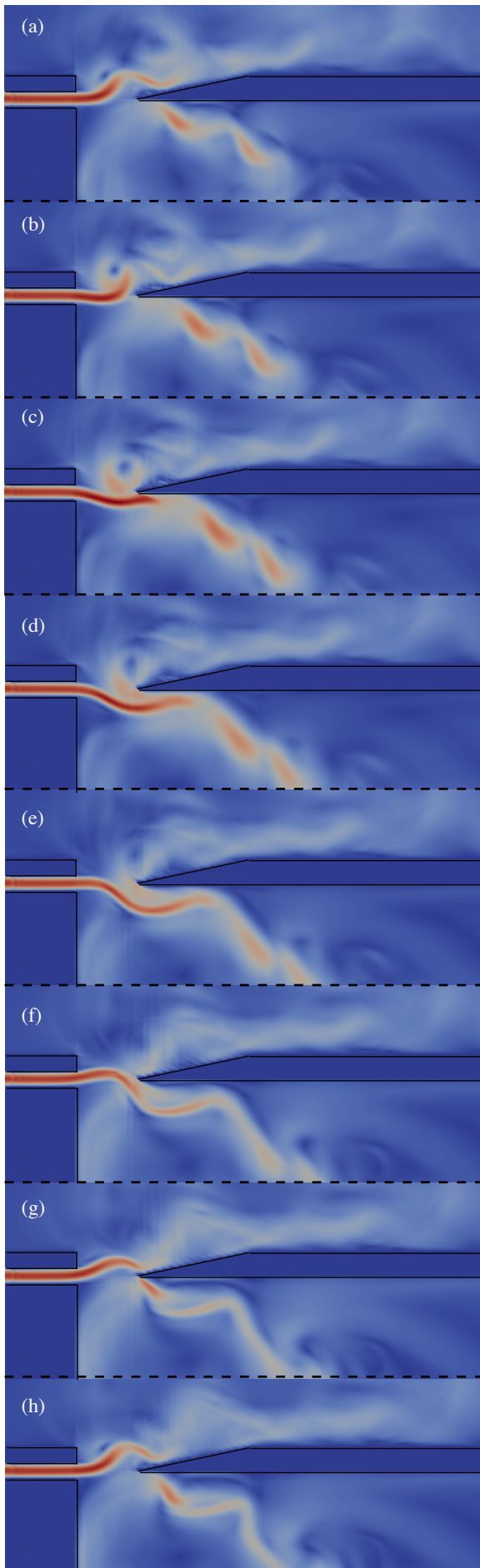


Figure 2. Images of the air speed during one period in the oscillation of the air jet. These images show only the region near the exit of the channel and the labium, with the surfaces of the recorder outlined in black. Dark red regions are ones of high speed and dark blue indicates region of low speed. The flow speed at the center of the channel was $u = 25$ m/s.

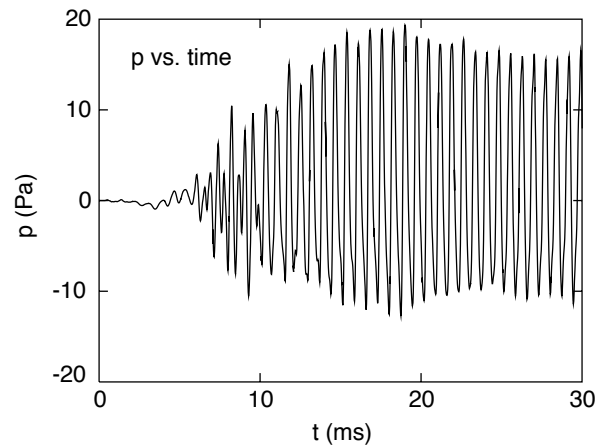


Figure 3. Variation of the acoustic pressure outside the recorder as a function of time. The flow speed in the center of the channel was $u = 30$ m/s.

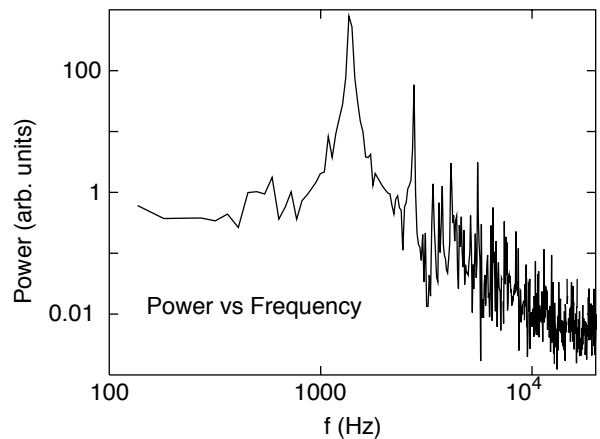


Figure 4. Spectrum of the acoustic pressure with $u = 30$ m/s at the center of the channel.

The frequencies of the first few harmonics and their spectral power as a function of blowing speed u are shown in Figures 5 and 6. A minimum blowing speed of about $u = 15$ m/s was required to maintain an oscillation. At larger values of u the component at the fundamental frequency ($f_1 \approx 1350$ Hz, a value that depends somewhat on the value of u ; see below) was strongest up to about $u = 35$ m/s. Above that speed the component at the second harmonic dominated, and was largest up to the highest blowing speed studied ($u = 45$ m/s). Note that the strength of the component at the third harmonic became very weak in the regime at which the second harmonic dominated, as expected.

The behavior found in Figures 5 and 6 is qualitatively similar to that reported experimentally and explained by models of sound production in flue pipes (see, e.g., [23,24]). For the future we plan to carry out simulations

for a systematic range of parameters to allow a quantitative comparison with those results.

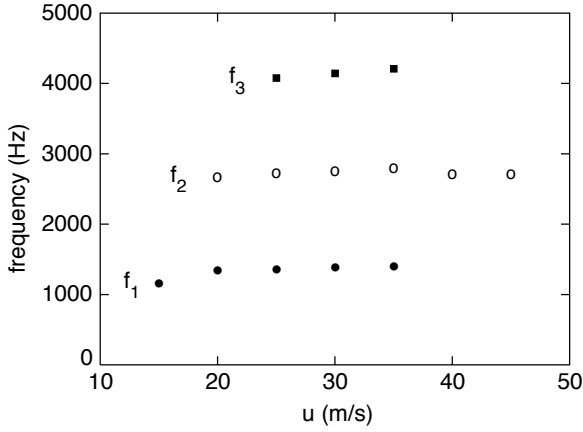


Figure 5. Frequencies of the spectral components at the fundamental (f_1), second harmonic (f_2), and third harmonic (f_3) as a function of the blowing speed u at the center of the channel.

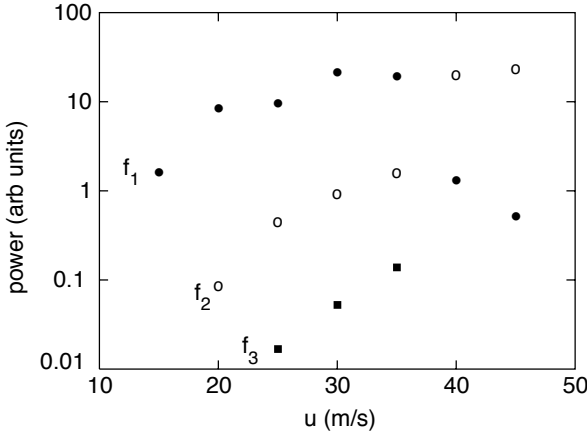


Figure 6. Power of the spectral components at the fundamental, second harmonic, and third harmonic as a function of the blowing speed u at the center of the channel.

The value of the fundamental frequency f_1 in Figure 5 varies slightly with u , an effect well known to recorder players. Data for that variation has been given by Bak [25] who showed that for the range of u in which a stable oscillation and resonance in the tube is obtained, the frequency of that oscillation varies approximately as

$$f_1 \approx ku^\alpha \quad (1)$$

Bak also found that $\alpha \approx 0.05$ for several notes played by several different recorders. Our results for f_1 in the regime in which the fundamental frequency is dominant are shown in Figure 7. In this log-log plot, Equation 1 is a straight line, which is seen to be consistent with our results. We find $\alpha \approx 0.085$ and while there is a considerable uncertainty of perhaps 0.01 or more, the value we find is somewhat larger than found by Bak. The reason for this difference is not clear, but may be due to fact that our model recorder is much shorter than a standard recorder.

We speculate that this may make the end correction to f_1 more sensitive to u , an issue that needs to be studied further. We plan to investigate this issue by extending our modeling to full size recorders, which will involve much longer simulation times.

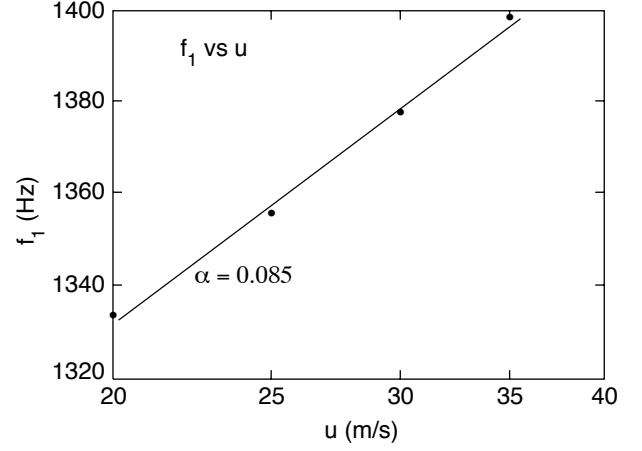


Figure 7. Log-log plot of the fundamental frequency f_1 as a function of blowing speed u at the center of the channel. The straight line corresponds to Equation 1 with $\alpha = 0.085$.

4. SUMMARY AND CONCLUSIONS

We have reported results from a quantitative modeling study of the recorder in three dimensions. Results for the dependence of the spectrum on blowing speed are in reasonable agreement with available experimental results. It should be possible to compare our results quantitatively to models of sound generation in recorder-like instruments, such as the jet-drive model [26], a problem we will address in the future. An extension of our modeling approach to other wind instruments also seems quite feasible.

Acknowledgments

The author thanks the Rosen Center for Advanced Computing at Purdue University for access to the computational resources essential for this work.

5. REFERENCES

- [1] N. Giordano and M. Jiang, "Physical modeling of the piano," *EURASIP J. Appl. Signal Process.* 2004(7), 926–933 (2004).
- [2] E. Bécache, A. Chaigne, G. Derveaux, and P. Joly, "Numerical simulation of a guitar," *Comput. Struct.* **83**, 107–126 (2005).
- [3] V. Doutaut, D. Matignon, and A. Chaigne, "Numerical simulations of xylophones. II. Time-domain modeling of the resonator and of the radiated sound pressure," *J. Acoust. Soc. Amer.* **104**, 1633–1647 (1998).

- [4] L. Rhaouti, A. Chaigne, and P. Joly, "Time-domain modeling and numerical simulation of a kettledrum," *J. Acoust. Soc. Amer.* **105**, 3545-3562 (1999).
- [5] N. Giordano, "Direct numerical simulations of the recorder in two and three dimensions," *Proceedings of Meetings on Acoustics* **19**, 035062 (2013).
- [6] P. A. Skordos, "Modeling of flue pipes: Subsonic flow, lattice Boltzmann and parallel distributed distributed computers," Ph.D. thesis, MIT, Cambridge, MA, 1995.
- [7] P. A. Skordos and G. J. Sussman, "Comparison between subsonic flow simulation and physical measurements of flue pipes," in *Proceedings of the International Symposium on Musical Acoustics*, Dourdan, France (1995), pp. 1-6.
- [8] Y. Obikane and K. Kuwahara, "Direct simulation for acoustic near fields using the compressible Navier-Stokes equations," *Comput. Fluid Dyn. J.* 2008/2009, 85-91 (2009).
- [9] Y. Obikane, "Computational aeroacoustics on a small flute using a direct simulation," in *Computational Fluid Dynamics*, edited by A. Kuzmin (Springer-Verlag, New York, 2010), pp. 435-441.
- [10] M. Miyamoto, Y. Ito, K. Takahashi, T. Takami, T. Kobayashi, A. Nishida, and M. Aoyagi, "Applicability of compressible LES to reproduction of sound vibration of an air-reed instrument," in *Proceedings of the International Symposium on Musical Acoustics*, Sydney and Katoomba, Australia (2010).
- [11] M. Miyamoto, Y. Ito, T. Iwasaki, T. Akamura, K. Takahashi, T. Takami, T. Kobayashi, A. Nishida, and M. Aoyagi, "Numerical study on acoustic oscillations of 2D and 3D flue organ pipe like instruments with compressible LES," *Acustica-Acta Acustica* **99**, 154-171 (2013).
- [12] H. Kühnelt, "Simulating the mechanism of sound generation in flutes using the lattice Boltzmann method," in *Proceedings of the Stockholm Music Acoustics Conference (SMAC 03)* (2003), SMAC1-SMAC4.
- [13] H. Kühnelt, "Simulating the mechanism of sound generation in flutes and flue pipes with the lattice-Boltzmann-method," in *Proceedings of the International Symposium on Musical Acoustics*, Nara, Japan (2004), pp. 251-254.
- [14] A. R. da Silva and G. Scavone, "Coupling lattice Boltzmann models to digital waveguides for wind instrument simulations," in *Proceedings of the International Symposium on Musical Acoustics*, Barcelona, Spain (2007), pp. 1-7.
- [15] A. R. da Silva, H. Kühnelt, and G. Scavone, "A brief survey of lattice-Boltzmann method," in *Proceedings of the International the Acoustics*, Madrid, Spain (2007), pp. 1-6.
- [16] R. W. MacCormack, "The effect of viscosity in hypervelocity impact cratering," *AIAA Pap.* **69**-354, 1-7 (1969).
- [17] R. W. MacCormack, "A numerical method for solving the equations of compressible viscous flow," *AIAA J.* **20**, 1275-1281 (1982).
- [18] A. Jameson, W. Schmidt, and E. Turkel, "Numerical solution of the Euler equations by finite volume methods using Runge Kutta time stepping schemes," *AIAA Pap.* **81**-1289, 1-14 (1981).
- [19] Swanson, R. C. and Turkel, E. (1987). "Artificial dissipation and central difference schemes for the Euler and Navier-Stokes equations," *AIAA paper* **87**-1107, 55-69.
- [20] D. Botteldooren, "Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid," *J. Acoust. Soc. Am.* **95**, 2313-2319 (1994).
- [21] N. Giordano, "Direct numerical simulation of a recorder," *J. Acoust. Soc. Amer.* **133**, 1113-1118 (2013).
- [22] B. Fabre, A. Hirschberg, and A. P. J. Wijnands, "Vortex shedding in steady oscillation of a flue organ pipe," *Acustica-Acta Acustica* **82**, 863-877 (1996).
- [23] M-P. Verge, B. Fabre, A. Hirschberg, and A. P. J. Wijnands, "Sound production in recorderlike instruments. I. Dimensionless amplitude of the internal acoustic field," *J. Acoust. Soc. Am.* **101**, 2914-2924 (1997).
- [24] R. Auvray, B. Fabre, and P.-Y. Lagrée, "Regime change and oscillation thresholds in recorder-like instruments," *J. Acoust. Soc. Am.* **131**, 1574-1585 (2012).
- [25] N. Bak, "Pitch, temperature and blowing pressure in recorder playing. Study of treble recorders," *Acustica* **22**, 296-299 (1969).
- [26] S. Dequand, J. F. H. Willerns, M. Leroux, R. Vullings, M. van Weert, C. Thieulot, and A. Hirschberg, "Simplified models of flue instruments: Influence of mouth geometry on the sound source," *J. Acoust. Soc. Am.* **113**, 1724-1735 (2003).

Investigation of bassoon directivity

Timo Grothe

Hochschule für Musik Detmold,
Erich-Thienhaus-Institut
grothe@hfm-detmold.de

Malte Kob

Hochschule für Musik Detmold,
Erich-Thienhaus-Institut
kob@hfm-detmold.de

ABSTRACT

Due to the distribution of tone holes across a long bent corpus, the bassoon has an irregular directivity. This is well-known for Tonmeister, who wish to record the sound of a bassoon with a small number of spot microphones. This study presents the directivity patterns of the bassoon obtained in an anechoic chamber. Using an artificial mouth with adjustable embouchure, the instrument was excited at different pressure levels that were held constant for a time span of minutes. Several fingerings, covering most of the playing range of the bassoon ($f_0 = 58 \dots 591$ Hz), have been sounded while the bassoon was mounted vertically and horizontally with respect to the rotary axis of a turntable. The use of the artificial mouth allowed studying influences of the fingering as well as of the playing dynamics on the pressure radiation pattern. Additionally an acoustic camera was used to support the interpretation the observed sound radiation patterns.

1. INTRODUCTION

Directivity measurements of musical instruments have become a valuable tool in two domains. One objective of the assessment of acoustic radiation properties of traditional or new instruments is to provide a base of knowledge for musicians and sound engineers and for optimum performance practice and recording. Apart from the directivity measurements by J. Meyer [1] more recent projects have created 3D databases with directivity data from most symphonic instruments [2]. A possible application is the use of directivity patterns for the design of virtual acoustic environments. Since room acoustical simulation programs have become quite powerful the accuracy of simulations now also depends on the availability of radiation data for the implementation of natural sounding virtual instruments.

A main problem of directivity measurements of acoustic music instruments is the need to repetitively excite the instrument with constant excitation parameters during the course of the measurement.

For the bassoon this is of major importance since the availability of an acoustic reference signal at a specific location outside of the instrument is not easy to provide. In singers the near field sound pressure at the mouth can serve as a

reference signal. If divided by this reference the sound pressure in the far field would always be normalized to the actual sound pressure generated at the only significant acoustic output of the singer [3]. Using this technique the individual radiation characteristics can be evaluated independent of the actual sound production technique and sound pressure level. The bassoon does not offer this unique acoustic reference point since the radiation depends critically upon the opened and closed tone holes for a specific fingering. Therefore an artificial mouth [4] was used to ensure a constant driving pressure during a complete measurement cycle. Similar to the approach introduced for singers, the unsteady pressure inside the double-reed is monitored and used as a reference. This allows to smooth variations in the radiated sound spectra due to minor variations in the artificial excitation.

Aim of this study is to validate and supplement the available directivity data. By excluding the absorbing effect of a musician, the radiation characteristics of the instruments body can be studied in detail. Another objective was to verify if a commercial microphone array would help to understand the complex radiation characteristics of the bassoon more in detail.

2. MATERIALS AND METHODS

To measure the directional radiation characteristics of musical instruments, it is important to excite the instrument constantly, while turning it across his main axes.

If a high precision is required, these measurements are difficult to obtain with a musician playing. We therefore decided to mount the bassoon on an automated turntable and used a blowing machine to provide a most constant excitation.

2.1 Experimental Setup

The artificial mouth is equipped with a rigidly mounted double reed, one or two artificial lips can be pressed to the reed, imitating the upper and lower jaw. The lip force can be adjusted within a range of 0.1 to 10 N for each lip. Pressurized dry air from a medical compressor is fed into the system; the blowing pressure in the mouth is regulated manually with a precision of ± 2 mbar. Once adjusted, a large diaphragm pressure regulator provides a stable mouth pressure within a tolerance of ± 0.5 mbar for a time span of several minutes. For reasons of repeatability, a synthetic bassoon double-reed is used.

While the turntable was running from 0 to 360 °, the radiated sound was measured by two measurement microphones (B&K Type 4190, Type 4938), at a distance from the rotary axis of 3.5 and 1.2 m, respectively. The two microphone distances were chosen for practical reasons: the more distant one was placed as far as possible from the bassoon, the other one as close as possible to the moving set-up. Due to the small openings of the tone holes for frequencies above ca. 100 Hz ka is much smaller than unity, and therefore directive sound radiation can be expected that should be similarly recorded by both microphones. However, the SNR is much better at high frequencies in the closer microphone which facilitates the evaluation of the signals.

The acoustic camera¹ used in this study comprises a microphone array with 40 MEMS - microphones irregularly distributed over an area of 0.6×0.6 m, with mic to mic distances between 5 and 10 cm. It was placed at a distance of 2.5 m from the bassoon. The principle of the acoustic camera is the evaluation the time differences in the microphone array when a sound wave from the far field propagates through it. The evaluation allows for a localization of sound sources and results in a color map that overlays a monochrome image of the observed sound radiating objects. The color indicates the radiated sound pressure level of the objects. Several evaluation algorithms exist that are optimized for individual analysis tasks.

The pressure inside the reed mouthpiece was sensed with a miniature pressure transducer (*kulite* XCG093) The operation noise of the turntable has been recorded separately, and was subsequently used as a sound pressure offset and subtracted from the measured sound spectra of the bassoon tones. The recordings were carried out in an anechoic chamber² with dimensions of 4.8×5.8×4 m³. All notes from the major scale in B♭ have been sounded over three octaves (B♭1, $f_0 = 58.7$ Hz to D5, $f_0 = 591$ Hz) on a modern German bassoon. The fingerings have been realized using adhesive tape to operate the keywork; fingerholes have been closed by a mouldable putty like adhesive (Bostik BluTack).

The lip force has been adjusted to excite each note in tune at a moderate level. The blowing pressures were in the range typical for bassoon playing [5], namely $1.5 < p_m < 6$ kPa, depending on the fingering. The intonation was referenced to the well-tempered scale (A4, 443 Hz); the pitch was controlled with a digital tuner, to be stable within ± 5 Cent. By monitoring lip force and blowing pressure, the bassoon could be excited with nominally identical embouchure in several runs. A fixture has been constructed which allows to mount a regular bassoon with artificial mouth at any angle onto the turntable. The setup with the bassoon fixed vertically is shown in Fig. 1).

2.2 Data Analysis

The sound pressure recorded in the experiments as recorded by a Brüel and Kjaer Pulse system ($f_s = 65.536$ kHz, 16 bit), were analyzed by means of standard algorithms

¹ "Noise Inspector", CAE Software and Systems, Gütersloh, Germany

² University of Applied Sciences OWL, Lemgo, Germany



Figure 1: Bassoon with artificial mouth excitation mounted vertically on a turntable. Left and right of the bassoon are the acoustic camera and the close microphone (1.2 m). The far microphone is not shown.

implemented in Matlab, to obtain a power spectral density estimate of the sound. The time-series of a full turn was approximately 180 s long and was analyzed with a running filter of 50% overlap and a window size of 1 second, which corresponds to a rotation angle of 2°. Analysis frames have been passed through a third octave filterbank with center-frequencies f_c referenced to 1 kHz and the respective equivalent pressures $P_{eq}(f_c)$ are calculated. Further, the DFT of each analysis frame has been calculated, and subsequent to an auto-correlation based f_0 estimation, the spectral pressures have been added in frequency bands with a width of f_0 around the i^{th} harmonic peak in the frequency domain to give $P_i(i f_0)$. In the following this will be referred to as harmonic analysis, marked by the subscript $(\cdot)_h$, while the third octave-filtered results are marked by $(\cdot)_{3rd}$.

As the excitation spectrum $P_r(\alpha)$ was not perfectly constant during a full turn of the turntable ($\alpha = 0 \dots 2\pi$), for each frequency or frequency band the magnitude of the acoustic sound spectra $P_{ac}(\alpha)$ as measured outside of the instrument has been centered using the spectral mean value $|\bar{P}_r| = \frac{1}{2\pi} \int |P_r| d\alpha$ of the pressure measured inside the reed, to obtain $|P_{ac}|^* = |P_{ac}| |P_r| / |\bar{P}_r|$. The variation of $|P_r| / |\bar{P}_r|$ during one full turn is within 0.7, 3.7 and 12 dB standard deviation for the frequency ranges 1 to 500 Hz, 0.5 to 5 kHz and 5 to 20 kHz.

Both third octave filtered and harmonic spectral results have subsequently been smoothed with a running average filter with a window width of 10°, and will be displayed as functions $L_{3rd}(\theta)$, $L_{3rd}(\phi)$ and $L_h(\theta)$, $L_h(\phi)$ in dB, where θ is the azimuth angle and ϕ the polar angle with respect to the main longitudinal axis of the instrument (from "butt" to "bell").

3. RESULTS AND DISCUSSION

Regarding the octave filtered directivity patterns, our measurements are in general agreement with previous results of Meyer [1]. Covering a dynamic range of 15 to 30 dB, our

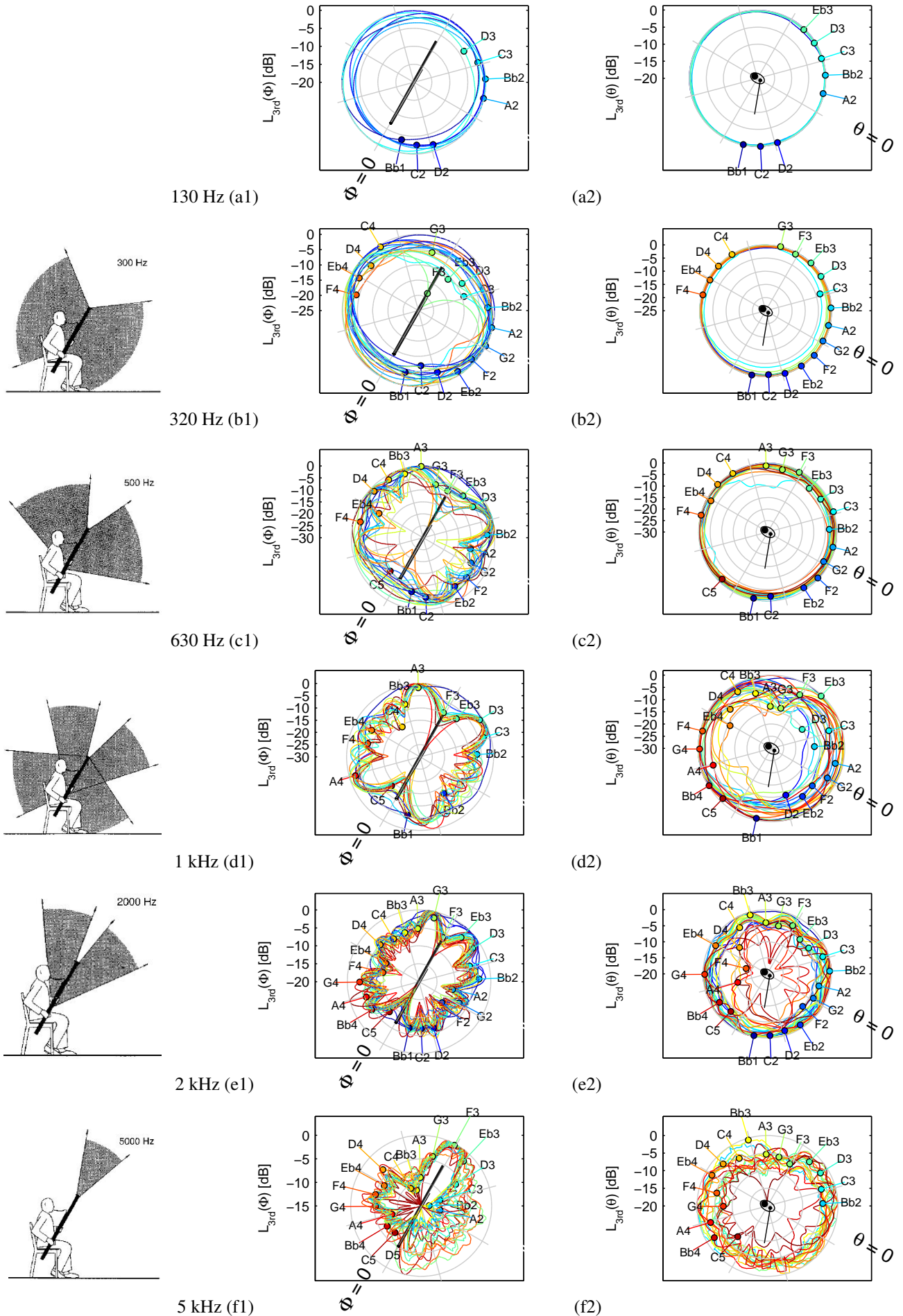


Figure 2: Third-octave filtered radiation patterns of the bassoon. Polar rotation (middle) and azimuthal rotation (right) are compared to Meyers results [1] (left) with a dynamic range of 3 dB. Notes of a B major scale have been played in *mezzoforte* across the whole playing range (Bb1 ($f_0 = 58$ Hz) to D5 ($f_0 = 591$ Hz)), at moderate constant blowing pressures. The experiment was carried out in an anechoic chamber using an artificial mouth.

results confirm the radiation patterns given there. In Fig. 2 – for better comparison with the literature results (left column) – the polar plots in the median plane (middle column) have been rotated by an offset of $\Phi_0 = 240^\circ$. Flipping this view over the $\Phi = 90^\circ$ - axis results in the topview of the horizontal plane (right column). In this graph, the oval in the center symbolizes the bassoon cross section, the line originating from the narrower bore symbolizes the bocal and depicts the position of the player (who was absent in this study).

3.1 The median plane

The radiation in the median plane of the instrument (polar rotation) is measured with the bassoon mounted horizontally on the turntable. The radiation pattern are circular for frequencies below $f_c = 130$ Hz. As the frequency increases, the sound pressure radiated from the open end ("bell") decreases. The circular pattern is constricted on the longitudinal axis of the instrument, and two symmetrical side lobes show up. The angle between these decreases with frequency, and they finally merge into one on-axis lobe around 4 kHz. For frequencies up to 1 kHz, the radiation pattern is symmetrical with respect to the instrument's transverse axis (orthogonal to the main axis). This means that the radiated sound pressure is equal both at the open ("bell") and closed ("butt") end. This discrepancy with Meyers results can partially be explained by the sound absorbing body of the player, who was present in his measurement, but not in ours. However we would not expect his effect to be large at relatively low frequencies around 600 Hz (Fig. 2 (c1)).

At higher frequencies a focusing of the pattern onto the outward pointing direction is observed. This focusing might be significantly enhanced due to the player's body in a real playing situation. However our results indicate, that still a significant amount of sound pressure (-5 dB) is still radiated from the tone-holes in the middle of the instrument. Note, that Figs. 2 (a1)-(e1) show measurements in at 3.5 m distance, while Fig. 2 (f1) shows the radiation pattern measured at 1.2 m distance.

3.2 The horizontal plane

The radiation in the horizontal plane of the instrument (azimuthal rotation) is measured with the bassoon mounted vertically on the turntable, as shown in Fig. 1. In contrast to the symmetrical radiation discussed above, the sound radiation patterns in the horizontal plane are nearly circular for frequencies below 1 kHz. Above this, the radiation patterns become very irregular, strongly depending on the fingering in the horizontal plane of the instrument (azimuthal rotation) for higher frequencies. This can be well understood taking into account the sharp bent at the instrument's lower end ("butt"). For any note higher than F2 ($f_0 = 87.8$ Hz), tone holes on both downstream and upstream side of the bent are open. Thus, possible sound sources are found all across the outer dimension of the bassoon, and are pointing at different azimuthal angles from the bores' center axis. Essentially, the strong lobes observed in Fig. 2 (e1) are the result of superpositions and

may rather not be interpreted as a directed sound impact from a single tone hole.

Above 2 kHz, the dynamic variations with respect to the turn angle are within 5 dB for most notes, and no clear directional characteristics can be observed.

Due to the relation of tone hole dimensions to the bassoons outer perimeter, we expect the influence of the radial directions of the tone hole chimneys on the radiation pattern in the median plane to be negligible in the frequency range of interest. Therefore we assume the radiation pattern to be the same in any median plane, and did not study a third plane orthogonal to the two above. In presence of a player this assumption should be reconsidered.

3.3 Harmonic analysis

The complex radiation pattern of the bassoon in the horizontal plane are even more prominent in the harmonic analysis. In Fig. 3 (a) and (b) the amplitude variations of single partials during the turn of the instrument are mapped. Whereas in the median plane (b) the radiation patterns of neighbored partials resemble each other, in the horizontal plane (a) significant deviations up to 20 dB are found. As the human ear is very sensitive in this frequency range, these differences are likely to influence the perceived timbre.

3.4 Influences of the playing dynamics

When blowing reed instruments, the mouth pressure plays a major role in determining the dynamic level of the radiated sound. As the mouth pressure is increased, the higher harmonics in the sound spectrum are more affected than the lower ones. In Fig. 4, azimuthal directivity patterns are shown for two harmonics (6th and 8th) near 1 kHz of the note Eb3. This note has been chosen because it has the specialty of being played with a forked fingering. It is interesting that although being close in frequency, the two distinct partials exhibit a quite different radiation pattern. For each of the five experiments, the mouth pressure was adjusted to a constant value within $1.6 < p_m < 5.8$ kPa, to cover the dynamical range of this note from *pianissimo* to *fortissimo* in five steps (*pp*, *p*, *mf*, *f*, *ff*) that appeared to be equal steps in loudness as perceived by the experimenter. It can be seen that this variation in dynamic level does not effect the general shape of the directional pattern of the single partials. Note, that the amplitude of the single harmonics do not necessarily scale with the mouth pressure. In particular for the 8th partial a decrease in amplitude is observed for a doubling of the blowing pressure, which corresponds to increasing the dynamic level from *forte* and *fortissimo*. In the low range, the curves almost coincide, which is an indication for the repeatability of the experiment.

3.5 Localizing phantom sources

The use of a microphone array combined with a camera allowed to localize sound sources on the bassoon. In the

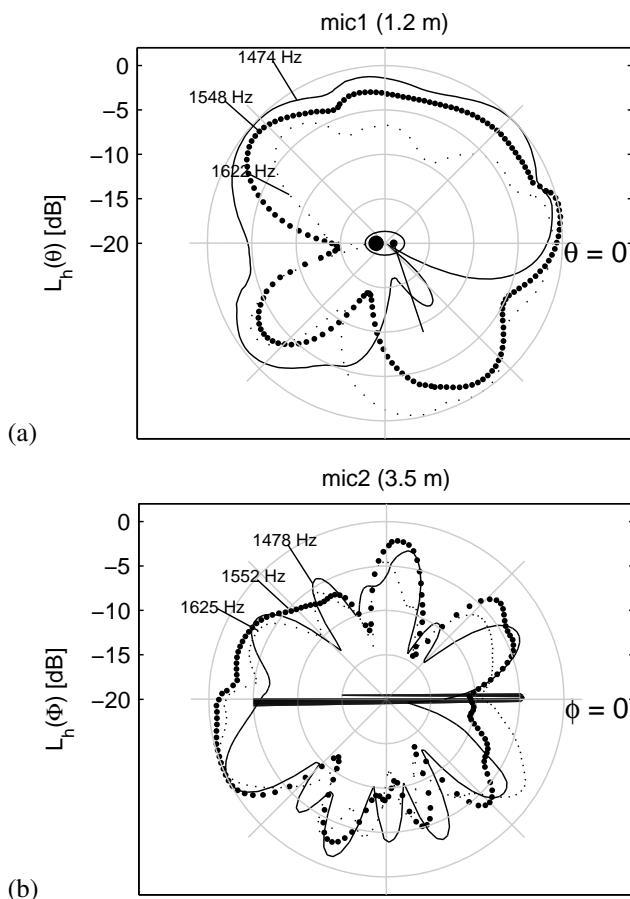


Figure 3: Azimuthal (a) and polar (b) radiation characteristics of single partials. Here are shown the 19th, 20th and 21th of the note D2 ($f_0 = 73.7$ Hz).

post-processing, the distance of the reference plane from the array is chosen such that it coincides with median planes of the rotating instrument. Under the assumption of point sources in this plane, their corresponding sound pressure level is calculated by the software. Maxima of sound pressure are mapped in red and allow to localize sound sources. Exemplary results are shown for a high note (Bb4, $f_0 = 469.3$ Hz), for which two register holes on the upper end of the bore ("wing"), are opened. At four different turning angles, the resulting sound levels in the range of 53 to 43 dB are shown in Fig. 5 (1-4) as a contour plot, overlayed to the camera picture. Note, that it is not possible to localize single tone holes. The spatial resolution of the acoustic camera is frequency-dependent: the lower the frequency the lower the radiation. Consequently, the microphone array cannot resolve small distances such as neighbored tone holes at low frequencies. The location on the color map would then be an average of several tone holes. The spatial resolution can be improved artificially by algorithms which exploit several assumptions. In this study, the standard beamforming algorithm has been applied to the measurements.

The images indicates a variable number and intensity of up to three sound sources along the length of the bassoon, depending on the azimuth angle. In the corresponding polar

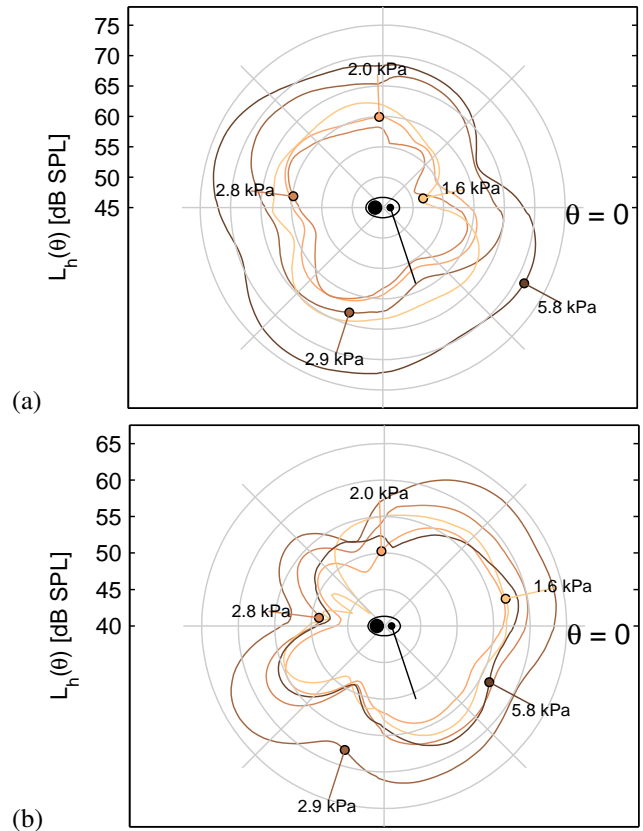


Figure 4: Influence of the playing dynamics on the azimuthal radiation characteristics of the 6th (a) and 8th (b) partial of the note Eb3 ($f_0 = 156.6$ Hz). Pressure magnitudes are displayed in dB SPL, measured in the far field (3.5 m).

plot, these sources merge to one level that varies strongly vs. frequency. For the angle 245° Fig. 5) there are two contradictory levels seen in the polar plot: the close microphone exhibits a destructive interference whereas the far microphone almost has its maximum level. For the other angles both plots are in rather good agreement. This can be explained by the superposition of two or more coherent sound sources that – at the microphone positions – interfere and add according to their relative phase. For the position of 245° this phase is exactly shifted by $\lambda/2$ between both microphone positions. At the other angles the difference has less dramatic consequences, so the level difference remains moderate.

Note, that sound sources exhibited by beamforming are phantom sound sources, which do not necessarily coincide with the real sound sources, the tone-holes. During one full turn, the phantom sources observed within the fixed dynamic level vary in quantity from one to three, as well as in location by more than 50 cm.

4. CONCLUSIONS

The bassoon is a complex multipole sound source. The artificial blowing machine allows to exclude absorbing effects of a player on the sound radiation, and thus to study

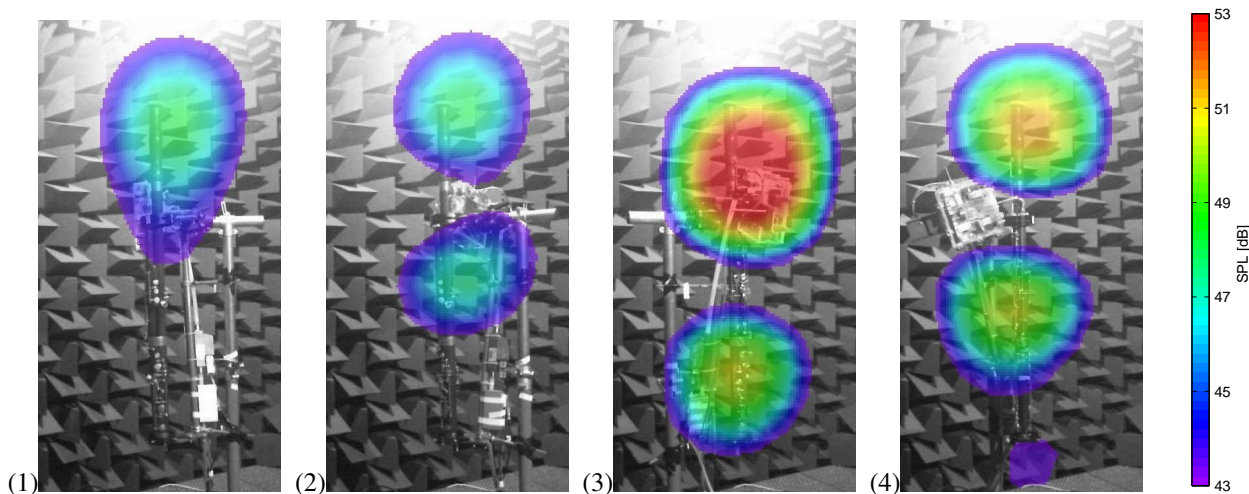
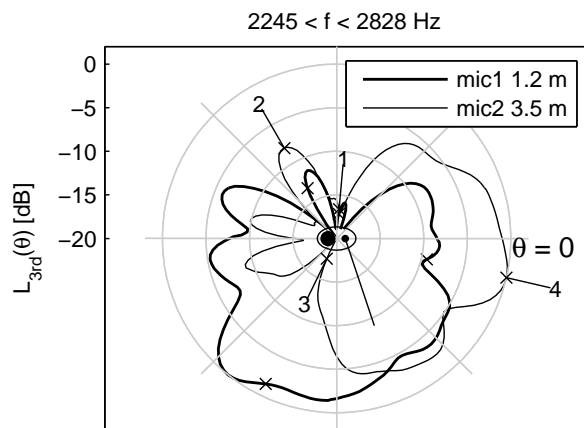


Figure 5: Azimuthal radiation characteristics of the 4th partial of the note Bb4 ($f_0 = 469.3$ Hz) and corresponding views of the acoustic camera for four different angles (1): 85°, (2): 120°, (3): 245°, (4): 345°.

the radiation characteristics of the sole instrument. Under these unrealistic circumstances a strong symmetry of the sound radiation characteristics is observed.

The directivity of the bassoon exhibits some interesting features that can be explained by its complex construction. The folded tube, together with the largely irregular configuration of tone holes provides a potential multi pole radiator with coherent sound sources for each partial. The interaction of these multi poles will lead to constructive and annihilative superposition of sound pressures at a distance. Polar diagrams sum up the participation of all sources of radiation for a given recording position. The use of the acoustic camera allowed to further interpret the occurrence of interference patterns of up to three coherent sound sources, depending on the fingering of the instrument. The identification of each relative source polarity, i.e. whether the contributions add in phase or out of phase, cannot be answered with one method alone; the combination of both, however, can explain interference patterns at various frequencies.

The variation of directivity from the recording distance exhibits similar patterns. Above 4 kHz the similarity decreases due to insufficient pressure levels at these frequen-

cies for the distant microphone. Another aspect to study in detail is the effect of the player's presence on the symmetry or asymmetry of the radiation patterns. The directivity measurements could be repeated with absorbing material covering the lower end of the bassoon to model the effect of the player. It can be expected that such measurements will result in directivity patterns that match the previous results obtained by Meyer [1].

Further investigations could use probe microphones at the position of those tone-holes that contribute to the radiation at a specific frequency to verify their relative phase relation. Together with their distance to the microphones that are used for the directivity measurements the potential effect of interference on the radiation pattern could be calculated. More in detailed studies are planned to also take into account the positions of the tone-holes relative to each other and to the horizontal plane, as well as the positions of pressure anti nodes in the bore.

Acknowledgments

Uwe Meier and Rainer Günther from the Acoustics Laboratory of university of Applied Sciences Ostwestfalen-Lippe are acknowledged for making the anechoic chamber

available for the experiments. CAE Software and Systems, Gütersloh, Germany and Nico Zurmühlen in particular is acknowledged for the extended rental period of the microphone array and support in the post-processing. We thank Bodo Besselmann from Schalltechnik Süd & Nord GmbH, Essen, Germany for providing an automated turntable for the duration of the measurements. The artificial mouth used in the experiments has been constructed at Technische Universität Dresden by one author. Jürgen Meyer is acknowledged for his kind permission to reproduce the directivity patterns from his book [1].

5. REFERENCES

- [1] J. Meyer, *Acoustics and the Performance of Music*. Springer, 2009.
- [2] S. Pelzer, M. Pollow, and M. Vorländer, “Auralization of a virtual orchestra using directivities of measured symphonic instruments,” in *Program of the 11th Congrès Français d’Acoustique and 2012 IOA annual meeting, Acoustics2012, Nantes*, 2012.
- [3] M. Kob and H. Jers, “Directivity measurement of a singer,” in *Proceedings of Forum Acusticum/ASA/DAGA in Berlin*. Acoustical Society of America, 1999, p. 2AMU_19.
- [4] T. Grothe, “Investigation of bassoon embouchures with an artificial mouth,” in *Program of the 11th Congrès Français d’Acoustique and 2012 IOA annual meeting, Acoustics2012, Nantes*, 2012.
- [5] L. Fuks and J. Sundberg, “Blowing pressures in bassoon, clarinet, oboe and saxophone,” *Acta Acustica united with Acustica*, vol. 85, pp. 267–277(11), March/April 1999.

EVALUATING A WAVELET-BASED ANALYSIS OF SENSOR REED SIGNALS FOR PERFORMANCE RESEARCH

Alex Hofmann

Institute of Music Acoustics (IWK),
University of Music
and Performing Arts Vienna,
Austria
hofmann-alex@mdw.ac.at

Werner Goebel

Institute of Music Acoustics (IWK),
University of Music
and Performing Arts Vienna,
Austria
goebel@mdw.ac.at

Michael Weilguni

Institute of Sensor and Actuator Systems,
Vienna University
of Technology,
Austria
michael.weilguni@tuwien.ac.at

ABSTRACT

Empirical investigations of saxophone and clarinet performance are important for a thorough understanding of the human motor skills required to musically operate woodwind instruments. In this paper, we discuss two methods of detecting tonguing related landmarks in a sensor saxophone reed signal. We give detail about the reed signal's characteristics under three typical playing instructions (legato, portato and staccato articulation) and define detection tasks for physical tone onsets and tone offsets. When the player's tongue contacts the reed, the oscillations are dampened and the reed is bent towards the mouthpiece (tongue-reed contact). Removing the tongue from the reed returns it to its equilibrium position (tongue-reed release). From these observations we derive two landmark detection functions: a heuristic for peak detection, based on thresholding the smoothed reed signal, and a wavelet-based analysis, operating on specific sub-bands of the reed signal. When evaluating both methods, the wavelet analysis revealed better results using our current test dataset.

1. INTRODUCTION

To investigate how humans interact with musical instruments during skilled musical performance, empirical studies require the analysis of large data sets [1]. The raw data captured during studies of music performance may include audio recordings [2], video capture [3] or signals retrieved from sensors attached to musical instruments [4, 5].

For the analysis of specific sensor signals, such as our sensor saxophone reed [6], landmark detection functions (LDF) have to be developed for each new measurement setup.

For instruments that do not provide symbolic data (i.e., MIDI), state-of-the-art algorithms from music information retrieval can be used to roughly transcribe performances from audio material [7]. However, these algorithms were developed to track perceptual note onset times and may not define physical note onsets and offsets with the precision that is required for performance analysis.

Specifically developed detection functions demand careful evaluation of the results to ensure their reliability when making statements about the underlying performance. In this paper we describe two different approaches of detecting tongue-reed interaction in sensor saxophone reed signals: A peak detection function, based on thresholding the smoothed reed signal, and a wavelet-based analysis. Finally we evaluate the analysis results of both methods under different levels of stringency.

Articulation on single-reed woodwind instruments

When playing on single-reed woodwind instruments, the sound of consecutive played tones depends on the articulation technique used [8–10]. In contrast to legato playing, when the player blows constantly and only varies the fingerings, portato articulation and staccato articulation require tongue–reed interaction in order to control note beginnings and note endings.

A reliable, automated detection of these physical note onsets and note offsets from the sensor reed signal is essential, to investigate the motor control mechanisms in expressive woodwind performance.

2. METHOD

Sensor Reed Signal Properties

Strain-gauge sensors attached to woodwind single-reeds have been shown to capture the bending of the reed during performance. These signals contain precise information about tongue–reed interactions without environmental interference [6]. We observed different reed signals for different articulation techniques (see Figure 1). Whereas in legato playing (top panel) no tongue-reed contact (TRC) occurred, tonguing in portato articulation and staccato articulation was clearly visible (middle and bottom panel). During TRC the tongue pressed the reed towards the mouthpiece and thereby dampened the reed vibrations. To start the next tone, the tongue released the reed (tongue-reed release, TRR) [10].

2.1 Signal smoothing method

In our previous study on finger and tongue coordination in saxophone performance, we analysed the sensor reed

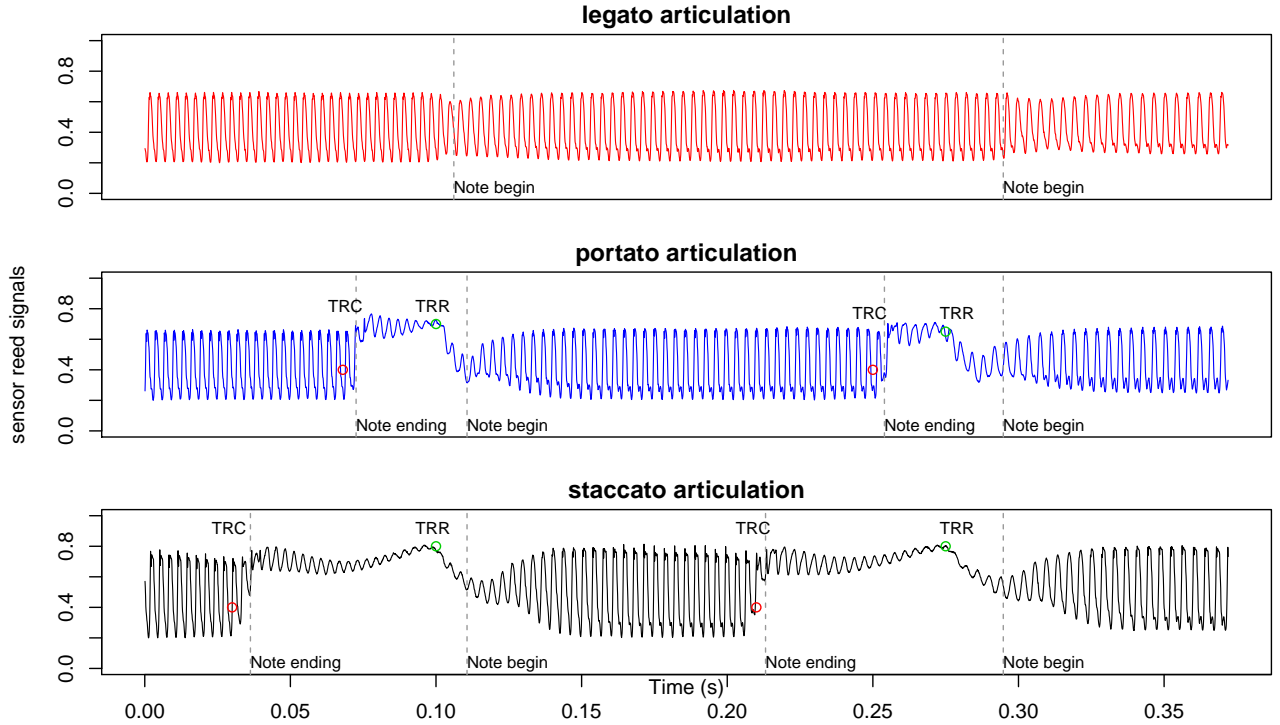


Figure 1. Alto-saxophone sensor reed signals showing the note transitions (e4–d4–c4) under a tempo instruction of 179 ms inter-onset interval (audio sampling rate 11 kHz). Legato articulation without tonguing (top); portato articulation with tongued note onsets (tongue reed release, TRR) and note offsets (tongue reed contact, TRC)(middle); staccato articulation, with extended tongue reed contact time (bottom)

signal with a rather simple approach, following two steps [11]:

Preprocessing

A smoothing function (Butterworth, low pass filter, 15 Hz) removed the high frequency oscillations, which occurred during sound production. Artefacts of the filter (i.e., phase shift) were ignored.

Landmark detection function

Local maxima above a certain threshold (40 % quantile of all maximum levels) were marked as potential TRRs and minima were marked as TRCs. Saddle points (pairs located on the same level of the smoothed signal) were automatically removed. For the calculations, external libraries (signal, msProcess) were used in the R-statistics software package [12].

The high error rate (multiple detections of one event) required a manual second step, in which the landmarks were verified by visual inspection before further processing was allowed. This time consuming procedure was possible for a relatively small data set containing 8700 tones, but is not a suitable method for further empirical studies with larger datasets.

2.2 Wavelet-based method

In the following section, we will discuss a LDF based on wavelet decomposition of the reed signal. A multiresolution analysis (MRA) is considered to be a suitable tool for time critical signal analyses (i.e., drum pattern transcription in audio material) [13–15]. The computational efficient pyramid algorithm calculates the wavelet coefficients, which represent the energy distribution over time and frequency, with $O(N \log_2 N)$ [16]. The external libraries (wmtsa, msProcess) were used in the R-statistics software package [12].

Preprocessing

The reed signal was decomposed into eleven sub-bands using the Maximal Overlap Discrete Wavelet Transform (MODWT, $J_0 = 11$). The advantage of the MODWT over a Discrete Wavelet Transform (DWT) is that the MODWT is well defined for any sample size and the resulting sub-bands (details \tilde{D}_J and smooth \tilde{S}_{J_0}) are associated with zero phase filters. The Daubechies least asymmetric 8-tap filter LA(8) was chosen, because its phase properties allow direct reference from the MODWT coefficients to actual times in the original signal [16].

The choice of level $J_0 = 11$ allows investigations of relevant sub-bands, starting from the smooth with a horizontal spacing between individual coefficients with \tilde{S}_{J_0} : $\lambda_{11} \Delta t = 2^{11} \cdot \frac{1000 \text{ ms}}{11025 \text{ Hz}} = 185.76 \text{ ms}$ (5.38 Hz). Large

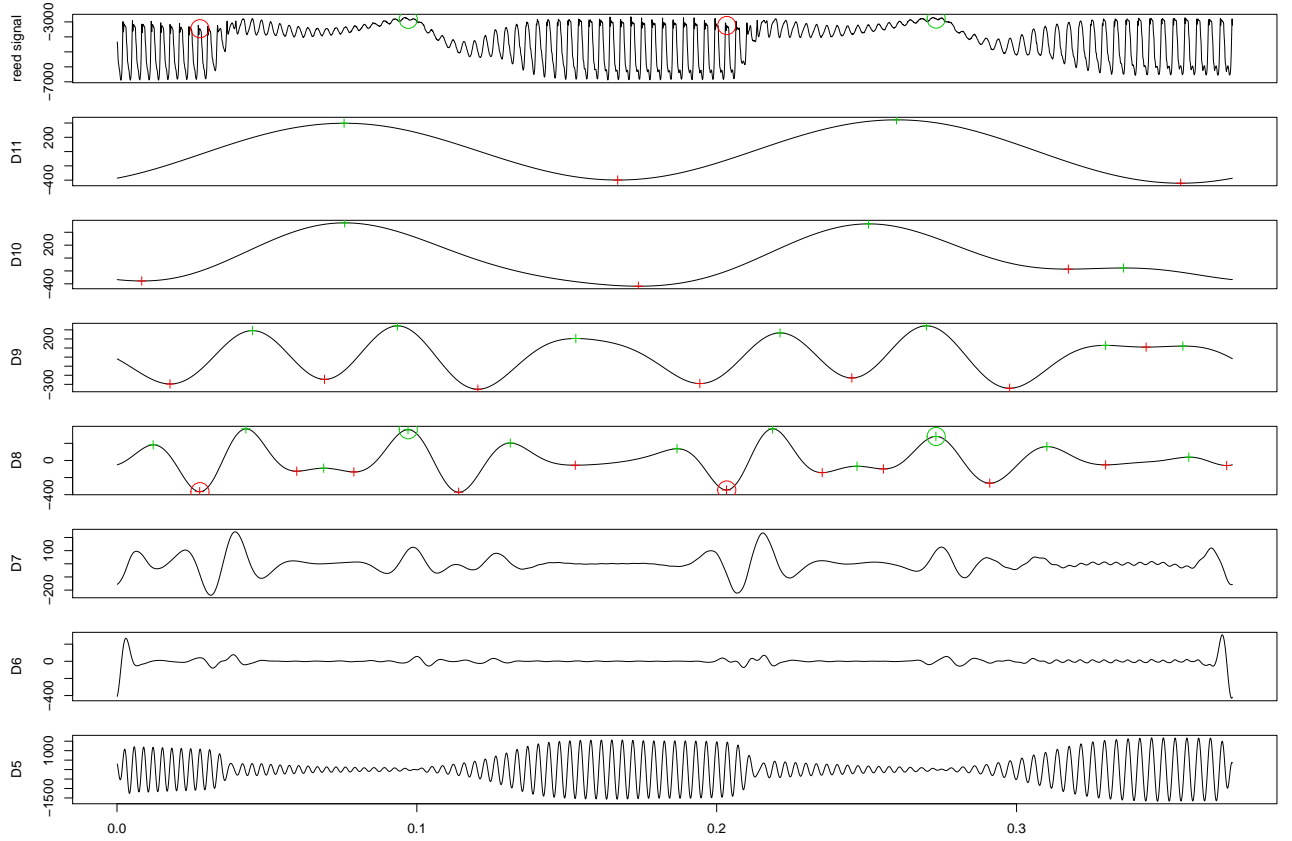


Figure 2. Maximal Overlap Discret Wavelet Transform of sensor reed signal showing staccato articulation: The figure shows the input signal (top) with detected landmarks (TRC: red circle, TRR: green circle) and the details \tilde{D}_{11-5} (below). The D11–D8 landmark detection function labeled maxima (green) and minima (red) in detail \tilde{D}_{11} . These positions were refined to extrema of \tilde{D}_{10} , \tilde{D}_9 and \tilde{D}_8 .

scale fluctuations, caused by temperature effects to the strain gauge [17], were thereby isolated in \tilde{S}_{J_0} and do not further affect the analysis of the details \tilde{D}_J .

The horizontal time spacing of the multiresolution analysis (MRA) sub-bands can be calculated, by seeing the MODWT as a constant-Q filter bank with octave spaced centres of the filters. Each higher sub-band contains double the time resolution as the previous sub-band ($\tilde{D}_{11}: \tau_{11}\Delta t = 92.88\text{ ms}$; $\tilde{D}_{10}: \tau_{10}\Delta t = 46.44\text{ ms}$; $\tilde{D}_9: \tau_9\Delta t = 23.22\text{ ms}$; $\tilde{D}_8: \tau_8\Delta t = 11.61\text{ ms}$).

Landmark detection function

\tilde{D}_{11} was used to detect the reed displacement caused by the tongue. Maxima of \tilde{D}_{11} (maximum displacement of the reed towards the mouthpiece) were marked as TRR, because the following signal decrease is an indicator that the player released the tongue. As a logical consequence, a TRC must have happened before a TRR. Consequently, minima were labelled as TRC and maxima as TRR. To precise the position, these labels were shifted to the extrema in sub-bands with a better time resolution (\tilde{D}_{10} , \tilde{D}_9 and \tilde{D}_8 : $\tau_8\Delta t = 11.61\text{ ms}$). This LDF will be abbreviated as D11–D8.

Figure 2 depicts the wavelet-based landmark detection for staccato tone transitions: First, maxima and minima of \tilde{D}_{11} were labelled. These rough landmarks were then

refined to extrema of \tilde{D}_{10} , \tilde{D}_9 and afterwards to those of \tilde{D}_8 .

3. EVALUATION

3.1 Dataset

Our test dataset contained 1744 visually annotated landmarks, based on sensor reed signals similar to the material from our previous study [11]. For this evaluation, we used eighth-note melodies, played with portato and staccato articulation, in three tempo conditions (IOI = 250 ms, 178.6 ms, 144.2 ms), performed by six alto-saxophone players in our laboratory.

3.2 Measures

To compare both LDFs, the standard measures *precision*, *recall* and *F-measure* were used. *Recall* describes the completeness of the search and *precision* gives status about the quality of the search results. *F-measure* combines the two previous measures by the following equation:

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (1)$$

Starting from the annotated ground truth, the existence and number of detected landmarks around the annotated

Landmark Det. Func. WindowSize	% F-meas. $\pm 25\text{ ms}$	% Prec. $\pm 25\text{ ms}$	% Rec. $\pm 25\text{ ms}$	% F-m. $\pm 15\text{ ms}$	% Prec. $\pm 15\text{ ms}$	% Rec. $\pm 15\text{ ms}$	% F-m. $\pm 10\text{ ms}$	% Prec. $\pm 10\text{ ms}$	% Rec. $\pm 10\text{ ms}$
TRC–Sig. smoothing	78.7	74.5	83.3	43.3	41.1	45.9	3.3	3.1	3.4
TRR–Sig. smoothing	85.5	81.0	90.5	72.3	68.5	76.5	33.3	31.5	35.2
TRC–Wavelet analysis	95.2	95.5	94.8	90.1	90.5	89.8	55.0	55.2	54.8
TRR–Wavelet analysis	95.4	95.7	95.1	76.6	76.9	76.4	51.2	51.4	51.0

Table 1. *F-measure, precision and recall* for both landmark detection functions proposed in Section 2. Results for both tasks: TRC (tongue-reed contact, note offset) and TRR (tongue-reed release, note onset) detection within a $\pm 25\text{ ms}$, $\pm 15\text{ ms}$ and $\pm 10\text{ ms}$ evaluation window are given and explained in Section 4.

events was checked. A single detection was counted as one true positive, whereas double detections were considered as one true positive and one false positive. A missing landmark was counted as one false negative. Remaining landmarks, not matched to annotated events, were counted as false positives. The strictness of such an evaluation is defined by the size of the evaluation window. Usually a $\pm 25\text{ ms}$ evaluation window is used to check onset-detection functions working on percussive material [7]. For articulation detection, a higher accuracy of the detected events might be necessary, so $\pm 15\text{ ms}$ and $\pm 10\text{ ms}$ evaluation windows were additionally considered.

4. RESULTS

Comparison of both LDFs

Table 1 shows the results for the two LDFs described in Section 2. For the smoothing method (see Section 2.1.), only the automated detection of potential landmark positions was applied to the dataset, without manual correction steps.

Overall, the wavelet-based analysis gave better results in the *F-measure* of both detection tasks ($> 95\%$ within a $\pm 25\text{ ms}$ evaluation window). The wavelet-based analysis also gave better results for precision ($> 95\%$ within a $\pm 25\text{ ms}$ evaluation window) than the smoothing method ($> 74\%$). False positive detections, a main problem of the smoothing method, were reduced by using the wavelet LDF.

The choice of the evaluation window had a significant influence on the quality measure. Within the $\pm 25\text{ ms}$ evaluation window both LDFs performed quite well (Smoothing LDF *F-measure*: $> 78\%$; Wavelet LDF *F-measure*: $> 95\%$). A reduction of the evaluation window to $\pm 15\text{ ms}$ showed a significant difference in the behaviour of both methods. The wavelet LDF showed a better score for TRC detection (*F-measure*: 90.1%) compared to the smoothing LDF (*F-measure*: 43.3%). Both LDFs showed similar results for TRR detection (Smoothing LDF *F-measure*: 72.3% ; Wavelet LDF *F-measure*: 76.6%). A further reduction of the evaluation window to $\pm 10\text{ ms}$ showed clear superiority of the wavelet approach’s accuracy, especially in TRC detection (Smoothing LDF *F-measure*: 3.3% ; Wavelet LDF *F-measure*: 55.0%).

5. DISCUSSION

We developed two different methods to extract physical note onsets (and note offsets) from sensor saxophone reed signals, with the aim of enabling automated examination of large datasets of woodwind performances.

We compared the reliability of both proposed detection functions under different parameters of detection accuracy and found that the wavelet-based LDF outperformed the signal smoothing method.

Direct comparisons of our detection results with state-of-the-art onset detectors are not possible because these algorithms were evaluated on larger testset, containing various types of musical recordings from different genres and environments (*F-measure* between 70% and 87% [7, 18]).

The archived *F-measure* of $> 90\%$ (for TRC detection with wavelets, within a $\pm 15\text{ ms}$ evaluation window) meets our criteria and will be used in future studies to analyse saxophone and clarinet performances’ sensor reed signals. A detailed examination of the reasons for the decreasing *F-measure* of 76.6% for the TRR detection task is intended.

Finding the right balance between a sufficiently powerful, but not over-fitted detection function still remains a difficult task. The current wavelet LDF (D11–D8) is designed to accommodate different playing speeds and articulation techniques, but may be limited to sensor reed signals within a certain amplitude range. To further investigate the question of the detection quality, an evaluation with complex musical pieces, including varying dynamics, is required.

In the future, we aim to optimize our approach towards an online articulation detection function. This may enable the development of woodwind sensor instruments, which provide feedback about the actual performance in a learning situation or can be used as interfaces to physical modelling based sound synthesis in contemporary music performances.

Acknowledgement

This research was supported by the Austrian Science Fund (FWF): P23248-N24. The authors made extensive use of free software under GNU/Linux. Further, we would like to thank Kai Siedenburger for pointing us to wavelet analysis.

6. REFERENCES

- [1] A. Gabrielsson, “Music performance research at the millennium,” *Psychology of music*, vol. 31, no. 3, pp.

- 221–272, 2003.
- [2] E. Scheirer, “Using musical knowledge to extract expressive performance information from audio recordings,” in *Readings in Computational Auditory Scene Analysis*, H. Okuno and D. Rosenthal, Eds., Citeseer. Erlbaum, Mahwah, NJ, 1995, pp. 153–160.
- [3] W. Goebel and C. Palmer, “Temporal control and hand movement efficiency in skilled music performance,” *PLOS ONE*, vol. 8, no. 1, p. e50901, 2013, (open access).
- [4] A. Almeida, R. Chow, J. Smith, and J. Wolfe, “The kinetics and acoustics of fingering and note transitions on the flute,” *The Journal of the Acoustical Society of America*, vol. 126(3), pp. 1521–1529, 2009.
- [5] H. Kinoshita and S. Obata, “Left hand finger force in violin playing: Tempo, loudness, and finger differences,” *The Journal of the Acoustical Society of America*, vol. 126(1), pp. 388–395, 2009.
- [6] A. Hofmann, V. Chatziioannou, M. Weilguni, W. Goebel, and W. Kausel, “Measurement setup for articulatory transient differences in woodwind performance,” vol. 19, no. 1. ASA, 2013, p. 035060. [Online]. Available: <http://link.aip.org/link/?PMA/19/035060/1>
- [7] S. Böck, F. Krebs, and M. Schedl, “Evaluating the online capabilities of onset detection methods,” in *In Proceedings of the 13th International Society for Music Information Retrieval Conference, Porto, Portugal.*, 2012.
- [8] K. Krautgartner, “Untersuchungen zur artikulation bei klarinetteninstrumenten im jazz,” Ph.D. dissertation, Universität zu Köln, 1982.
- [9] D. Liebman, *Developing a personal saxophone sound*. Medfield: Dorn Publications, 2006.
- [10] A. Hofmann, V. Chatziioannou, W. Kausel, W. Goebel, M. Weilguni, and W. Smetana, “The influence of tonguing on tone production with single-reed instruments,” in *5th Congress of the Alps Adria Acoustics Association*, Zadar, Croatia, 2012.
- [11] A. Hofmann, W. Goebel, M. Weilguni, A. Mayer, and W. Smetana, “Measuring tongue and finger coordination in saxophone performance,” in *12th International Conference on Music Perception and Cognition (ICMPC) and 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*, Thessaloniki, Greece, 2012, pp. 442–445.
- [12] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012. [Online]. Available: <http://www.R-project.org>
- [13] G. Tzanetakis, G. Essl, and P. Cook, “Audio analysis using the discrete wavelet transform,” in *Proc. Conf. in Acoustics and Music Theory Applications*, 2001, pp. 318–323.
- [14] R. Kronland-Martinet, J. Morlet, and A. Grossmann, “Analysis of sound patterns through wavelet transforms,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 1, no. 02, pp. 273–302, 1987.
- [15] A. Paradzinets, H. Harb, and L. Chen, “Use of continuous wavelet-like transform in automated music transcription,” in *Proceedings of EUSIPCO*, 2006.
- [16] D. Percival and A. Walden, *Wavelet Methods for Time Series Analysis*, ser. Cambridge Series In Statistical And Probabilistic Mathematics. Cambridge: Cambridge University Press, 2006.
- [17] A. Window, *Strain Gauge Technology*, 2nd ed., ser. Elsevier Applied Science. Essex: Elsevier Science Publishers, 1989.
- [18] S. Böck, A. Arzt, F. Krebs, and M. Schedl, “Online real-time onset detection with recurrent neural networks,” in *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12)*, York, UK, 2012.

THE VOICE OF THE MECHANICAL DRAGON

Michel Hirschberg, Oleksii Rudenko

MTP, Dept. Applied Physics,
Technische Universiteit Eindhoven,
Postbus 513, 5600 MB Eindhoven
O.Rudenko@tue.nl

Gunes Nakiboglu, Ad Holten

Gunesnakib@gmail.com

Jan Willems, Avraham Hirschberg

A.Hirschberg@tue.nl

ABSTRACT

The sound radiation of a swinging corrugated tube (Hummer) has been measured under anechoic and semi-anechoic conditions. The whistling is induced by synchronized vortex shedding at each corrugation coupled to an acoustic standing longitudinal wave. In an earlier paper the Hummer was hand-driven. In order to eliminate this human factor, the instrument is driven mechanically. This considerably enhances the agreement between measured sound radiation and prediction by a model assuming at the open ends acoustic velocity fluctuations of 5% of the main flow velocity through the tube. These pipe terminations act as monopole sound sources. Significant deviations from theory still remain. In particular the amplitude modulation of the sound, due to interference between the two sources, is deeper than predicted. Experiments also show the possible coexistence of two acoustic modes, which is not considered in the theory. The possible elongation of the tube by centrifugal force and flutter due to vortex shedding appears to be negligible.

1. INTRODUCTION

In an earlier paper [1] we considered the sound production and radiation of a swinging corrugated tube. This musical toy is called a Hummer [2]. In musical applications it is called the Voice of the Dragon [3], [4]. The Hummer considered here has a length $L = 0.750\text{m}$ and an inner diameter $D = 25\text{ mm}$. The corrugation pitch length is $W_p = 7.0\text{ mm}$. The pipe inlet has a horn shape ($L_{con} = 20\text{ mm}$) followed by a smooth pipe segment ($L_{smth} = 30\text{ mm}$), forming a handle (Fig. 1).

When the pipe is held at one end and swung, it acts as a centrifugal pump. This induces a flow through the pipe with a velocity U [1]. Vortex shedding, at each cavity formed by the corrugations, couples with longitudinal acoustic standing waves: modes of the open-open tube with resonant frequency $f_n \simeq nc_{eff}/(2L)$ ($n = 1, 2, \dots$), where c_{eff} is the effective speed of sound in the tube [1]. This synchronizes the vortex shedding along the pipe for critical ranges of the Strouhal number $f_n W_p/U$. The resulting

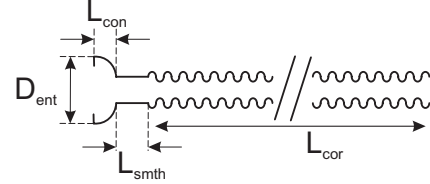


Figure 1. Geometry of the Hummer: $L = L_{for} + L_{smth} + L_{con} = 75\text{ cm}$, $D = 25$ and $D_{ent} = 40\text{ mm}$.

loud whistling saturates as a result of non-linearity. The acoustic velocity fluctuations u' , have then typically a dimensionless amplitude, u'/U , of the order of 5%, where u' is the amplitude of the velocity oscillation at the open pipe terminations. One expects that the non-linear saturation mechanism, which stabilizes the limit cycle, prevents the coexistence of two acoustic modes [5], [6], [7]. Therefore the internal acoustic field should be a line spectrum dominated by multiples of the fundamental $f = mf_n$ ($m = 1, 2, \dots$).

The open terminations radiate as coherent monopoles. The radiation is isotropic because $2\pi f_n D/c_0 \leq 0.5$, where c_0 is the speed of sound in free space [8]. The interference pattern induced by these two sources rotates with the pipe. In addition to that the rotation induces a Doppler effect. The perceived frequency of the rotating sound source varies up to a semi-tone. A theoretical model (section 2) taking these effects into account has been proposed in our earlier paper [1]. The radiated sound spectrum is accurately predicted, but the model was unable to explain the observed amplitude modulation. As the instrument was swung by hand, we expect that the rotation speed cannot be constant. Also the position of the microphone is difficult to determine accurately. Furthermore the pipe is bended and oscillates. We therefore expect a variation in the flow velocity, U , through the pipe. As observed by Kristiansen [9] such a slow modulation of the flow velocity U induces a complex acoustic response of the pipe. There is a considerable delay between variations in U and the modulation in the amplitude u' of the acoustic oscillation.

In order to eliminate such uncertainties, experiments have been carried out with a mechanically driven Hummer. The experiment is inspired by the pioneering work of Silverman and Cushman [3]. The Hummer is mounted on a bicycle wheel driven by an electrical motor (section 3). The effect of reflections on the floor of the semi-anechoic room was investigated by placing a sound absorbing mat

on the floor between the Hummer and the microphone. The present paper describes primarily the results obtained (section 4).

We furthermore discuss some mechanical aspects of the problem (section 5). It has been suggested that the elongation of the Hummer under influence of the centrifugal acceleration results in a flattening of the pitch when increasingly high pipe modes are excited [10]. Another mechanical problem is that vortex shedding in the wake of the rotating tube can couple with mechanical oscillation modes of the tube, resulting into flutter and associated transversal motion of the tip of the tube.

2. THEORY

Based on the measurements presented by Nakiboglu et al. [1] we estimate the amplitude, u' , of the velocity fluctuations at the open pipe termination from:

$$u'/U = 0.05. \quad (1)$$

The flow velocity U is calculated from the measured angular rotation velocity Ω :

$$U = \Omega \sqrt{\left(\frac{R_2^2 - R_1^2}{1 + 4c_f \frac{L}{D}} \right)} \quad (2)$$

where R_1 is the radius of rotation of the pipe inlet, R_2 is the radius of rotation of the pipe outlet, L is the pipe length, $c_f = 0.0178$ is the friction coefficient of the pipe measured by Nakiboglu et al. [1]. The amplitude Q of the oscillating monopole sources is:

$$Q = u' \frac{\pi D^2}{4} \quad (3)$$

The pressure oscillation $p'_i(\vec{x}, t)$ radiated by monopole Q_i ($i = 1, 2$ from position \vec{x}_i , reaching the observer at \vec{x} at time t is given by [11]:

$$p'_i(\vec{x}, t) = (-1)^{i(n+1)} \frac{\partial}{\partial t} \left[\frac{Q \cos(\omega t_{e,i})}{4\pi |\vec{x} - \vec{x}_i(t_{e,i})| |1 - M_i(t_{e,i})|} \right] \quad (4)$$

where the angular frequency is $\omega = 2\pi f_n$ and the emission time is found by solving the equation:

$$t_{e,i} = t - \frac{|\vec{x} - \vec{x}_i(t_{e,i})|}{c_0} \quad (5)$$

In principle the oscillation frequency is close to the pipe resonance frequency $f \simeq f_n$. In our model, we will however use the measured oscillation frequency f , corresponding to the peak in the spectrum of the radiated sound. The formula takes into account that the oscillating velocities for even modes $n = 2, 4, 6, \dots$ have opposite phase at inlet and outlet of the pipe. The source position is given by:

$$\vec{x}_i(t_{e,i}) = (R_i \cos(\Omega t_{e,i}), R_i \sin(\Omega t_{e,i} + \phi_i), z_i) \quad (6)$$

In our experiments the Hummer is in the horizontal plane $z_i = h$. The relative Mach number is given by:

$$M_i = \frac{1}{c_0} \frac{\partial |\vec{x}_i(t_{e,i}) - \vec{x}|}{\partial t_{e,i}} \quad (7)$$



Figure 2. Fixation of the Hummer on the Wheel and detection of wheel rotation.

where c_0 is the speed of sound of ambient air. The acoustic field without reflections on the floor is simply the sum of the two contributions p'_i . When the floor is reflecting we have to add to this the contribution of the images with $z_i^* = -h$ at $\vec{x}_i^* = (R_i \cos(\Omega t_{e,i}^*), R_i \sin(\Omega t_{e,i}^* + \phi_i), z_i^*)$. In order to calculate the various contributions at the same observers time, we can solve equation (5) for the emission times for each source. Alternatively we can calculate for given emission times the contributions as functions of the observer time t . The reception time t is calculated explicitly for each source from equation (5) for given $t_{e,i}$. The acoustic field at a given observer time is obtained from these data by interpolation.

3. EXPERIMENT

The Hummer was attached to the spikes of a 26" bicycle wheel. The inlet $i = 1$ was fixed by means of a plastic ring close to the axis of the wheel ($R_1 = 0.10$ m). Another plastic ring was fixing the tube at the end of a spike (Fig. 2). The outlet $i = 2$ is at a distance $R_2 = 0.71$ m from the wheel axis. At rest $\phi_2 = 0$ and $\phi_1 = -2\pi/3$. The positions of the monopole sources are corrected for an estimated end-corrections of the pipe, $\delta = 0.3D$ cm for each side. Therefore the distance $|\vec{x}_2 - \vec{x}_1|$ exceeds the pipe length by 15 mm.

The angular velocity Ω is positive in the (x, y, z) right-handed coordinate system with the vertical z -axis along the wheel axis. As shown in the photograph (Fig. 3) a second Hummer is placed on the wheel. This Hummer is plugged and does not whistle but provides mechanical balance to the wheel.

The wheel is mounted on a horizontal optical bench, a 1 m long aluminum tube/rod of "square" cross section (0.1 m x 0.1 m). This horizontal rod mounted to a 1.90 m tall vertical rod. The sides of these hollow rods are plugged by sound absorbing foam. The wheel axis is at 0.2 m from the front end of the horizontal rod. The vertical rod is at 0.6 m from this front end. The electrical motor is fixed to the horizontal rod near the back end. A bicycle chain is used to transmit the rotation to the wheel. The Hummer lays at the height $h = 2.23$ m above the floor. The axis of the wheel in a corner of the semi-anechoic room at about 1 m from both walls. The semi-anechoic room has a volume of 100 m³



Figure 3. Overview of the set-up with reflecting floor.

and a cut-off frequency of 300Hz. A wind tunnel nozzle and a table, which could not be removed from the room, are placed at the opposite side of the room. They are covered by acoustic foam (10 cm thick mattresses). Unfortunately the table forms a resonating cavity with an height of 0.54 m which pollutes results with frequencies around 630 Hz. The frequency of this half wave length mode of the table almost coincides with the frequency $f_3 = 635$ Hz of the third acoustic mode of the Hummer.

A "distant" microphone (BK 1/2" type 4190) is placed at $\vec{x}_D = (1.58, 0, 1.68)$ m. An identical "close" microphone is placed at $\vec{x}_C = (0.42, 0.72, 2.03)$ m. The holder of the close microphone can be seen on the right of the set-up (Fig. 3). The microphone signals are preconditioned by an amplifier/batteries power supply (BK type 5935L). A four channel 16 bit ADC (National Instruments) is used for data acquisition. The signals were sampled at 10 kHz for 20 s. One channel is used to detect the electrical signal picked up by a coil upon passage of a small magnet fixed to one of the spikes. This signal is used to measure the period of rotation of the wheel $T = 2\pi/\Omega$.

Experiments were carried out with and without a sound absorbing mat placed halfway between the distant microphone and the vertical support rod. The mat of porous foam had a surface of 1 m² and was 8 cm thick.. It was laid 8 cm above the ground on a few 10 cm wide strips of the same foam material. On top of the mat three (30 cm high sound absorbing) wedges as the ones covering the wall are placed (Fig. 4). Once the rotation speed of the wheel was established a first measurement was carried out with this non-reflecting floor. The mat was then removed without stopping the wheel and the experiment was repeated with reflecting floor at the same rotation speed. The rotation speed was chosen to obtain a loud stable pure whistling tone.

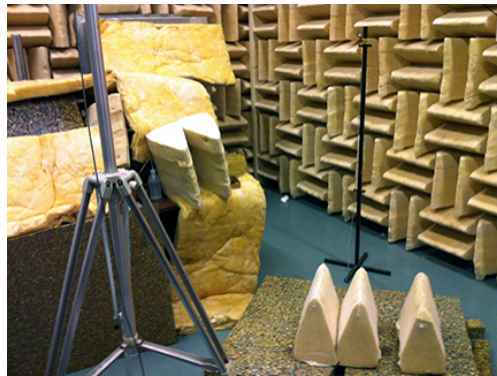


Figure 4. Sound absorbing mat on the floor between the wheel and the distant microphone.

Most experiments were carried out with the original Hummer with handle near the wheel axis. Some additional measurements were carried out with inverted Hummer and sound absorbing floor. A final series of measurements was carried out with a Hummer from which the horn had been removed (2 cm shorter pipe).

4. RESULTS

In Fig. 5 we compare the measured and predicted signals of the close microphone for $n = 2, 3, 4$ and 5, the four whistling modes, for the non-reflecting floor. In Fig. 6 we show the same results for the distant microphone.

Stable signals were obtained for the acoustic modes $n = 2, 3$ and 4. The first mode ($n=1$) did not whistle. In some experiments the fifth mode ($n = 5$) is "polluted" by a significant contribution of the fourth mode ($n = 4$). Both modes were stable and are clearly seen in the spectrum (Fig. 7). This contradicts the assumption that due to non-linear saturation effects, two modes cannot coexist. Increasing the rotation velocity one obtains a "pure" mode $n = 5$ oscillation.

Globally the theory does reasonably predict the Sound Pressure Level (within 2 dB). As explained above, the resonance of the air under the table did pollute the results around 630 Hz, which corresponds to f_3 . The theory does qualitatively well predict the signals for modes $n = 2$ and 4 in absence of reflections on the floor. A difference is that the measured modulation of the amplitude of the signals is deeper than the predicted modulation. For higher modes $n = 4$ and 5 we observe more modulations than predicted by theory.

We attempted to obtain a better prediction by assuming different monopole strength or modify the monopole positions. A difference in monopole strength is not impossible as we do not have necessarily a pure standing wave in the tube. A difference in monopole strength did not fully explain the observed difference between theory and experiments. We only show results of theory assuming two equal monopole strength. Furthermore the experiments with inverted Hummer did not show important differences compared to the results with normal Hummer position (Fig. 8). The most important difference between theory and experi-

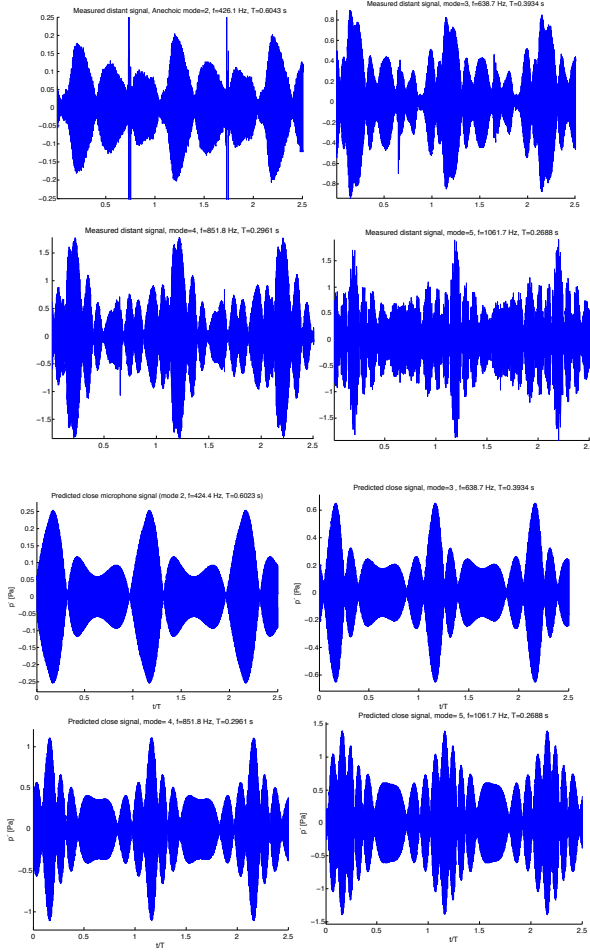


Figure 5. Close microphone signals as function of $t/T = tf$ for non-reflecting floor: experiment (four upper traces) and theory (four lower traces). In each block (experiment and theory) we have mode $n = 2$ at the upper left, $n = 3$ upper right, $n = 4$ lower left and $n = 5$ lower right. We observe the strange behavior of the signal for $n = 3$, which is expected to be due to resonant reflections by the table around 630 Hz..

ments, is the deeper modulation of the amplitude observed in the experiments. This modulation is dominated by the interference between the sources. This effect depends critically on the distance between the two sources compared to the acoustic wave length. A distance of exactly an integer number of half wave length would induce much stronger modulations than a shorter distance. As $c_{eff} < c_0$ we do expect a shorter distance. Furthermore the drag force of the flow around the pipe will bend the pipe reducing further the distance between the two sources. We therefore do not understand the observed deeper modulation. In Fig. 9 we show the measured signals of the distant microphone for a reflecting floor. The signals of the close microphone, which are not significantly affected by the reflections on the floor are not shown. The reflections on the floor increase the complexity of the observed signals. In view of the significant effect of reflections on the signal one should consider further the possible influence of the rods support-

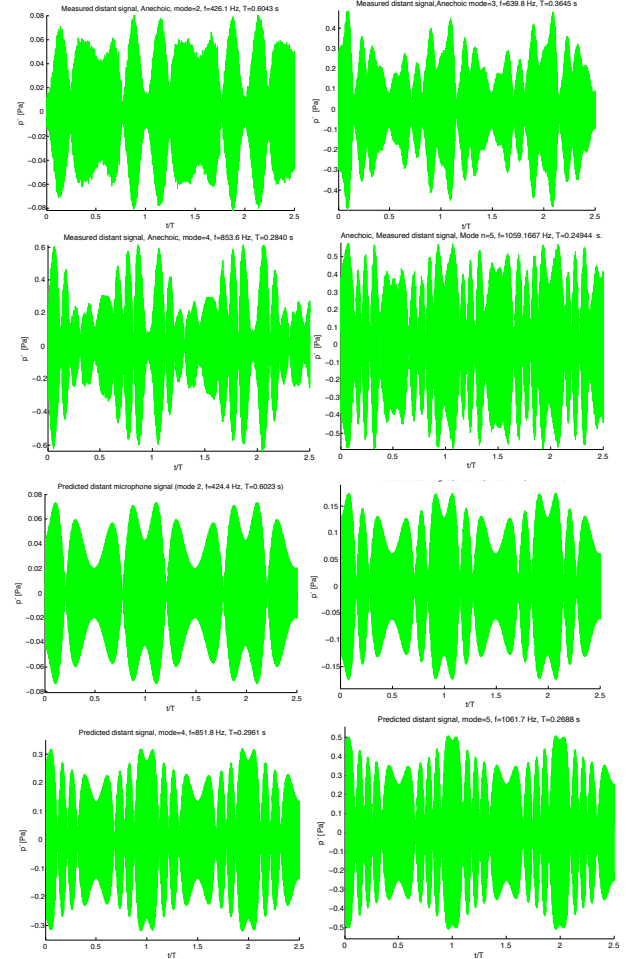


Figure 6. Distant microphone signals as function of $t/T = tf$ for non-reflecting floor: experiment (four upper traces) and theory (four lower traces). In each block (experiment and theory) we have mode $n = 2$ at the upper left, $n = 3$ upper right, $n = 4$ lower left and $n = 5$ lower right. The modulation of the measured signal is stronger than of the predicted signal.

ing the wheel on the measured signal (especially in the case of a non-reflecting floor).

5. MECHANICAL EFFECT

We now consider two mechanical effects.

It was argued by Hartman [10] that a flattening of the pitch can be observed upon increase of the whistling mode of the Hummer. This was explained as a result of the increase in pipe length due to the centrifugal force. Nice stable loud whistling was found for: $f_2 = 426.3$ Hz, $f_3 = 635.7$ Hz, $f_4 = 851.1$ Hz and $f_5 = 1061.5$ Hz. We have: $2f_3/(3f_2) = 0.994$, $f_4/(2f_2) = 0.998$ and $2f_5/(5f_2) = 0.996$, so that within the measurement accuracy we do not observe any flattening. Hanging a 1.2 kg weight to the pipe corresponding to the centrifugal force for $n = 5$ we only observed an elongation of 4 ± 1 mm, which is indeed negligible compared to the pipe length ($L = 750$ mm). A flattening in pitch of 5% (corresponding to an elongation of 4 cm) as reported by [10], [13] is expected to be due

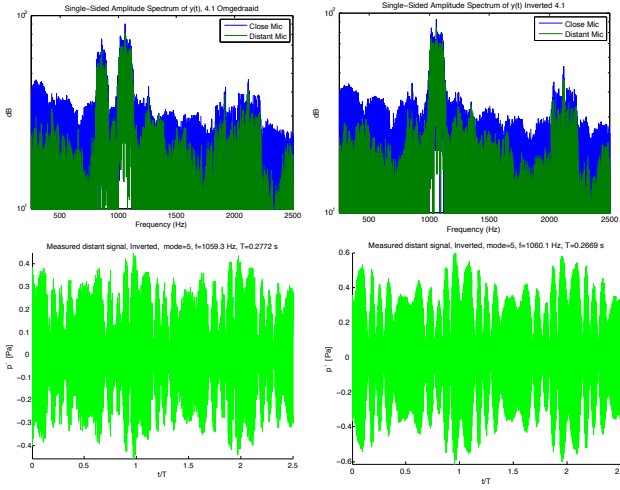


Figure 7. Left: Spectrum and distant signal p' for $n = 5$ showing the coexistence with the $n = 4$ mode. Right: same for stable $n = 5$ mode.

to a larger flexibility of the pipe used in Hartman's experiments. Also the way our pipe is fixed to the wheel can reduce the elongation due to centrifugal acceleration.

Some of the deviations between theory and experiments could be due to transversal oscillation of the free pipe end driven by vortex shedding in the wake of the rotating pipe. The Von Karman vortex shedding frequency f_{VK} corresponds to a Strouhal number $St_D = 0.2$ based on the outer diameter of the pipe $D_{outer} = 33$ mm:

$$f_{VK} = 0.2 \frac{\Omega R_2}{D_{outer}} \quad (8)$$

For $n = 3$ we find $f_{VK} = 60$ Hz. Assuming a clamped beam, we find for the first pipe bending mode for the pipe segment outside the wheel [12]:

$$f_0 = 0.55966 \sqrt{\frac{EI}{(R_2 - R_w)^4 \rho}} \quad (9)$$

We measured $EI = 0.62$ kg m³s⁻² by bending the pipe statically. We measured the pipe weight per unit length and found $\rho = 0.065$ kg m⁻¹. We find $f = 11$ Hz. The next mechanical resonance of a clamped beam is $6.267f_0$. This would correspond to 70 Hz, which is not far from the Von Karman vortex shedding frequency. Hence we cannot exclude significant flow induced vibration of the pipe due to vortex shedding. Such vibration would result into vertical oscillation. Note that the assumption of a clamped beam is not accurate. We furthermore do observed on video pictures of the experiments oscillations of the tip of the pipe. So a more careful study should be carried out.

In the case of a Hummer driven by hand the mechanical resonance modes are very low, so that we do not expect such a flow induced vibration. However one very clearly observes that the acceleration of the pipe by the movement of the hand induces bending traveling waves along the tube. Hence a model assuming a simple uniform rotation speed of the pipe termination is certainly not accurate.

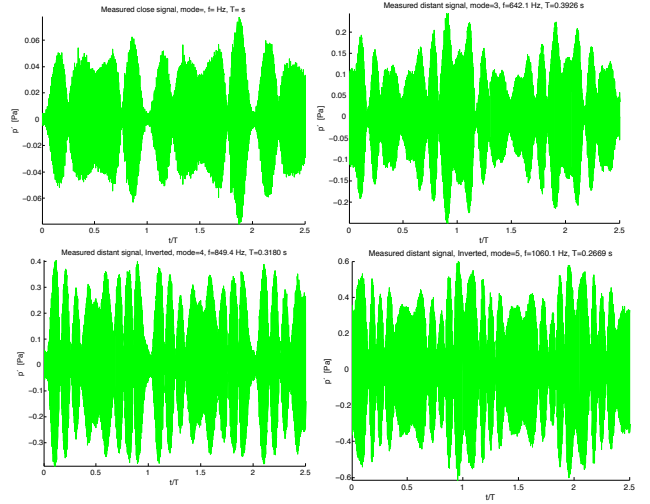


Figure 8. Distant microphone signals p' for inverted pipe: $n = 2$ (upper left), $n = 3$ (upper right), $n = 4$ (lower left) and $n = 5$ (lower right) for non-reflecting floor as functions of t/T .

6. CONCLUSIONS

Measurements of the sound radiated by a mechanically driven Hummer demonstrate that a simple theory assuming the radiation of two equally strong monopoles (in phase or opposite phase depending on the acoustic mode) explains most of the amplitude modulation of the sound radiated by a Hummer in an anechoic environment. While theory explains qualitatively the observed signal, it does not predict the depth of the modulations observed in the experiments.

Reflections on a floor or wall dramatically increase the complexity of the signal. We therefore suspect the reflections to be an important cause of deviation between theory and experiment. One cannot for example exclude that part of the deviation between theory and experiment is due to the reflections or diffraction of the sound on the rods supporting the set-up.

Experiments with modified pipe terminations (inverted Hummer or Hummer without horn) did not show a dramatic difference with the original measurements. A slight decrease in amplitude was observed due to the increased convective inlet losses.

For the highest mode which could be reached $n = 5$ we observe in some experiments a stable coexistence with mode $n = 4$. This contradicts the assumption that the non-linearity, needed to stabilize whistling, precludes the coexistence of self-sustained oscillations of two acoustic modes (non-harmonically related frequencies).

The model ignores the mechanical deformation of the Hummer. The reported flattening of the pitch [10], due to pipe elongation upon increase of pipe rotation speed, is not observed in our experiments.

Flow induces oscillation of the free pipe termination cannot be excluded and should be considered in a more accurate model of the Mechanical Dragon. Traveling bending waves induced by periodic impulses of the hand should be considered in the modeling of the hand driven Hummer.

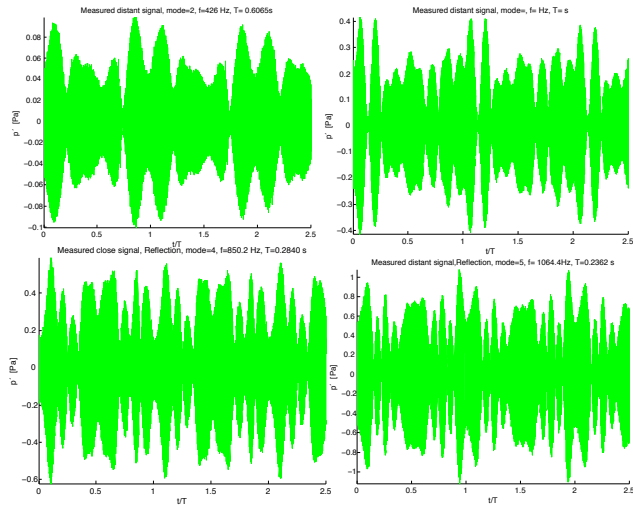


Figure 9. Distant microphone signals p' for inverted pipe: $n = 2$ (upper left), $n = 3$ (upper right), $n = 4$ (lower left) and $n = 5$ (lower right) for reflecting floor as functions of t/T .

- [8] L.L. Beranek, Acoustics, McGraw-Hill, NY (1954) 103-104.
- [9] U.R. Kristiansen, P.-O. Mattei, C. Pinhede, M. Amieth, Experimental study of the influence of low frequency flow modulation on the whistling behavior of a corrugated pipe, *The Journal of the Acoustical Society of America* 130 (2011) 1851-1855. J. Acoust. Soc. Am. 130, 1851 (2011)
- [10] W.M. Hartman, Acoustics for everyone *The Newsletter of the Acoustical Society of America: Echoes* 22, Number 2 (2012) 4-5.
- [11] A. P. Dowling and J. F. Williams, Sound and Sources of Sound, Ellis Horwood, Ltd., West Sussex, England, (1983) 187-190.
- [12] P. McCord Morse, K.U. Ingard, Theoretical Acoustics, MacGraw-Hill, NY (1968).
- [13] W.M. Hartman, private communication (2013).

Acknowledgments

This project has been financed by the Dutch Technology Foundation STW (project 08126). Freek van Uittert provided the data acquisition system. The students from Fontys: G. Diogaardi, I. de Groot, E. Manders, M. Mies and D. Selderehuis.

7. REFERENCES

- [1] G. Nakiboglu, O. Rudenko, A. Hirschberg, Aeroacoustics of the swinging corrugated tube: Voice of the Dragon, *The Journal of the Acoustical Society of America* 131 (2012) 749-765.
- [2] F. S. Crawford, Singing corrugated pipes, *American Journal of Physics* 62 (1974) 278-288.
- [3] M. Silverman, G. Cushman, Voice of the dragon: The rotating corrugated resonator, *European Journal of Physics* 10 (1989) 298-304.
- [4] S. Serafin, J. Kojs, Computer models and compositional applications of plastic corrugated tubes, *Organ. Sound* 10 (2005) 67-73.
- [5] N.H. Fletcher, Th. Rossing, The Physics of Musical Instruments, Springer, NY (1998).
- [6] J. C. Bruggeman, A. Hirschberg, M. E. H. van Dongen, A. P. J. Wijnands, J. Gorter, Self-sustained aeroacoustic pulsations in gas transport systems: experimental study of the influence of closed side branches, *Journal of Sound and Vibration*, 150 (1991) 371-393.
- [7] D. Tonon, A. Hirschberg, J. Golliard, S. Ziada, Aeroacoustics of pipe systems with closed branches, *International Journal of Aeroacoustics* 10 (2011) 201-276.

GRAPH-BASED MODELS FOR WOODWINDS

G. Le Vey

Ecole des Mines de Nantes/IRCCyN UMR-CNRS 6597, FRANCE

levey@emn.fr

ABSTRACT

A model for woodwinds with tone and register holes is presented. It is inspired by the original idea of A.H. Benade considering the set of toneholes as a sequence of 'T-shaped' sections. This idea can be deepened thanks to more recent works on mathematical modelling and analysis of repetitive structures such as networks of strings, beams, membranes, pipes or canals. An essential feature of the model is that it keeps at the one dimensional level. The purpose of this work is to build upon the idea of Benade, inside a precise mathematical framework using concepts and methods from graph theory, for modelling the bore of woodwinds together with its holes in order to address questions like length corrections due to the lattice of (closed or open) toneholes or toneholes interactions. The long term objective is to use this type of model for simulation, characterization of the natural frequencies of woodwinds and, in a control theoretical setting, for musical acoustics questions such as design as well as for theoretical purposes. Applications of the approach are exposed as a program for future research.

1. INTRODUCTION

Wind instruments have been the subject of numerous studies for a long time and are now rather well understood [1–4]. A woodwind can be considered in the first approximation as a duct, the length of which can be made variable by opening or closing one or several toneholes, thereby adjusting the pitch of the instrument. But the acoustical behavior of woodwinds is known to be strongly influenced by the design of its system of tone holes, their sizes and spacings, while these last two parameters cannot be designed independently for a tube of a given thickness [5]. One first straightforward way of studying and simulating these instruments is to write the linear wave equation with convenient initial, boundary conditions on a domain that nevertheless can be somewhat complex. Nevertheless, this direct approach may not give sufficient enlightening to phenomena that interest the musician, the instrument maker or the acoustical physicist and especially how to design a woodwind for musical purposes. Therefore, A.H. Benade proposed to look at a woodwind as a sequence of what he called 'T-shaped sections', each consisting of a

piece of the main bore and one tonehole, with given radii and lengths. In order to obtain qualitative results, a simplified version was used with identical such T-shaped sections, as a kind of periodic medium. Several effects were studied such as those due to the closed-holes or open-holes length corrections, to fork fingerings, or to the function of register hole for higher register functioning ([2] and [3], chap. 7 for a survey). The main objectives are to obtain information about the playing frequencies and the influence toneholes characteristics have on them. These characteristics essentially are of a geometric nature (size, spacing along the main duct). This approach has been the main basis of most works since (see the recent [6–8] e.g.). On another hand, the effects of discontinuities in acoustical ducts (either musical or not) have been investigated in [9] for the case of junctions of two, three and four guides, using modal decomposition and conformal maps. A second aspect in the study of wind instruments (brasses or woodwinds) concerns the bore cross-section. Although it is admitted that the only musically useful bores are members of the Bessel horn family (including cylindrical and conical bores), practical instruments show that the bore is not precisely cylindrical nor conical and that small variations from these idealized shapes arise, e.g. from deliberate alterations brought by the instrument maker in order to improve the tone or the tuning of an instrument [5]. This, together with the presence of toneholes or register holes, affects the natural frequencies of the whole instrument [6, 8]. Thus it is interesting to have means to study this question and the related design problem in a precise way.

For all these questions, the commonly used way of studying woodwinds is through a modal approach, using the electric-acoustic analogy with equivalent circuits and their impedances within the transmission lines formalism, while approximating a duct as a sequence of cylinders or cones; this is quite natural as musical instruments work usually in harmonic regimes. Mode matching is then used to make coherent the different modal decompositions.

In the present work, one takes for modelling a route different from the above-mentioned. Hopefully it will allow to give another light to the above questions and answer some other questions that still remain open. At the present stage, the results are of a theoretical nature but look promising. The original idea of A.H. Benade [5] was to consider a woodwind as a lattice or network of several tubes connected together with junctions. The present work is an attempt to build upon this idea in a mathematically precise way, with one main difference: it keeps a (one dimensional) PDE formalism instead of turning to the discrete, lumped-parameter, transmission lines framework. To this

end, the main ingredient comes from mathematical studies of equations on networks [10–13] and more recently on assemblages of several similar components such as strings, beams, plates, membranes (see the monograph [14] and references therein), pipes or canals [15] and, more specifically connected to the present work, vibrating systems. The basic idea is to consider the graph of this network connecting together the elementary components and to study for example the spectrum of this set through properties of the graph itself and of the components. Hence, the *skeleton* of a given woodwind can remain the same, whereas the model for individual components can be changed according to what effects (e.g. linear vs nonlinear) are to be studied. This way, one remains in the one dimensional setting although complex geometries are in order. Together with this modelling approach, we adopt a control theoretical point of view, considering the geometric parameters such as the duct and toneholes cross-sections or spacings between toneholes, that are indeed design parameters, as control parameters that have to be optimized in some sense that will be detailed in future work. A first step in that direction was done in [16] where the duct cross-section of general wind instruments without toneholes was considered as a control parameter for musical design purposes. This allowed to look at bore design as an optimal control problem. In a similar fashion, the cross-section of tone holes and their spacing can be considered as control parameters for design. We restrict in this work to linear acoustics, although the approach can be extended to a non-linear context.

The presentation is organized as follows : in a first step, we recall the linear model for a unique duct, that can have a non uniform cross-section. In a second step (section 3), a model of a woodwind as a simple network of elementary components is presented. Then, in section 4, we show how to compute the natural frequencies in that context, going into more details for the special case of a one tonehole instrument. Last, applications and future works such as design problems are presented and briefly discussed within a control theoretical setting.

2. LINEAR MODEL IN ONE DUCT

For the sake of self-containedness, we recall here the well-known one dimensional linear model (without sources) used for studying propagation in ducts, while making appear what we consider here as *control parameters*. The fluid is assumed to be barotropic i.e. the pressure is a function of the density ρ only : $p = p(\rho)$ and in the usual conditions of linear acoustics dealing with small perturbations of variables about their mean values. In the sequel, ρ_0 is the density of the gas at rest, c the velocity of sound in free space, $p(x, t)$ the pressure, $v(x, t)$ the volume velocity, $A(x)$ the cross-section area of a tube at abscissa x . The fluid is as usual assumed irrotational, i.e. there exists a velocity potential ϕ : $v = \partial_x \phi$. A model for the plane wave lossless propagation inside a one dimensional non uniform

acoustic wave guide is the following [3] :

$$\begin{cases} \partial_t v + \frac{A}{\rho_0} \partial_x p &= 0 \\ \partial_t p + \frac{\rho_0 c^2}{A} \partial_x v &= 0 \end{cases} \quad (1)$$

For models with losses see [17] or [3], chap. 5. It can be reduced to the well-known horn equation for the pressure : $\frac{1}{c^2} \partial_{tt} p - \frac{1}{A} \partial_x (A \partial_x p) = 0$ which is also valid for the velocity potential. When the cross-section is constant, i.e. for cylindrical ducts, it reduces to the wave equation : $\frac{1}{c^2} \partial_{tt} p - \partial_{xx} p = 0$. Later, the cross-section area A , as a design parameter, will be considered as a *control parameter* in the same spirit as in [16].

3. GRAPH BASED MODELS OF WOODWINDS

For modelling woodwinds within graph theory, consider the *skeleton* of a woodwind as in figure 1 : a scheme of a main tube (the duct of the woodwind) with several other tubes (the toneholes or the register hole) joining it at different locations is constituted of edges of a graph that meet at vertices (or nodes). One elementary situation with one

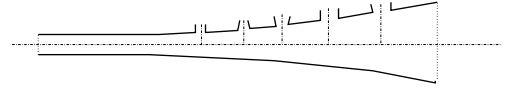


Figure 1. scheme of a woodwind with its graph

main duct and one tonehole appears in figure 2. As one can

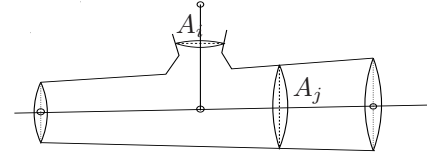


Figure 2. scheme of a woodwind with one tonehole

see on figure 2, each section is not 'T-shaped' in the terminology of A.H. Benade : in such a section, two tubes with different, non uniform cross-sections join together. This is in contrast with usual models that are cylindrical. One hopes to have this way a finer description of what happens e.g. when 'undercutting' is done for toneholes, which amounts to having non cylindrical toneholes.

3.1 Graph description

For the graph description, one follows closely [12] (see [18] e.g. for graph theory). We consider (see figure 3) that each portion of the main duct between two adjacent toneholes is modelled in a schematic way by an edge with two ends modelled by two vertices. Each tonehole or register hole is modelled schematically as an edge joining two vertices. This is relevant for one dimensional models considered here. Then the union of all these edges and vertices constitutes a graph. Observe first that the graph so associated to a woodwind is of a very special type : it is a *tree* as it is connected and contains no cycle [18]. The

resulting underlying tree of a general woodwind is illustrated in figure 3 : for a wind instrument with n holes (tone and register holes), the associated tree has $N = 2n + 2$ vertices (or nodes), denoted V_i , and $N - 1 = 2n + 1$ edges, denoted E_i . Each edge and its associated quanti-

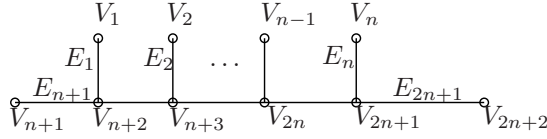


Figure 3. Graph of a woodwind with toneholes

ties are indexed by an integer : $i \in \mathcal{I} = \{1, \dots, N - 1\}$. Therefore, one defines for each edge, the length l_i , the running variable $x_i \in [0, l_i]$, the cross-section area A_i , the pressure p_i , the particle velocity v_i and the velocity potential ϕ_i ($v_i = \partial_{x_i} \phi_i$), $i \in \mathcal{I}$. One can also assume, for the sake of generality, that the sound velocity, c_i , is different in each tube, although this is likely not to be the general case in woodwinds. The locations of end points of each tube, i.e. the vertices of the tree, are labelled by $j \in \mathcal{J} = \{1, \dots, N\}$. Looking at figure 3, one sees that simple vertices belong to the boundary, $\partial \mathring{\mathbf{G}}$ of the graph and that multiple vertices belong to its interior $\mathring{\mathbf{G}}$ ($\mathbf{G} = \mathring{\mathbf{G}} \cup \partial \mathring{\mathbf{G}}$). Hence, one distinguishes multiple vertices, indexed by $j \in \mathcal{J}_M = \{n+2, n+3, \dots, 2n, 2n+1\}$, where several tubes meet, from simple vertices, indexed by $j \in \mathcal{J}_S = \{1, 2, \dots, n, n+1, 2n+2\}$, which are the external ends of the tubes. Notice that for all woodwinds, exactly three edges obviously meet at one multiple vertex. An *ocarina* could be modelled as several tubes that join at one vertex, the cavity, thus as one multiple vertex with more than three joining edges but in such an instrument, the flow cannot be considered one dimensional. For $j \in \mathcal{J}$, it is useful to define :

$$\mathcal{I}_j = \{i \in \mathcal{I} : \text{the } i\text{th tube meets the } j\text{th vertex}\}$$

For $i \in \mathcal{I}_j$, one sets $x_{ij} = 0$ or l_i corresponding to which end meets the other tubes at the j th vertex. One also sets $\epsilon_{ij} = 1$ if $x_{ij} = l_i$ or $\epsilon_{ij} = -1$ if $x_{ij} = 0$, useful for the description of the set of natural frequencies in section 4.

3.2 A woodwind linear model

A linear model for a complete woodwind can be obtained deductively through a variational approach. As this is standard, we just give the results with no demonstration. First, one has a horn wave equation in each tube :

$$\frac{1}{c_i^2} A_i \partial_{tt} \phi_i - \partial_{x_i} (A_i \partial_{x_i} \phi_i) = 0, i \in \mathcal{I} \quad (2)$$

and a Kirchhoff-type flow conservation law at each multiple node :

$$\sum_{i \in \mathcal{I}_j} \epsilon_{ij} A_i(x_{ij}, t) \partial_{x_i} \phi_i(x_{ij}, t) = 0, \forall j \in \mathcal{J} \quad (3)$$

that reduces to a Neumann condition at a simple node, e.g. :

$$\partial_{x_i} \phi_i(x_{ij}, t) = 0 \quad (4)$$

corresponding to a closed hole. Imposed boundary conditions at the external simple nodes, corresponding to open holes radiation conditions, or to the excitation mechanism, can be added through the introduction of a suitable work functional into $\mathcal{L}(\phi)$. All the above represents a lossless wind instrument without active components.

4. NATURAL FREQUENCIES OF A WOODWIND

Determining the natural frequencies of a wind instrument - i.e., in mathematical terms, the spectrum of the differential operator defined in (2) - is central from the viewpoints of physics, of instrument making and is important for musical practice. One straightforward way for their computation is through a finite element approximation of the continuous underlying system, with possibly complicated boundaries, followed by solving the associated generalized eigenvalue problem with a convenient numerical algorithm. But this may be tricky and not much informative about the influence of geometric parameters such as hole sizes and inter-hole spacings. Thus approximating the natural frequencies by corrections from those of idealized bore shapes such as cylinders and cones has been for a long time the favoured approach in the musical acoustics community and several formulae have been given for that purpose [19, 20]. The main reason for this is that no real instrument has an exact cylindrical or conical bore shape [19] whereas for simple duct shapes, exact formulae are known. Thus it is worthwhile to investigate the normal modes of tubes which depart from these exact shapes [19, 21]. The effects of holes also greatly affects the natural frequencies of an instrument [6, 8]. In that context, the graph modelling approach presented above can be an interesting method for computing the natural frequencies, as it consists in one dimensional equations thus of quite lower complexity than full 3D models, while connecting it to geometric parameters of interest to instrument makers, such as tone hole dimensions and inter hole spacings.

In [12], it is shown that the structure of the underlying graph of an elliptic operator on a network plays a distinctive role in the spectrum of the associated eigenvalue problem. This is done using an equivalent boundary value problem for a matrix differential equation. This idea is at the basis of the present section. Notice the important point that mode matching is automatically satisfied with this approach. One essential aspect of the method is to rescale the spatial variables associated with each element of the structure to the uniform interval $[0, 1]$, while accounting for an orientation on each element. Then, a special matrix calculus - on an element by element basis - due to J. Hadamard, allows to pose and solve the corresponding eigenvalue problem leading to the searched spectrum.

First, using equations (2), (3) in section 3, the set of natural frequencies of a woodwind model is the solution of the following eigenvalue problem :

$$\begin{cases} \phi_i \in C^2([0, l_i]), \forall i \in \mathcal{I}, \phi = (\phi_i) \\ \phi \text{ is continuous on } \mathbf{G} \\ \partial_{x_i x_i} \phi_i + \frac{\partial_{x_i} A_i}{A_i} \partial_{x_i} \phi_i = -\frac{\lambda^2}{c_i^2} \phi_i, \forall i \in \mathcal{I} \\ \sum_{i \in \mathcal{I}_j} \epsilon_{ij} A_i(x_{ij}, t) \partial_{x_i} \phi_i(x_{ij}, t) = 0, \forall j \in \mathcal{J} \end{cases} \quad (5)$$

It can be formulated in a synthetic fashion as follows. Consider the tree (see figure 3), $\mathbf{G} \subset \mathbb{R}^2$, of a woodwind, with its set of N vertices, $V(\mathbf{G}) := \{V_i, i = 1, \dots, N\}$, and its set of $N - 1$ edges, $E(\mathbf{G}) := \{E_j, j = 1, \dots, N - 1\}$. The edges are parameterized by $\pi_j : [0, l_j] \rightarrow \mathbb{R}^2$, where the running variable $x_j \in [0, l_j]$ represents the arc length. The maps π_j are assumed to be C^2 -smooth. We introduce the incidence matrix $\mathcal{D} = (d_{ij})$:

$$d_{ij} = \begin{cases} 1 & \text{if } \pi_j(l_j) = V_i \\ -1 & \text{if } \pi_j(0) = V_i \\ 0 & \text{elsewhere} \end{cases} \quad (6)$$

which is the matrix version of the ϵ_{ij} 's of the previous section, and the adjacency matrix $\mathcal{E} = (e_{ih})$:

$$e_{ih} = \begin{cases} 1 & \text{if there exists } s = s(i, h) \in \mathcal{I}, \\ & \text{with } k_s \cap V(\mathbf{G}) = \{V_i, V_h\}, \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

which describes how edges connect vertices. Whenever $e_{ih} = 0$, set $s(i, h) = 1$. Now we recall the Hadamard operations for $n \times n$ matrices $P = (p_{ij})$: whereas the ordinary matrix product will be denoted by simple concatenation (AB) , the product $P.Q$ (with a dot) is done element by element, $(P.Q)_{ij} = p_{ij}q_{ij}$ and for any function $f : \mathbb{R} \rightarrow \mathbb{R}$, the matrix $f(P)$ is given by $f(P) = (q_{ih})$ with:

$$q_{ih} = \begin{cases} f(p_{ih}) & \text{if } e_{ih} = 1 \\ 0 & \text{if } e_{ih} = 0 \end{cases} \quad (8)$$

and especially when $f(x) = x^r$, $r \in \mathbb{R}$ for the matrix powers P^r . Define the vector $e = (1, \dots, 1)^T$ and, for any n -vector, $y = (y_i)$ the diagonal matrix $\text{diag}(y) = (\delta_{ij}y_i)$ with δ the Kronecker delta function. Define also the matrices:

$$\begin{aligned} \mathcal{A} = (A_{ih}) &= (A_{s(i,h)}e_{ih}) \\ \mathcal{C} = (c_{ih}) &= (c_{s(i,h)}e_{ih}) \\ \mathcal{L} = (l_{ih}) &= (l_{s(i,h)}e_{ih}) \end{aligned} \quad (9)$$

and for $\phi : \mathbf{G} \rightarrow \mathbb{R}$ and $x \in [0, 1]$, $\Phi(x) = (\phi_{ih}(x))$ with:

$$\phi_{ih}(x) = e_{ih}\phi_{s(i,h)} \left[l_{ih} \left(\frac{1 + d_{is(i,h)}}{2} - x d_{is(i,h)} \right) \right]$$

such that

$$\Phi(0) = \left(\phi_{ih} \left(\pi_{s(i,h)}^{-1}(E_i) \right) \right) = (\phi_i(x_{ij}, t)) e^T \mathcal{E} = \psi e^T \mathcal{E}$$

$\psi = (\phi_i(x_{ij}, t))$ denoting the vector of values of ϕ_i 's at the vertices. Notice the symmetry $\phi_{hi}(x) = \phi_{ih}(1 - x)$, $x \in [0, 1]$. Last, as the independent variables x_i have all been rescaled to $[0, 1]$, we denote the spatial derivatives with primes ($u' = \partial_x u$) in the rest of this section, to conform with the usual notation. With this set of notations, the eigenvalue problem (5) is equivalent to the following:

$$\begin{cases} \phi_{ih} \in C^2([0, 1]) \text{ and } (e_{ih} = 0 \Rightarrow \phi_{ih} = 0) \forall i, h \in \mathcal{I} \\ \mathcal{L}^{-2} \mathcal{C}^2 \Phi''(x) + \mathcal{L}^{-1} \mathcal{C}^2 \mathcal{A}^{-1} \mathcal{A}' \Phi'(x) = -\lambda^2 \Phi(x), \\ \forall x \in [0, 1] \\ \exists \psi \in \mathbb{R}^N : \Phi(0) = \psi e^T \mathcal{E} \\ \left[\mathcal{L}^{-1} \mathcal{A} \mathcal{C} \Phi'(0) \right] e = 0 \\ \Phi^T(x) = \Phi(1 - x), \forall x \in [0, 1] \end{cases} \quad (10)$$

The solutions of this problem furnish the natural frequencies of the modelled woodwind. A detailed analysis within this generality is deferred to future work. Instead, let us look here at an illustrative and important particular case, when all the c_i 's are equal to the same constant c , which is the most frequent assumption. Assume also that all ducts are cylindrical, which implies: $\mathcal{A}' = 0$. The corresponding eigenvalue problem reduces to:

$$\begin{cases} \phi_{ih} \in C^2([0, 1]) \text{ and } (e_{ih} = 0 \Rightarrow \phi_{ih} = 0) \forall i, h \in \mathcal{I} \\ \mathcal{L}^{-2} \Phi''(x) = -\frac{\lambda^2}{c^2} \Phi(x), \forall x \in [0, 1] \\ \exists \psi \in \mathbb{R}^N : \Phi(0) = \psi e^T \mathcal{E} \\ \left[\mathcal{L}^{-1} \mathcal{A} \Phi'(0) \right] e = 0 \\ \Phi^T(x) = \Phi(1 - x), \forall x \in [0, 1] \end{cases} \quad (11)$$

Using the Hadamard calculus above, the solution of this problem can be given explicitly as:

$$\Phi(x) = \cos\left(\frac{x\lambda}{c}\mathcal{L}\right) \cdot \Phi(0) + \frac{c}{\lambda} \mathcal{L}^{-1} \cdot \sin\left(\frac{x\lambda}{c}\mathcal{L}\right) \cdot \Phi'(0) \quad (12)$$

Thus one has an explicit expression of the eigenvector $\Phi(x)$ corresponding to an eigenvalue λ . This is very important as it gives the solution of a somewhat complex eigenvalue problem in a comprehensible form that moreover can be related to existing results for simpler systems as a simple cylinder, for the sake of comparisons for example. From this, the detailed structure of the set of natural frequencies and eigenvectors can be 'read into' the structure of the underlying tree, through the structure of the matrices: mode matching is automatically satisfied. Omitting the demonstrations, set $\mathcal{B} = \frac{1}{c}\mathcal{L}$ and define the matrix:

$$\mathcal{M}(\lambda) = \mathcal{A} \cdot (\sin(\lambda\mathcal{B}))^{-1} - \text{diag} \left(\left[\mathcal{A} \cdot (\sin(\lambda\mathcal{B}))^{-1} \cdot \cos(\lambda\mathcal{B}) \right] e \right)$$

with given physical and geometric parameters $c, \mathcal{L}, \mathcal{A}, \mathcal{E}$. Thanks to the expression of $\Phi(x)$ and to the boundary conditions in (11), the eigenvalues of problem (11) can be shown to be of one of the following two types:

1. $\lambda = l_{s(i,h)}^{-1} c \pi k$ for some $i, h = 1 \dots, N$; $k \in \mathbb{Z}$ which are the eigenvalues of elementary ducts.
2. λ is a solution of the transcendental equation:

$$|\mathcal{M}(\lambda)| = \det \mathcal{M}(\lambda) = 0 \quad (13)$$

Therefore the complete set of natural frequencies of the woodwind model is explicit. These results parallel those for a parabolic problem on networks in [12], extended in [14] for networks of hyperbolic mechanical systems made of strings and beams.

4.1 A simple example : the one tonehole case

As a most simple illustration of the general method, we present the (non realistic) case of a duct closed at both ends with one closed tonehole ($n = 1$), both cylindrical, that has been thoroughly studied. The matrices for the graph are:

$$\mathcal{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathcal{E} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (14)$$

for the orientation of the three edges for which the origin is at the only multiple vertex, according to the scheme in figure 3. The sound velocity is assumed to be c in all ducts. The matrices \mathcal{A} , \mathcal{L} carry the same structure as \mathcal{E} by definition. For readability, define the intermediate quantities : $c_i = \cos(\frac{\lambda_i}{c})$, $s_i = \sin(\frac{\lambda_i}{c})$, $t_i = \tan(\frac{\lambda_i}{c})$, $S_c = \sum_{i=1}^3 c_i$, $S_t = \sum_{i=1}^3 A_i t_i$. Then the matrix $\mathcal{M}(\lambda)$ writes after some computations :

$$\mathcal{M}(\lambda) = \begin{pmatrix} -A_1 \frac{S_c}{s_1} & 0 & \frac{A_1}{s_1} & 0 \\ 0 & -A_2 \frac{S_c}{s_2} & \frac{A_2}{s_2} & 0 \\ \frac{A_1}{s_1} & \frac{A_2}{s_2} & -S_t & \frac{A_3}{s_3} \\ 0 & 0 & \frac{A_3}{s_3} & -A_3 \frac{S_c}{s_3} \end{pmatrix} \quad (15)$$

It is symmetric and its determinant is :

$$|\mathcal{M}(\lambda)| = -S_c^2 \prod_{i=1}^3 \frac{A_i}{s_i^2} \sum_{\sigma_c} [A_{\sigma(1)} s_{\sigma(2)} s_{\sigma(3)} (1 + c_{\sigma(1)} S_c)]$$

where σ_c are the three circular permutations of the set $\{1, 2, 3\}$. The natural frequencies are then the solutions of $|\mathcal{M}(\lambda)| = 0$. Further simplifying the problem, assume now that all lengths of individual edges of the graph are equal to a given length l , which implies that $\mathcal{L} = lI$, with I the identity matrix. Then the natural frequencies can be determined very simply as the solution of an algebraic eigenvalue problem. First, we search for matrices $\Phi(0)$ and $\Phi'(0)$ (indeed the function ψ) that satisfy the conditions of the third and fourth line in (11). After some computations, using the symmetry condition in (11), one gets :

$$\phi'_{ih}(0) = \frac{\frac{\lambda}{c}}{\sin(\frac{\lambda}{c})} \left(e\psi^T - \cos(\frac{\lambda}{c})\psi e^T \right) \cdot \mathcal{E}$$

and this allows to write the general solution as a solution of ψ by simple substitution :

$$\phi_{ih} = \cos(\frac{\lambda x}{c}) \mathcal{E} \cdot \psi e^T + \frac{\sin(\frac{\lambda x}{c})}{\sin(\frac{\lambda l}{c})} \left(e\psi^T - \cos(\frac{\lambda l}{c})\psi e^T \right) \cdot \mathcal{E}$$

The same substitution into the fourth line of (11) gives :

$$(\mathcal{A} \cdot [e\psi^T - \cos((\lambda l)/c)\psi e^T] \cdot \mathcal{E}) e = 0$$

which can be rewritten as :

$$\mathcal{A}\psi = \cos((\lambda l)/c) \text{diag}(\mathcal{A}e)\psi$$

Defining now the matrix $Z := (\text{diag}(\mathcal{A}e))^{-1} \mathcal{A}$, the previous equation translates to an eigenvalue problem for ψ :

$$Z\psi = \mu\psi$$

where : $\lambda = c \arccos(\mu)/l$. Using the results from [12] and the fact that the graph of a wind instrument, as defined above, is *bipartite* [18], we get the following characterization of the eigenvalues, i.e. the natural frequencies of the instrument model. From [12], Z has eigenvalues $\mu_i : i = 1, \dots, N$ satisfying : $1 = \mu_1 \geq \mu_2 \geq \dots \geq \mu_N = -1$, with corresponding eigenvectors : $\xi_1 = e, \xi_2, \dots, \xi_N$. Label the eigenvalues as λ_{sk} , where s denotes the edge number and $k \in \mathbb{N}$ the index of eigenvalues for that edge and ϕ_{sk} is the corresponding eigenelement and let α be a generic real number. Then one has :

- $\lambda_{sk} = -\frac{c}{l} \arccos(\mu_s) + (k+1)\pi$ if $k \equiv 1 \pmod{2}$, $2 \leq s \leq N$ and $-1 < \mu_s, k \in \mathbb{N}, \phi_{sk} = \alpha \xi_s$
- $\lambda_{sk} = -\frac{c}{l} \arccos(\mu_s) + k\pi$ if $k \equiv 0 \pmod{2}$, $2 \leq s \leq N$ and $-1 < \mu_s, k \in \mathbb{N}, \phi_{sk} = \alpha \xi_s$

5. DISCUSSION

The model developed in section 3 accounts in a simple way for important geometric parameters : toneholes spacings are given by the length of internal edges $E_k, k = n+2, \dots, 2n+1$ and height of toneholes are given by the length of external edges $E_k, k = 1, \dots, n$ (see figure 3), all gathered in the matrix \mathcal{L} . The diameter of these last ones is explicitly given in the dynamic equations ((2), (3), (5) or (10)). The main duct diameter is given by the sequence of diameters corresponding to the internal edges. All diameters are elements of matrix \mathcal{A} . Thus the set of natural frequencies of a woodwind can be computed and varied as a function of all these parameters. It is not clear at the moment how to describe the geometry at the junctions but notice that mode matching is automatically satisfied. This model makes it possible to study anew usual questions in musical acoustics related to the natural frequencies of woodwinds such as : quantify length corrections due to the closed holes or open holes lattices ; study the case of one main duct with one tonehole, as in [6] ; analyze the effect of different cross-fingerings on the playing frequency ; quantify the experimental observation that the tuning properties of a woodwind are predominantly affected by the properties of only the first two or three open toneholes [5]. Due to the one dimensional nature of the model, the complexity of computing the natural frequencies or of a simulation with the presented model is low : for an instrument with 8 toneholes, the corresponding model is made of 18 equations for the dynamics and as much unknowns. These points and comparisons with the usual transmission lines approach will be investigated in future works.

5.1 Woodwinds design as a control problem

As it has been shown in [16], focussing on bore shape design, control theory can be a useful framework for design problems in musical acoustics. The model that has been presented in section 3 allows to pursue this line of investigation by including in the design process important geometric parameters such as toneholes dimensions and spacings between them. These parameters, i.e. the matrices \mathcal{A} and \mathcal{L} , can be considered as control parameters for a series of inverse problems. Optimal control theory can then be used as in [16] when a suitable optimization criterion is defined. A typical one for woodwinds is the precise alignment of fundamental frequencies for the first and second registers. One important point is that such design problems make appear controls that are *distributed* in the space variable. On the contrary, previous works on network models of pipes or canals, based on the nonlinear St Venant equations, focussed on boundary control in the time variable [15, 22], because the geometry of canals was there given and fixed. Hence the above model for wood-

winds can be useful for numerous initial-boundary control problems for purposes of design and simulation.

6. CONCLUSION AND FUTURE RESEARCH

Modelling a woodwind using graph-theoretical concepts opens new possibilities to treat questions such as length corrections and can be useful for simulation. The computational complexity of the model is relatively low. Nevertheless it is likely that it cannot compete with the transmission lines approach for real-time sound synthesis. But we think that for off-line analysis and for design it can be helpful. Also, it allows to treat design questions as control problems hence can be useful as a tool for instrument making as well as for better insight into the physics of the instruments. Future research will focus on the relationship between the above effects and the geometric parameters, in the same spirit as in [5] and recent research [8]. In that respect, the matrix formulation of the present work fits well a perturbation analysis, useful for studying the influence of geometry on the natural frequencies through the matrix M . Also, the excitation mechanism and related questions have to and will be accounted for in this model. Notice eventually that the proposed approach is applicable to other acoustical studies not related to musical acoustics.

The previous developments have been limited to linear models of elementary ducts because the natural frequencies is an utmost important characteristic of a wind instrument. But it is known that several nonlinear effects appear too in playing situations. The graph-based approach can be extended straightforwardly to the nonlinear situation, at the price of a greater complexity, by considering the nonlinear equations in an elementary duct, together with the same tree *skeleton*. This is currently under investigation.

7. REFERENCES

- [1] N. Fletcher and T. Rossing, *The physics of musical instruments*. Springer, 1991.
- [2] A. H. Benade, *Fundamentals of musical acoustics*. Dover, 1990, 2d edition.
- [3] A. Chaigne and J. Kergomard, *Acoustique des instruments de musique*. Belin, 2008.
- [4] H. Bouasse, *Instruments à vent, I, II*. Delagrave, 1929, new edition Librairie A. Blanchard, 1986.
- [5] A. Benade, "On the mathematical theory of woodwind finger hole," *J. Acoust. Soc. Amer.*, vol. 32, pp. 1591–1608, 1960.
- [6] V. Dubos, J. Kergomard, A. Khettabi, J. P. Dalmont, D. H. Keefe, and C. J. Nederveen, "Theory of sound propagation in a duct with a branched tube using modal decomposition," *Acta Acustica*, vol. 85, pp. 153–169, 1999.
- [7] E. Moers and J. Kergomard, "On the cutoff frequency of clarinet-like instruments. geometrical versus acoustical regularity," *Acta Acustica u. Acustica*, vol. 97, no. 6, pp. 984–996, 2011.
- [8] A. Lefebvre, G. P. Scavone, and J. Kergomard, "External toneholes interaction in woodwind instruments," 2012, preprint, <http://arXiv.org/pdf/1207.5490> or <http://hal.archives-ouvertes.fr/hal-00720132>.
- [9] J. Kergomard, A. Khettabi, and A. Garcia, "Formulation générale des jonctions de guides d'ondes aux basses fréquences," *C.R. Acad. Sci. Paris*, vol. 319, no. II, pp. 887–892, 1994.
- [10] W. H. Wittrick and F. W. Williams, "Natural frequencies of repetitive structures," *Quart. J. Mech. and Applied Math.*, vol. XXIV, no. 3, pp. 285–308, 1971.
- [11] S. Nicaise, "Some results on spectral theory over networks applied to nerve impulse transmission," in *Lecture notes in mathematics*, C. Brezinski, Ed., 1985, pp. 533–541.
- [12] J. Von Below, "A characteristic equation associated to an eigenvalue problem on c^2 networks," *Linear algebra and its applications*, vol. 71, pp. 309–325, 1985.
- [13] —, "Classical solvability of linear equations on networks," *J. Diff. Eqns*, vol. 72, pp. 316–337, 1988.
- [14] J. L. Lagnese, G. Leugering, and E. J. P. G. Schmidt, *Modeling, analysis and control of dynamic elastic multilink structures*. Birkhäuser, 1994.
- [15] G. Leugering and E. J. P. G. Schmidt, "On the modelling and stabilization of flows in networks of open canals," *SIAM J. Control Optim.*, vol. 41, no. 1, pp. 164–180, 2002.
- [16] G. Le Vey, "Optimal control theory : a method for the design of wind instruments," *Acta Acustica u. Acustica*, vol. 96, pp. 722–732, 2010.
- [17] D. H. Keefe, "Acoustical wave propagation in cylindrical ducts : transmission line parameter approximations for isothermal and nonisothermal boundary conditions," *J. Acoust. Soc. Am.*, vol. 75, no. 1, pp. 58–62, 1984.
- [18] J. W. Wilson, *Introduction to graph theory*. Academic Press, 1972.
- [19] A. H. Benade, "On woodwind instruments bores," *J. Acoust. Soc. Amer.*, vol. 31, no. 2, pp. 137–146, 1959.
- [20] A. H. Benade and E. V. Jansson, "On plane and spherical waves in horns with non uniform flare," *Acustica*, vol. 31, no. 2, pp. 79–98, 1974.
- [21] H. Bouasse, *Tuyaux et résonateurs*. Delagrave, 1929, new edition Librairie A. Blanchard, 1986.
- [22] J. de Halleux, C. Prieur, J. M. Coron, B. d'Andréa Novel, and G. Bastin, "Boundary feedback control in networks of open channels," *Automatica*, vol. 39, pp. 1365–1376, 2003.

VOCAL TRACT EFFECTS ON THE TIMBRE OF THE SAXOPHONE

Weicong Li

Jer-Ming Chen

John Smith

Joe Wolfe

School of Physics, The University of New South Wales, Sydney
weicong.li@student.unsw.edu.au jerming@unsw.edu.au john.smith@unsw.edu.au
J.Wolfe@unsw.edu.au

ABSTRACT

For notes sounded over the normal and altissimo playing range, experienced saxophonists can produce changes in the spectral envelope of the radiated sound by adjusting their vocal tract configuration. Measurements of the vocal tract acoustic impedance, Z_{mouth} , during performance showed that, when Z_{mouth} was comparable with the input impedance of the bore, Z_{bore} , *i.e.* several $\text{MPa}\cdot\text{s}\cdot\text{m}^{-3}$ or more, harmonics of the radiated sound falling near these peaks in Z_{mouth} were substantially enhanced. In contrast, the broadband noise in the radiated sound produced by upstream turbulence was attenuated in the frequency range over which the magnitude of Z_{mouth} was large.

1. INTRODUCTION

The shape of the vocal tract during woodwind performance is widely regarded by woodwind instrument performers to be of great importance to both pitch selection and control and to the quality of the sound. Recent advances in understanding this relationship involve direct measurements of the acoustic properties of the vocal tract during performance: the acoustic impedance of the vocal tract [1,2] and measurements of the ratio of the sound pressure inside the mouth and to that in the mouthpiece [3].

Expert saxophonists and clarinetists have been shown to use their vocal tracts for executing advanced performance techniques that involve controlling the fundamental frequency: performing in the altissimo register, bugging, multiphonics, pitch bending and glissando [2-5]. In a simple model, Benade [6] showed that the bore (downstream) and the vocal tract (upstream) impedances (Z_{bore} and Z_{mouth}), as ‘seen’ by the reed, are effectively in series. For these techniques, players tune a vocal tract resonance near to the desired frequency and this, in series with the bore, determines the playing pitch [2,4,5].

Woodwind players also report varying their vocal tract configuration to produce different timbre in different musical styles and contexts. These vocal tract configurations are often described by players in terms of various vowel-like mouth shapes (*e.g.* ‘oo’, ‘ee’, ‘ah’). Presumably, varying the upstream geometry changes Z_{mouth} and thus influences the amplitude of higher harmonics in the produced sound.

In the previous SMAC (2003), we used artificial playing systems to show how changes in the upstream geometry, analogous to a changing tongue position, can produce different spectral envelopes (as well as different pitches) in trombone and didjeridu playing [7]. We later showed that for didjeridu playing, formants (or broad peaks in the envelope) of the radiated sound coincided with minima in the acoustic impedance measured in the mouth, and that the broad minima in the spectral envelope corresponded with impedance peaks in the mouth [8].

How might changes in vocal tract shape influence the amplitude of a harmonic of the output sound? One possibility is that a peak in Z_{mouth} at a particular frequency could vary the Fourier component of the pressure acting on the reed at that frequency and thus affect the shape of its vibration. Another possibility is that a peak in Z_{mouth} inhibits flow past the reed at that frequency, which explains the timbre changes of the didjeridu. A third possibility is that the variation in the upstream geometry changes the pitch of the note played. A small change in fundamental frequency has an n times greater effect on the n th harmonic, which may be enough to shift it from coinciding with a bore resonance to not coinciding, or *vice versa*. Whether or not a harmonic coincides with a bore resonance has a large effect on the radiated sound because of the production of standing waves [8].

To investigate this effect, the acoustic impedance in the mouth of experienced saxophonists was measured while they played notes with different timbre.

2. MATERIALS AND METHODS

For this study, a Yamaha Custom EX Tenor Saxophone with Yamaha 5C mouthpiece is used. The mouthpiece is fitted with a Légère synthetic saxophone reed (hardness 3), chosen for its stability, hygiene and stable physical properties. The input impedance data of the tenor saxophone bore used in this study comes from a database [9] obtained using the three-microphone-two-calibration method with non-resonant calibration loads [10].

To measure directly the acoustic impedance of the player’s vocal tract during playing, an acoustic impedance measurement head based on the capillary method was modified from Chen *et al.* [2]. The measurement head (Figure 1) consists of a narrow stainless steel tube with internal cross sectional area of 2 mm^2 integrated into the saxophone mouthpiece which supplies the acoustic current source with harmonics from 500 to 4000 Hz at a spacing of 5.38 Hz. This is injected into the player’s

mouth during playing. Next, an Endevco 8507C-2 miniature pressure transducer of 2.42 mm diameter similarly fitted in the mouthpiece, adjacent to the stainless steel tube, is used to measure the sound pressure in the player's mouth, which includes both that produced by the vibrating reed and the response of the vocal tract to the injected probe current. These modifications increase the thickness of the mouthpiece by 2 mm at the bite point. However, this is not regarded as a significant disturbance by players [5], some of whom use different geometry mouthpieces for different styles of music.

The impedance measurement system is calibrated by connecting the modified mouthpiece to an acoustically infinite pipe (length 197 m, internal diameter 26.2 mm). To make a measurement, the player is asked to sustain a note for at least 6 seconds, while the broadband probe signal is injected into the player's mouth and its response recorded.

Three expert saxophonists, each with more than 10 years' classical and/or jazz background, participated in the study. Using the modified mouthpiece described above, the players were asked to achieve different timbre by only adjusting the tongue position while keeping other control parameters constant (*e.g.* biting force on the reed, embouchure, pitch and loudness).

Another microphone (Rode NT3) was positioned one bell radius from and on the axis of the bell of the saxophone to record the radiated sound for spectral analysis.

The raw acoustic impedance spectra were then analysed and treated [4] to remove the harmonics generated by the vibrating reed and to smooth the airflow turbulence measured inside the mouth.

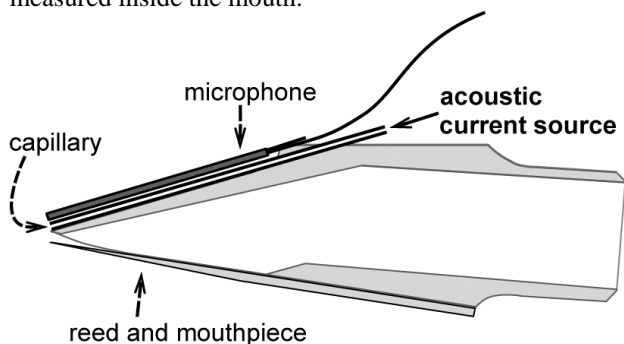


Figure 1. A schematic cross-section of the mouthpiece shows the capillary method used in the measurement of the vocal tract impedance during playing.

3. RESULTS AND DISCUSSION

All our subjects were able to produce different timbre by changing their tongue position, for notes across the normal and altissimo playing range, without much difficulty.

One general observation applies to all players: variation of the sound spectrum inside the mouth was much greater than that of the radiated sound. The player can hear the sound inside the mouth via transmission from mouth to ear through the bones and tissues, which explains why timbre changes may sometimes seem significant to the player, but rather less so to listeners.

Further, the players were usually able to produce larger timbre changes when they also changed the pitch, typically by ten or so cents. An obvious explanation is that large changes in the vocal tract configuration produce a change in the acoustic load on the reed that not only changes the harmonic content, but also the frequency of vibration. When constrained to keep constant pitch, the players were probably limited to a smaller range of vocal tract changes, for which the pitch could be compensated using other control parameters, such as the biting force on the reed.

Nevertheless, all the players in this study could produce significant changes in the spectral envelope of the radiated sound while maintaining a constant pitch. Here, we restrict discussion to the changes at constant pitch only.

The execution of these changes seems to vary from one player to another, and slight adjustments of the tongue can result in subtle timbre variations. This highlights the difficulty of comparing across different players, so in this study, we compare only the timbre variation with different tongue positions (vocal tract configuration) for the same player. In particular, two of our subjects report that the 'ah' tongue position is their default position during normal performance, especially for jazz playing. While the 'ee' tongue position is unusual, some players use it to create a subtle timbre variation. One subject noted that the 'ah' tongue position provides brightness while the 'ee' tongue position gives a 'dark and nasal' sound.

Figure 2 shows the sound spectra of one note (written C5 on tenor saxophone, sounding A#3, 233 Hz) and the measured vocal tract impedance spectra of one subject playing that note with different timbre by adjusting tongue position, (a) for the normal 'ah' tongue position and (b) for the 'ee' tongue position. The spectral envelopes in Figure 2 (a) and (b) clearly show the timbre variation. The amplitude of the third and fourth harmonics in the spectrum of the 'ee' tongue position is much greater: about 10 dB larger than those of the 'ah' tongue position. Other harmonics in the frequency range from 1.5 to 3.0 kHz all have about 5 dB difference than those of the 'ah' tongue position. The sound spectrum of the 'ee' tongue position also shows broad peaks (formants) at about 1.8 and 2.3 kHz, whereas that of the 'ah' tongue position, shows formants at about 1.6 and 2.5 kHz. How do these modifications in the spectral envelope relate to the resonances in the vocal tract?

In the Z_{mouth} spectrum for the 'ee' tongue position, the second and third impedance maxima shift to about 0.9 and 1.8 kHz, and their magnitudes are increased to about 8 and 3 MPa·s·m⁻³, respectively, which is comparable with the magnitude of Z_{bore} . Figure 2(b) shows that, over the frequency range where Z_{mouth} exceeds Z_{bore} , the amplitude of harmonics here are substantially increased above their amplitude for the normal 'ah' tongue position. In other words, the biggest increase of the amplitude of the harmonics seems to correspond to the maxima in the vocal tract impedance of the 'ee' tongue position, when the magnitude of Z_{mouth} is sufficiently large. This is consistent with an observation by Scavone [3]: for one of the players in that study, the fifth and eighth harmonics in the radiated sound increased when the ratio of the pressure in the mouth to that in the mouthpiece increased.

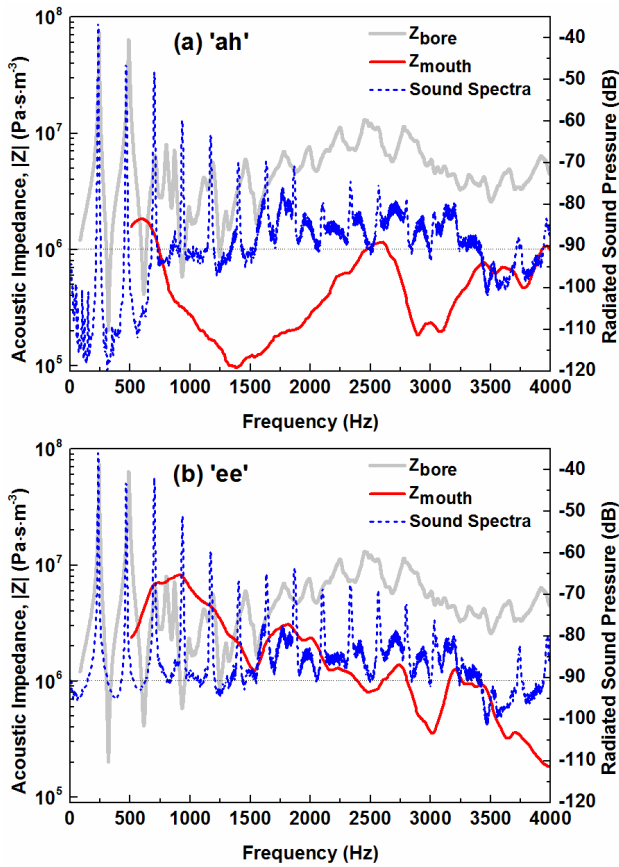


Figure 2. The sound spectra of the note (dashed blue line, A#3, nominally 233 Hz) and the measured vocal tract impedance spectra (continuous red line, measured from 0.5 to 4.0 kHz). Subject A plays the same note with different timbre by adjusting tongue position: (a) ‘ah’-like and (b) ‘ee’-like vowel. In both (a) and (b), the pale grey line shows the impedance spectra of the bore.

When we compare the broadband noise in the sound spectra with the impedance spectra, it is found that the background noise exhibits the complementary behaviour, *i.e.* the minima of the background noise correspond to the maxima of the Z_{mouth} and *vice versa*. This can be explained in the following way: the maxima of Z_{mouth} give minima in the acoustic flow at the reed, so there is little power input into the mouthpiece over this range of frequencies and thus little power in the radiated sound spectra. Similar behaviour was also observed in didjeridu playing [8], but for the harmonics of the played note.

Figure 3 shows the frequencies of the harmonics having the largest increase in amplitude when compared with the normal sound, plotted against the frequency of the peaks in Z_{mouth} for the cases where Z_{mouth} is comparable with Z_{bore} . From the figure, we can see that the correlation between the enhanced amplitude of the harmonics of the sound and peaks in Z_{mouth} occurs over the entire frequency range measured.

From Figure 3, we conclude that strong peaks in Z_{mouth} enhance the power of harmonics falling in or near the range of those peaks. The correspondence is not exact: for high notes, the harmonics are widely spaced, so few may fall near a peak.

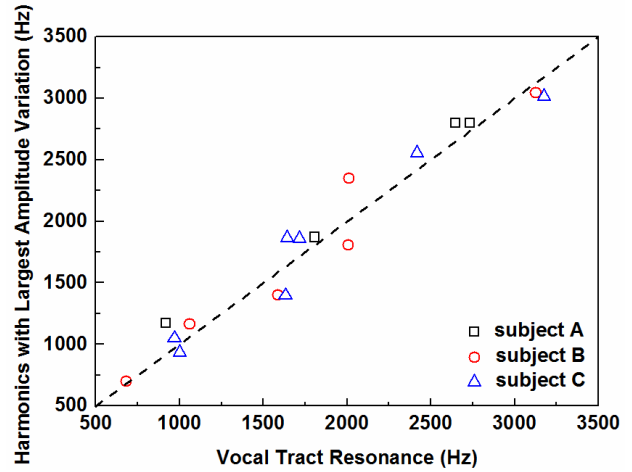


Figure 3. The harmonics with largest increase in amplitude in the radiated sound spectra against the nearest vocal tract resonance (maxima). Different sets of symbol stand for the results of different subjects.

In the case of didjeridu playing [8], the impedance maxima of the vocal tract correspond closely to minima in the spectral envelope of the radiated sound, while maxima in the radiated sound spectrum are correlated with the minima in the impedance spectra. Thus, the vocal tract configuration has a distinct effect on the timbre variation of the didjeridu. For the trombone [7,11], different vocal tract configurations (high tongue and low tongue) also produce timbre variations, but the effect is much less dramatic than during didjeridu playing because the Z_{bore} of the trombone is much higher than that of the didjeridu and consequently changes in Z_{mouth} have a reduced effect. The Z_{bore} of the saxophone also has a larger magnitude than that of the didjeridu [12], so the timbre change in Western wind instruments is substantially less than that in the didjeridu, but still of considerable musical importance.

4. CONCLUSIONS

When a player alters their vocal tract configuration, the changes in the spectral envelope of the sound radiated from the saxophone are rather smaller than those of the sound inside the player's mouth. Nevertheless, experienced players can vary the spectral envelope substantially at constant pitch and loudness. When the amplitude of the peaks in Z_{mouth} is at least comparable with those of Z_{bore} , the harmonics of the radiated sound are enhanced in the radiated sound. In contrast, the broadband noise is suppressed within this range.

Acknowledgments

We thank the Australian Research Council for support and our volunteer subjects.

5. REFERENCES

- [1] A. Tarnopolsky, N. Fletcher, L. Hollenberg, B. Lange, J. Smith, and J. Wolfe, “The vocal tract and

- the sound of a didgeridoo,” *Nature*, 2005, vol. 436, p39.
- [2] J.-M. Chen, J. Smith, and J. Wolfe, “Experienced saxophonists learn to tune their vocal tracts,” *Science*, 2008, vol. 319, p726.
 - [3] G.P. Scavone, A. Lefebvre, and A.R. Da Silva, “Measurement of vocal-tract influence during saxophone performance,” *J. Acoust. Soc. Am.*, 2008, vol. 123, pp. 2391–2400.
 - [4] J.-M. Chen, J. Smith, and J. Wolfe, “Pitch bending and glissandi on the clarinet: roles of the vocal tract and partial tone hole closure,” *J. Acoust. Soc. Am.*, 2009, vol. 126, pp. 1511–1520.
 - [5] J.-M. Chen, J. Smith, and J. Wolfe, “Saxophonists tune vocal tract resonances in advanced performance techniques,” *J. Acoust. Soc. Am.*, 2011, vol. 129, pp. 415–426.
 - [6] A. H. Benade, “Chapter 35: Air column, reed, and player’s windway interaction in musical instruments,” in *Vocal Fold Physiology, Biomechanics, Acoustics, and Phonatory Control*, ed. I.R. Titze and R.C. Scherer (Denver Center for the Performing Arts, Denver, CO), 1985, Chap. 35, pp. 425–452.
 - [7] J. Wolfe, A.Z. Tarnopolsky, N.H. Fletcher, L.C.L. Hollenberg, and J. Smith, “Some effects of the player’s vocal tract and tongue on wind instrument sound,” in *Proc. Stockholm Music Acoust. Conf. (SMAC 03)*, Stockholm, 2003, pp. 307–310.
 - [8] A. Tarnopolsky, N. Fletcher, L. Hollenberg, B. Lange, J. Smith, and J. Wolfe, “Vocal tract resonances and the sound of the Australian didgeridu (yidaki) I: Experiment,” *J. Acoust. Soc. Am.*, 2006, vol. 119, pp. 1194–1204.
 - [9] J.-M. Chen, J. Smith, and J. Wolfe, “Saxophone acoustics: introducing a compendium of impedance and sound spectra,” *Acoust. Aust.*, 2009, vol. 37, pp. 18–23.
 - [10] P. Dickens, J. Smith, and J. Wolfe, “High precision measurements of acoustic impedance spectra using resonance-free calibration loads and controlled error distribution,” *J. Acoust. Soc. Am.*, 2007, vol. 121, pp. 1471–1481.
 - [11] V. Fréour, and G.P. Scavone, “Vocal-tract Influence in trombone performance,” in *Int. Symp. Music Acoust.*, Sydney and Katoomba, 2010.
 - [12] J.-M. Chen, J. Smith, and J. Wolfe, “How players use their vocal tracts in advanced clarinet and saxophone performance,” in *Int. Symp. Music Acoust.*, Sydney & Katoomba, 2010.

Embracing the Digital in Instrument Making: Towards a Musician-tailored Mouthpiece by 3D Printing

Valerio Lorenzoni

L'Chaim

valeriolorenzoni@gmail.com

Zjenja Doubrovski

TU Delft

e.l.doubrovski@tudelft.nl

Jouke Verlinden

TU Delft

J.C.Verlinden@tudelft.nl

ABSTRACT

At present, the manufacturing of musical instruments still strongly relies on the tacit knowledge of experienced handcrafts while is commonly based on standard machining or casting techniques. This limits the musician-tailoredness to a small group of players, while others take compromises by employing stock parts.

The present article describes a new methodology for the design and production of woodwind instruments mouthpieces. By embracing digital modeling and manufacturing, this methodology encompasses four phases, which can be cut short when necessary. The aim of the presented methodology is to link the geometry of the mouthpiece to tone properties. Based on 3D printing, the inside geometry can be altered to complex and reproducible detail to obtain the desired acoustic features - eventually leading to mouthpiece geometries tailored to the player's sound and playability requirements.

The results of aerodynamic investigations together with the subjective experience of saxophone players have been used to design mouthpieces with modified inside geometries of both baffle and chamber. Prototypes have been produced at the Delft University of Technology (TU Delft) using several 3D printing technologies and different materials. Both professional and amateur saxophone players tested these at the Royal Conservatoire of The Hague and at the North Sea Jazz festival (Rotterdam 2012). Based on the judgment of the players, specific geometrical features were revealed to emphasize specific tone characteristics. A number of professional players are actively performing with our mouthpieces.

The application and further development of the methodology will lead to a better empirical basis to reason about acoustics and playability, and can be applied to other instruments as well. Future work includes additional measurements and developing a parameterized database of 3D models.

1. INTRODUCTION

It is widely recognized among wind instrument players that the inside geometry of the mouthpiece has a strong influ-

ence on the response, timbre, and intonation of a woodwind instrument.

Mouthpiece manufacturing nowadays still often relies on the knowledge of experienced handcrafts and is commonly based on lathe work or casting techniques. In the last decades several mouthpiece manufacturers have adopted computer-aided design systems and computer numerical control (CNC) machines for the production of saxophone and trumpet mouthpieces. The use of 3D printing however, compared to the other manufacturing techniques offers advantages in terms of reducing production costs and allowing the construction of inside geometries that are hardly achievable by machining.

Our methodology, presented in this paper, focuses on modifications of the mouthpiece inside geometry, which influence the flow structures and turbulence level inside it. The aim is to improve the acoustic properties of a mouthpiece according to specific sound requirements, in a controlled, measured and reproducible way, by exploiting the capabilities of 3D printing.

2. BACKGROUND

2.1 Mouthpiece performance

Due to the complexity of the mouthpiece functioning principle, it is hard to directly relate its internal geometry to the sound quality of the coupled mouthpiece-instrument system.

First attempts to identify such a connection dealt with variations of the chamber volume and shape. The work of Benade [1] revealed that a short and open mouthpiece chamber is associated with a dark tone quality, while a long and narrow chamber is associated with a relatively brighter tone quality. A variation of cavity size effects the lowest resonances in equal proportions and does not alter their ratios. At higher frequencies, however, the mouthpiece length correction becomes frequency dependent and can have important consequence on tuning and tone color.

Wynman [2] performed acoustic measurements on five different geometrical-types of alto saxophone mouthpiece models. He found that changes in the acoustic spectrum due to mouthpiece chamber modifications depend on the dynamic level of playing. Wynman also stressed that the amount of brightness in the tone is primarily controlled by the baffle shape and that a small baffle-to-reed angle tends to promote a bright tone.

An interesting overview of mouthpiece investigations can be found in the PHD dissertation of Scavone [3], which de-

scribes the modeling of single-reed wind instruments based on acoustic principles in the digital waveguides domain. Scavone also stresses the influence of the bore shape (conical and cylindrical) on the reed behavior, which strongly affects the harmonic distribution of the sound spectrum.

The study of Hasbrook [4] provides a large database of measurements on different saxophone and mouthpiece combinations. The analysis primarily focuses on mouthpiece pitch and compares a traditional "classical" sound versus a traditional "jazz" sound. An observation was that the tip opening usually increases from classical to jazz mouthpieces, as the length of the lays.

The effect of mouthpiece chamber modifications on the sound produced by saxophone has been the main focus of the established mouthpiece manufacturer François Louis. Some of the design concepts are provided by François Louis himself in a web seminar masterclass [5].

A similar investigation on the effect of the mouthpiece depth modification on the sound of the trumpet was carried out by Poirson *et al.* in [6]. It was found that increasing the depth of the mouthpiece chamber tends to produce a more "dark" sound as revealed by both measurements with an artificial blowing device and measurements on real players. They concluded that the influence of the musician on the produced sound is comparable to the variations of the mouthpiece geometries and therefore it is important to use an artificial blowing device to determine the qualities of a wind instrument for design and validation of the modeling tools.

2.2 3D printing

3D Printing is a collection of production technologies also known as Additive Manufacturing (AM). Although there are differences between different AM technologies, they all fabricate physical objects directly from a 3D computer file by adding material layer upon layer. Originally these technologies were used for prototyping, but improvements in speed, price and material properties have caused an ever increasing application of 3D printing for the production of end-products.

Compared to traditional manufacturing, such as machining and (injection) moulding, AM offers unprecedented freedom in shape complexity and custom geometry. Ref. [7] provides a literature overview of proposed methods to utilize this geometrical complexity. Combined with a high reproducibility and low costs for custom fabrication, 3D printing has a large potential for the production of customized musical instruments, as emerging from recent applications on a fully 3D printed flute (<http://youtu.be/zwHg szH0aqI>) and a fully 3D printed violin (<http://www.wired.co.uk/news/archive/2011>).

In earlier work we have explored the use of four different 3D printing technologies for the production of saxophone mouthpieces [8]. The technologies included: PolyJet, Fused Deposition Modeling, Selective Laser Sintering and Selective Laser Melting. Both the PolyJet and Fused Deposition Modeling machines are available in-house at the TU Delft. For the Selective Laser Sintering and Selective Laser Melting we have used the services of AM service

bureaus.

Figure 1 shows some of the prototype models produced at TU Delft using the above mentioned techniques. The materials used for the shown mouthpieces are (from left to right): ABS, Titanium, Objet VeroBlue plastic, Polyamide.



Figure 1: Preliminary mouthpiece models produced using different 3D printing technologies and materials.

The mouthpieces used for the tests described in Section 4 were produced using the PolyJet technology. Based on the ASTM categorization for AM technologies <http://www.astm.org/COMMITTEE/F42.htm>, the PolyJet technology belongs to the category Material Jetting. While moving over two axis (X and Y) an inkjet head deposits a layer of a liquid UV curable resin. After the deposition of the layer, a UV lamp cures the resin into a solid polymer. Once a layer is complete, the build tray lowers (Z axis) and the jetting of a new layer is initiated. This process is repeated until the entire model is created. A schematic illustration of the working principle of a PolyJet machine is shown in Figure 2. We found that the PolyJet technology, pro-

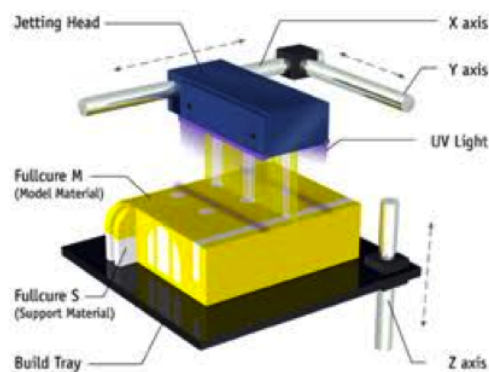


Figure 2: PolyJet 3D printer machine.

duced by Objet® (now Stratasys, <http://objet.com>) is suitable for producing high quality mouthpieces that do not require post processing.

3. METHODOLOGY FOR MOUTHPIECE DESIGN

The proposed methodology consists of four phases and it is sketched in Figure 3

In the first phase mouthpieces with specific geometrical features are produced using 3D printing. The geometric

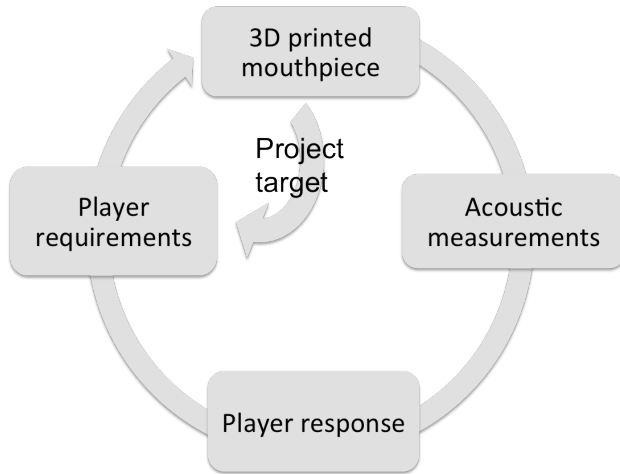


Figure 3: Sketch of the methodology for mouthpiece design.

characteristics of the mouthpiece are varied according to specific acoustic targets. The inputs can derive from either previous experiments or innovative design ideas. Some of the saxophone mouthpiece designs produced at TU Delft in the first tests, were based on the results of the aerodynamic measurements described in Section 4.1.

As second step acoustic tests are carried out on the mouthpieces in an anechoic environment by means of an artificial blowing device (artificial mouth) and microphones. This step is needed to objectively and quantitatively determine the spectral characteristics of the designed mouthpieces, without the influence of the player. The objective characterization of the spectra compared to the musician response would shed light on the timbre characteristics of each design. In general, it is difficult to discuss musical tone quality without using subjective terms such as: "dark", "mellow", "focused", "bright", "compact", etc. These tests would elucidate some of the above definitions often used by musicians to describe an instrument timbre, from a more scientific point of view. Aerodynamic visualization or computational fluid-dynamics simulations could also be performed in this phase to investigate the flow features associated to each geometry.

Once the spectra of the different mouthpieces have been objectively analyzed, the next crucial phase of the method consists of tests with real players. These would enable to link the measured acoustic spectra to the player-based characteristics.

An extensive investigation on different saxophone models and mouthpieces have been performed by Talley [9]. The scope of the latter work was to identify the tone fingerprints of different musicians, which is the distinctive tone quality of every player. The analysis of Hasbrook [4] demonstrates how the internal voicing of a player can dramatically alter the tone quality, independently of mouthpiece choice. From the above studies it seems not possible, in a first stage, to define the quality of a mouthpiece without considering its effect on the player.

Performing the same tests on several musician would allow to statistically link the measured spectra of each mouth-

piece to the response on the player. Characteristics found by several players would be considered to belong to the specific design. The ease-of-play could also be inferred, which is not directly detected from measurements on artificial blowing devices. Furthermore such analysis will indicate the sensitivity of the musician to modifications of the mouthpiece geometry.

By knowing what a geometric modification would cause on the spectral characteristics of the sound and also knowing the response of musicians to it, one could systematically alter the mouthpiece shape in a "controlled" and "reproducible" way, using the 3D printing technology, directly obtaining timbre characteristics that satisfy the player's requirements. This represents the main target of this project.

4. FIRST APPLICATION

A first application of the proposed methodology is described in this Section.

4.1 Preliminary aerodynamic measurements

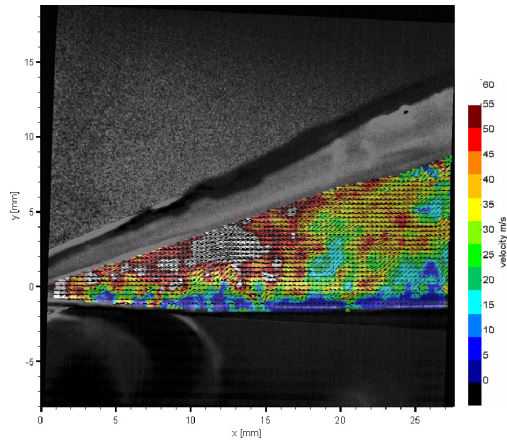
Aerodynamic tests on a saxophone mouthpiece have been performed at the aerospace faculty of TU Delft using the particle image velocimetry (PIV) experimental technique. The mouthpiece was connected to a real tenor saxophone and blown by an artificial mouth. Experimental setup and detailed results of the flow field measurements can be found in ref. [10]. Contours of the velocity field inside the mouthpiece baffle are shown in Figure 4.

The instantaneous velocity contours of Figure 4(a) revealed that the flow velocity features maxima of over 60 m/s, mainly localized on the upper lay of the baffle when the reed is approaching closure. Figure 4(b) shows the mean velocity contours calculated over about 300 reed opening/closing cycles. These indicate that the mean velocity across a reed cycle has maxima localized on the upper surface of the baffle at a distance of about 6 reed apertures (at rest) downstream of the tip (12-13 mm in this case) and that about 11 reed apertures downstream of the tip flow recirculation occurs, starting in the vicinity of the reed.

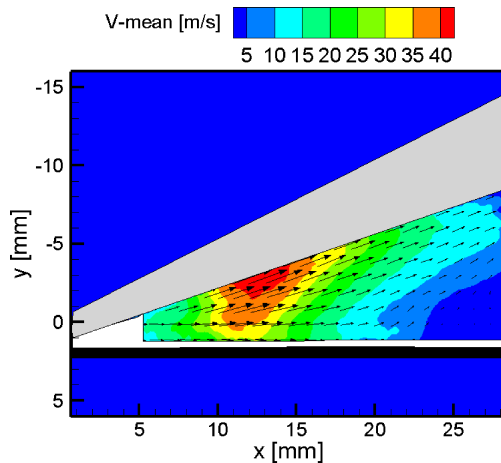
The first part of the baffle roof has been identified as the most influential on the sound production and it was believed that slight modifications of such a region might have a strong impact on the reorganization of the flow structures in the mouthpiece and therefore on the sound properties of the instrument. This seems to be in line with the findings of Wynman [2] and the experience of mouthpiece manufactures and handcrafts.

These aerodynamic results have served as input for geometric modifications of standard mouthpieces and led to the design of eleven different prototypes, three of which are shown in Figure 6 in Section 4.

The use of 3D printing might improve future PIV experiments by producing mouthpiece models with flat external sides. This would increase the optical accessibility through the mouthpiece and overcome some of the experimental issues reported by Lorenzoni and Ragni in [10].



(a) Instantaneous velocity snapshot



(b) Mean velocity

Figure 4: Flow visualizations inside the mouthpiece by particle image velocimetry.

4.2 3D printed mouthpiece

Since the first target user of our mouthpiece design method was David Liebman, one of the mouthpiece models normally used by him was chosen as reference for the geometrical modifications and a 3D scan of it was made at TU Delft using a "Phoenix Nanotom S" CT scanner (<http://www.ge-mcs.com/en/radiography-x-ray/ct-computed-tomography/nanotom-s.html>). This mouthpiece was a "Jazz Chamber" Lebayle[®] soprano mouthpiece of aperture 8.

The scanned model was reconstructed in SolidWorks[®] and modifications have then been made to the baffle and chamber of the original mouthpiece shape, according to the findings described in Section 4.1.

For producing these mouthpieces we used an Objet Eden 260 machine. The machine is able to produce objects up to the size of 600 x 252 x 200 mm at a resolution of 600dpi and layer thickness of 16 micrometer. The material we have used is a biocompatible resin marketed by Objet as MED610. It is a rigid transparent material developed and approved for prolonged contact with human tissue (<http://objet.com/3d-printing-materials/bio-compatible>). Using this machine allowed us to fabricate one mouthpiece in approximately one hour, or a batch of 10 unique mouthpieces in less than 4 hours.

There is no general agreement on the effect of the material on the sound characteristics of a mouthpiece. According to Larry Teal [11] the material by which the mouthpiece is constructed has little influence on the tone quality. This seems to disagree with what reported by David Liebman [12], who instead states that each material offers a unique response and that especially metal mouthpieces have a more brilliant and compact sound. A wide range of modern mouthpieces are produced in vulcanized rubber which offers a good compromise between flexibility and hardness.

3D printing offers a wide range of possible materials and it will be used to further investigate the effect of the material on the sound characteristics.

4.3 Acoustic measurements

The acoustic properties of each mouthpiece geometry still need to be assessed. As also stressed by Poirson *et al.* in [6], the use of an artificial blowing device is an essential requirement for the determination of the objective qualities of a wind instrument.

An artificial mouth was produced by students at TU Delft and is shown in Figure 5.

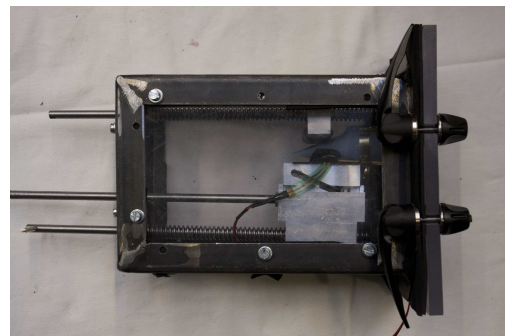


Figure 5: Artificial mouth produced at TU Delft for mouthpiece acoustic tests.

The artificial mouth in Figure 5 has been produced following the indications of the group of Gazengel [13] and features improvements compared to the model used by Lorenzoni and Ragni in [10]. This is a 5 mm thick Plexiglas[®] box reinforced with a steel frame. The side where the mouthpiece is installed is fixed to the rest by quick-release fasteners, which allow to easily interchange the mouthpiece inside the box. A metal support is placed underneath the mouthpiece and artificial lips of silicon-like material are installed on it. The position of the support can be adjusted both horizontally and vertically. An extra support is placed above the mouthpiece to simulate the teeth and to hold the mouthpiece in place during the tests. The pressure inside the box is controlled by a feedback system which allows to keep the level steady during the tests.

Acoustic tests using this device are planned in Delft in September 2013. The test set-up will consist of acoustic radiation measurements by means of microphones at the bell end and acoustic impedance measurements on the mouthpiece alone using an impedance probe.

4.4 Player response

Eleven different 3D printed mouthpiece geometries were produced using the Objet Eden 260 machine and the Objet[®] MED610 material. This material is also used for medical prothesis, it is safe for the user and provides good strength and stiffness for the present purpose.

Liebman evaluated the eleven mouthpieces, during a workshop at the music conservatory in The Hague. The musician gave our team positive feedback and provided us with a first user-based characterization of the mouthpieces from the point of view of an experienced musician. Three of the eleven soprano mouthpieces tested by Liebman are shown in Figure 6

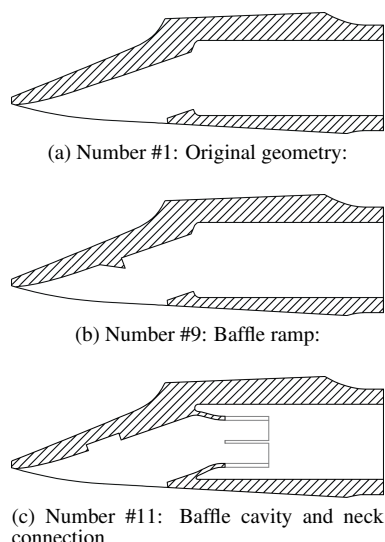


Figure 6: Three of the mouthpiece designs tested by David Liebman at the music conservatory in The Hague.

The comments by Dave Liebman about the three mouthpieces are the following:

Number #1: Original geometry *"As good as the original"*

Number #9: Baffle roof ramp *"By far the best. Better ability for dynamics: full, round with color"*

Number #11: Baffle roof cavity and neck connection *"Mel-low sound, pleasant. Does not have expansion. Not bad but nasal. Not as warm"...*

Number #1 is a close reproduction of the original geometry, the only significant difference is the material used. The inside geometries of the number # 9 and number #11 are based on the aerodynamic results of Section 4.1 and with respect to the number #1 both have a modified baffle roof geometry, which is the region where the flow speed was shown to have a maximum.

In particular number #9 features a small ramp which induces a flow recirculation and aims at increasing the turbulence level in the chamber. The presence of multiple turbulent flow scales enriches the spectral content of the sound. This would explain the "full" sound mentioned by Liebman.

The mouthpiece number #11 has a cavity in the baffle and a connection to the saxophone neck. The function of the connection is to smoothly converge the flow towards the neck. The flow in this case does not expand in the chamber and is more accelerated toward the neck due to the sectional reduction. Less flow recirculation is expected to occur which implies a spectrum more focused on a specific frequency band, generating what Liebman defined as nasal sound.

By eliminating the sectional area discontinuity, the neck connection revealed to noticeably improve the response of the model, increasing the "ability for dynamics", as mentioned by our test musician.

The geometric modifications applied to the mouthpiece revealed to produce the foreseen effects on the musician response. A more objective characterization of the models will derive from future acoustic measurements. The changes in the flow pattern also need to be assessed by targeted aerodynamic experiments or computational fluid dynamics simulations.

The 3D mouthpiece team has also presented the new mouthpiece at the North Sea Jazz Festival (Rotterdam 2012), where musicians have shown particular interest to the production technique and the good acoustic quality of the mouthpieces. The soprano player Jure Pukl (www.jurepukl.com) tried the mouthpieces at our stand at the North Sea Jazz. He was positively impressed by the sound of the number #9 model and he is currently performing with this mouthpiece.

5. CONCLUSIONS AND FUTURE WORK

A methodology has been presented for the design of new single-reed woodwind instruments mouthpiece geometries based on aerodynamic investigations in combination with acoustic measurements and players' evaluations.

Modified mouthpiece inside geometries were designed based on the results of aerodynamic experiments and innovative design ideas and manufactured using additive manufacturing. Musicians who played on the modified mouthpieces noticed the difference with the standard designs and the judgement seemed to be strictly correlated with the expected flow behavior.

3D printing revealed to be a powerful tool for the production of saxophone mouthpieces. The main advantages reside in the high speed of production, accuracy, large variety of usable materials, low costs and capability of producing complicated shapes compared to standard manufacturing techniques. The firsts feedbacks by the musicians revealed that 3D printing is valuable for the production of actual mouthpieces and that the present method has the potential for improving the design of mouthpieces, towards the production of geometries tailored to the sound requirements of the musicians. Patent application has been made by TU Delft for the application of 3D printing to saxophone mouthpiece design.

Future work will consist of acoustic tests on the eleven mouthpieces evaluated by Liebman, using the above described artificial mouth. Tests are planned to start in September 2013. In a second stage numerical and experimental

aerodynamic investigations will also be employed trying to directly relate the acoustic spectrum of the saxophone-mouthpiece system to the flow pattern and turbulence intensity level in mouthpiece baffle and chamber.

Acknowledgments

The authors would like to thank Prof. Gary Scavone for the interest shown in this project and for his help finding important reference material. The whole team is grateful to Dave Liebman for having shown special interest in the project and having taken the time to test our mouthpieces. A special thank also goes to Jure Pukl for being the first professional musician playing on one of our 3D printed mouthpieces. The whole 3D printed mouthpiece team would like to acknowledge the work performed on the artificial mouth by the students at the industrial design faculty at TU Delft and Wim Verwaal at CiTG for the CT scanning.

6. REFERENCES

- [1] A. H. Benade, *Fundamentals of Musical Acoustics*. New York: Oxford University Press, 1976.
- [2] F. Wynman, "An acoustical study of alto saxophone mouthpiece chamber design," *PHD thesis - Eastman School of Music*, 1972.
- [3] G. P. Scavone, "An acoustic analysis of single-reed woodwind instruments with an emphasis on design and performance issues and digital waveguide modeling techniques," *PHD Thesis - Stanford University*, 1997.
- [4] R. V. Hasbrook, "Alto saxophone mouthpiece pitch and its relation to jazz and classical tone qualities," *PHD Thesis - University of Illinois at Urbana-Champaign*, 2005.
- [5] F. Louis, *Choosing your saxophone mouthpiece. François Louis' master class at Mariachi*. <http://vimeo.com/22764049>, 2011.
- [6] E. Poirson, J. Petiot, and J. Gilbert, "Study of the brightness of trumpet tones," *Journal of Acoustic Society of America*, vol. 118, pp. 2656–2666, 2005.
- [7] E. Doubrovski, J. Verlinden, and J. Geraedts, "Optimal design for additive manufacturing: Opportunities and challenges," in *ASME DETC2011*, vol. 4, 2011, pp. 112–124.
- [8] E. Doubrovski, J. Verlinden, J. Geraedts, I. Horvath, and K. V.L.M., "Acoustic investigation of novel saxophone mouthpiece design," in *TMCE*, vol. 1, 2012, pp. 1–8.
- [9] K. M. Talley, "A comparative study of saxophone tone," *PHD thesis - Michigan State University*, 2005.
- [10] V. Lorenzoni and D. Ragni, "Experimental investigation of the flow inside a saxophone mouthpiece by particle image velocimetry," *Journal of the Acoustic Society of America*, vol. 131, pp. 715–721, 2012.
- [11] L. Teal, *The Art of Saxophone playing*. Warner Bros, 1963.
- [12] D. Liebman, *Developing a personal saxophone sound*. Dorn Publications, 1989.
- [13] B. Gazengel and J. F. Petiot, "Objective and subjective characterization of saxophone reeds," *HAL UNIV Internal report*, vol. 47, 2013.

SIMULATIONS OF MODAL ACTIVE CONTROL APPLIED TO THE SELF-SUSTAINED OSCILLATIONS OF THE CLARINET

Thibaut Meurisse, Adrien Mamou-Mani, Rene Causse

IRCAM - 1 Place Igor Stravinsky, 75004 Paris, France

Thibaut.Meurisse@ircam.fr

David Sharp

The Open University - Walton Hall, Milton Keynes, MK7 6AA, UK

David.Sharp@open.ac.uk

ABSTRACT

Modal active control enables modifications of the damping and the frequencies of the different resonances of a system. A self-sustained oscillating wind instrument is modelled as a disturbance coupled to a resonator through a non-linear coupling. The aim of this study is to present simulations of modal active control applied to a modeled simplified self-sustained oscillating wind instrument (e.g. a cylindrical tube coupled to a reed, which is considered to approximate a simplified clarinet), incorporating collocated speaker, microphone and a reed. The next goal will be to apply this control experimentally and to test it with musicians.

1. INTRODUCTION

Modal active control enables modifications of the damping and the frequencies of the resonances of a system [1, 2]. However, there have been relatively few applications for musical instruments [3–5] and no application to wind instruments to the authors' knowledge.

Self-sustained oscillating wind instruments, like the clarinet, are modeled as a disturbance coupled to a resonator through a non-linear coupling [6–9] (see Figure 1). By applying modal active control to a self-sustained simplified clarinet (e.g. a cylindrical tube coupled to a reed), it should be possible to modify its emitted sound and playability.

The aim of this study is to present simulations of modal active control applied to a simplified clarinet.

After presenting the clarinet model and the principles of the modal active control, a coupling between them is proposed. Then, simulations are presented of the control of the frequency and the damping of the first resonance, the control of the damping of the second resonance of a cylindrical tube and finally maps showing the control limits of the first resonance in frequency and damping.

Copyright: ©2013 Thibaut Meurisse, Adrien Mamou-Mani, Rene Causse et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

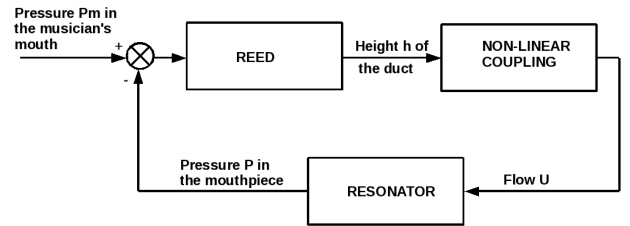


Figure 1. Model of a self-sustained wind instrument [6, 11].

2. MODELING

2.1 Self-Sustained Wind Instrument Model

Models of self-sustained wind instruments like the clarinet have been reported for over 30 years [6–10]. Classically, a self-sustained wind instrument can be described in terms of both linear (reed, resonator) and non-linear (coupling) elements (see Figure 1). In particular, the model used in this paper is the one described by [11].

In a clarinet, a single reed controls the flow of air from the player's mouth into the instrument. Let $h(t)$ be the position of the reed. Then,

$$\frac{1}{\omega_r^2} \frac{d^2 h(t)}{dt^2} + \frac{q_r}{\omega_r} \frac{dh(t)}{dt} + h(t) - h_0 = -\frac{1}{K_r} (P_m - P(t)) \quad (1)$$

where ω_r is the resonance frequency of the reed, q_r its damping, h_0 its equilibrium position and K_r its stiffness, P_m is the pressure in the player's mouth, assumed to be constant, and $P(t)$ the pressure in the mouthpiece.

The pressure in the mouthpiece is obtained through the poles s_n and the residus C_n of the input impedance of the resonator. Let P_n be the pressure of the mode n . Then,

$$\frac{dP_n(t)}{dt} = s_n P_n(t) + Z_c C_n U(t) \quad (2)$$

where $Z_c = \rho_0 c / S$ is the characteristic impedance of the tube with S its cross-sectional area, ρ_0 the density of the

acoustic medium, c the velocity of sound in the medium and $U(t)$ the flow through the reed duct. The pressure is then

$$P(t) = 2 \sum_n \Re(P_n) \quad (3)$$

Using the same hypothesis as described in [11], the flow can be determined from

$$U(t) = \text{sign}(P_m - P(t)) Wh(t) \sqrt{\frac{2|P_m - P(t)|}{\rho_0}} \quad (4)$$

where W is the width of the reed duct.

2.2 Modal Active Control

Modal active control makes it possible to control the damping and the frequency of the eigenmodes of a system. To apply this control, it is necessary to build a model of the system. A state-space model of the system is implemented.

2.2.1 State-Space Model

The state-space model of the acoustic duct used in this paper is inspired by [1] and adapted to the simulation needs, using [2, 5, 12, 13]. The diameter $2R$ of the duct is sufficiently small compared to its length L_t , that the duct can be considered to be a one-dimensional waveguide with spatial coordinate z , where $0 \leq z \leq L_t$. The control speaker is placed at $z = z_s$. The microphone is placed at the same location ($z_m = z_s$).

The pressure in the duct is described by:

$$\frac{1}{c^2} \frac{\partial^2 p(z, t)}{\partial t^2} = \frac{\partial^2 p(z, t)}{\partial x^2} + \rho_0 \frac{dv_s(t)}{dt} \delta(z - z_s) \quad (5)$$

where p is the acoustic pressure and v_s the speaker baffle velocity.

Using separation of variables, let

$$p(z, t) = q(t)V(z) \quad (6)$$

where ([1, 14])

$$V_i(z) = c \sqrt{\frac{2}{L_t}} \cos(k_i z) \quad (7)$$

where $V_i(z)$ is the amplitude of the mode i at position z and $k_i = (2i + 1)\pi/2L_t$. To obtain a state-space description of the acoustic duct, without considering the mode of the speaker, let

$$x(t) = \begin{bmatrix} q \\ \dot{q} \end{bmatrix} \quad (8)$$

where $x(t)$ is the state vector so that

$$\dot{x}(t) = Ax(t) + Bu_s(t) + G\omega(t) \quad (9)$$

$$y(t) = Cx(t) \quad (10)$$

where

$$u_s = \rho_0 \dot{v}_s \quad (11)$$

is the command and $\omega(t)$ a disturbance signal at $z_d = 0$,

$$A = \begin{bmatrix} 0_{r,r} & I_{r,r} \\ -\text{diag}(\omega_i^2) & -\text{diag}(2\xi_i \omega_i) \end{bmatrix} \quad (12)$$

is the *system matrix* [15], I is the identity matrix, ξ_i is the damping of mode i and ω_i its frequency,

$$B = \begin{bmatrix} 0_{r,1} \\ V_i(z_s) \end{bmatrix} \quad (13)$$

is the *actuator matrix*,

$$C = [V_i(z_m) \quad 0_{1,r}] \quad (14)$$

is the *sensor matrix*, and

$$G = \begin{bmatrix} 0_{1,r} \\ V_i(z_d) \end{bmatrix} \quad (15)$$

is the *disturbance matrix*.

2.2.2 Control of the Eigenmodes

The control is carried out using pole placement. The coordinates of the poles are defined by the damping and the angular frequency of each mode [2, 12]:

$$\text{Re}(\text{pole}_i) = \xi_i \omega_i \quad (16)$$

and

$$\text{Im}(\text{pole}_i) = \pm \omega_i \sqrt{1 - \xi_i^2} \quad (17)$$

where ξ_i and ω_i are the damping and the angular frequency of the i^{th} mode.

It is then possible to dictate the damping and the frequency of each mode. Practically, the observer generates a control gain vector K and an observation gain vector L . The control gain vector K is chosen such that

$$\det[sI - (A - BK)] = 0 \quad (18)$$

where $s = j\omega$. In the simulations, K is obtained using the Matlab *place* function [16] with the target poles. The observation gain vector L is used in the observer control loop such that

$$L(y - \hat{y}) \rightarrow 0 \quad (19)$$

where \hat{y} is the observer estimation of y (see figure 2). In the simulations, L is obtained using *place* function with poles which real parts are two times the value of the poles used to obtain K .

The simulations show transfer functions obtained in open loop (without control) and closed loop (with control). Open loop transfer function H_{OL} is [13]:

$$H_{OL} = C(sI - A)^{-1}G \quad (20)$$

Closed loop transfer function H_{CL} is :

$$H_{CL} = C(sI - (A - BK(sI - (A - BK - LC))^{-1}LC))^{-1}G \quad (21)$$

2.3 Active Control on the Self-Sustained Model

2.3.1 Dimensionless Equations

To make it easier to implement, the self-sustained model is made dimensionless. Let

$$\begin{aligned} x_h(t) &= h(t)/h_0 \\ y_h(t) &= h'(t)/v_0 \\ p_n(t) &= P_n(t)/P_M \\ p(t) &= P(t)/P_M \\ u(t) &= U(t)Z_c/P_M \end{aligned} \quad (22)$$

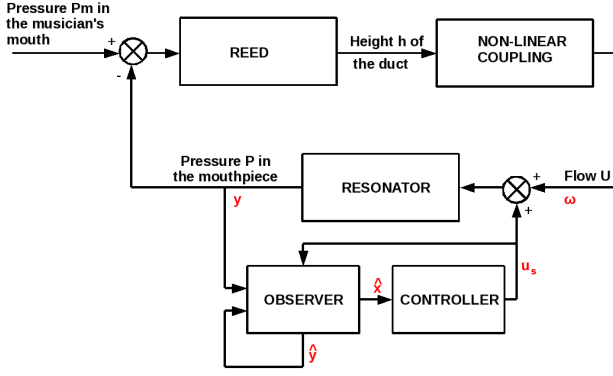


Figure 2. Model of a self-sustained wind instrument with control system. ω , y and u_s are defined in eq.(9) and (10). \hat{y} and \hat{x} are the observer estimations of y and x .

where $v_0 = h_0 \omega_r$ is the reed's speed in free behaviour and $P_M = K_r h_0$ is the pressure required to completely close the reed channel in static regime. The final system is then

$$\begin{aligned} \frac{1}{\omega_r} x'_h(t) &= y_h(t) \\ \frac{1}{\omega_r} y'_h(t) &= 1 - x_h(t) + p(t) - \gamma - q_r y_h(t) \\ p'_n(t) &= C_n u(t) + s_n p_n(t) \\ p(t) &= 2 \sum \Re(p_n(t)) \\ u(t) &= \zeta \text{sign}(\gamma - p(t)) x_h(t) \sqrt{|\gamma - p(t)|} \end{aligned} \quad (23)$$

where $\gamma = P_m/P_M$ represents the pressure in the musician's mouth, $\gamma \simeq 1/3 + \epsilon$ with $\epsilon \ll 1$ [8].

$\zeta = Z_c W \sqrt{\frac{2h_0}{\rho_0 P_M}}$ [9] represents the musician's lips on the mouthpiece.

2.3.2 Adapting the Control

The design of the controller does not allow the control of the resonator described by the poles and residus. It is then necessary to adapt the state space implementation of the controller to the dimensionless self-sustained system. Eq.(23) shows that the input of the resonator is a flow u . Eq.(5) shows that the state-space resonator must have an acceleration as an input, the acceleration of the speaker baffle \dot{v}_s . From (11),

$$u_s = \rho_0 \frac{dv_s}{dt} = \frac{\rho_0}{S} \frac{dU}{dt} \quad (24)$$

where S is the cross-sectional area of the tube. To make the flow U dimensionless, it has been divided by the characteristic impedance of the tube Z_c between eq.(2) and eq.(23). Then,

$$u_s = \frac{\rho_0}{SZ_c} \frac{du}{dt} \quad (25)$$

In *Simulink*, the *discrete derivative* object is used to apply the differentiation.

It is then possible to couple the reed and the state-space resonator *via* the flow (see Figure 2).

Next section presents some results of the simulations.

3. SIMULATIONS

Simulations are carried out using *Matlab* and *Simulink*. The simulations are made with a closed-open tube with

f_r	1300Hz
q_r	1
K_r	$3.3 \times 10^8 \text{N/m}$
h_0	$3 \times 10^{-4} \text{m}$
W	0.0168m

Table 1. Parameters used to characterise the bass clarinet reed.

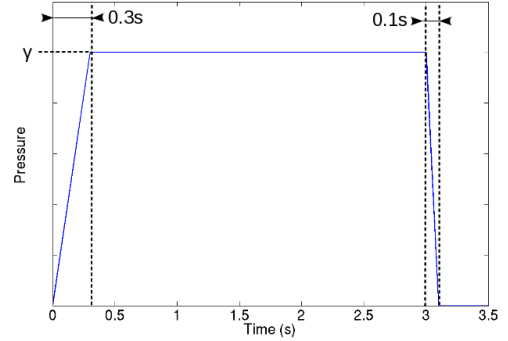


Figure 3. Shape of the dimensionless pressure input used for the simulations.

length $L_t = 1.09\text{m}$ and radius $R = 0.0109\text{m}$. The speaker and microphone are placed at the entrance of the tube ($z_s = z_m = 0$). The modal parameters (ξ_i, ω_i) of a calculated input impedance [17] are extracted thanks to a Rational Fraction Polynomials (RFP) algorithm [18]. The efficiency of the RFP algorithm has already been demonstrated [13, 19]. Only the 10 first resonances are modeled. The pole placement is obtained using the Matlab *place* function. Simulink is used in the *Fixed-step* mode, with the *ode3 (Bogacki-Shampine)* solver. The sample time is 1/44100, and the simulation time is 3s. The considered length is close to the length of a bass clarinet without bell and bocal, the modeled reed is then adapted to a bass clarinet reed. Table 1 shows the values that have been used in the simulation to characterise the reed. Figure 3 shows the shape of the dimensionless pressure input used.

Two cases are first presented. First, the frequency and the damping of the first resonance are controlled, and second, the damping of the second resonance is controlled. Effects of the control on the sound spectrum and the attack transient are studied. Then two maps are presented. The first shows the limits of the control of the frequency of the first resonance. The second shows the limits of the control of the damping of the first resonance.

3.1 Control of the Frequency and the Damping of the First Resonance

The control is applied such that the frequency of the first resonance is changed, from 78Hz to 70Hz, and its damping is increased five times.

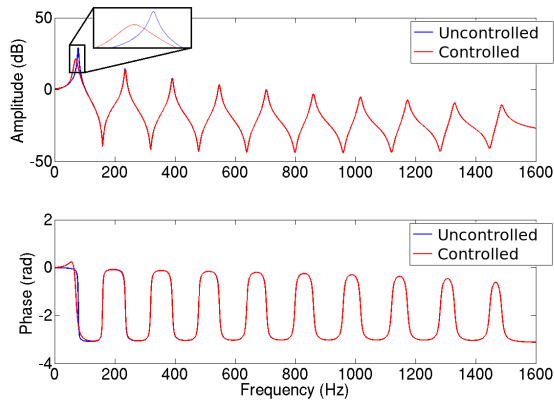


Figure 4. Top : Transfer functions of the uncontrolled (blue) and controlled (red) systems. The control aims to change the frequency of the first resonance from 78Hz to 70Hz and to increase its damping five times. Bottom : Phases of the transfer functions of the uncontrolled (blue) and controlled (red) resonators.

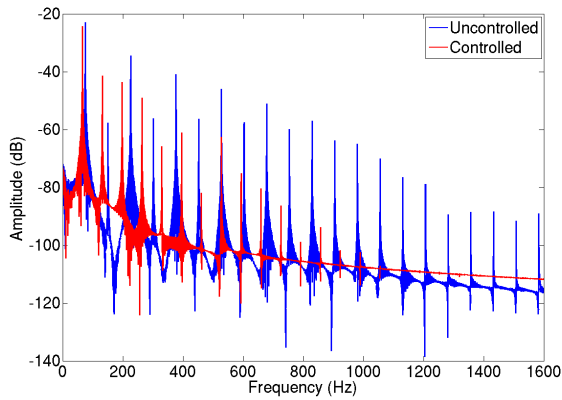


Figure 5. Sound spectra of the uncontrolled (blue) and controlled (red) self-sustained oscillating systems, with $\gamma = 0.3683$.

Figure 4 shows the transfer functions of the uncontrolled and controlled systems. With the control, the frequency of the first resonance is changed to 70.2Hz, and its amplitude is decreased by 7.7dB. The second resonance is also affected, with a decrease of 0.7dB.

Figure 5 shows the sound spectra of the steady states of the uncontrolled and controlled self-sustained systems, with $\gamma = 0.3683$. With the control applied, the simulation still plays on the first resonance of the instrument. Here, it is 66Hz, and 1.3dB lower than without control. As this resonance is no more tuned with the other resonances, there are fewer harmonics compared to the uncontrolled instrument, and no harmonics beyond 1000Hz. The last harmonic at 989Hz is 37dB lower than the closest harmonic of the uncontrolled instrument which is at 980Hz.

Figure 6 shows the attack transients of the uncontrolled and controlled self-sustained systems, with $\gamma = 0.3683$.

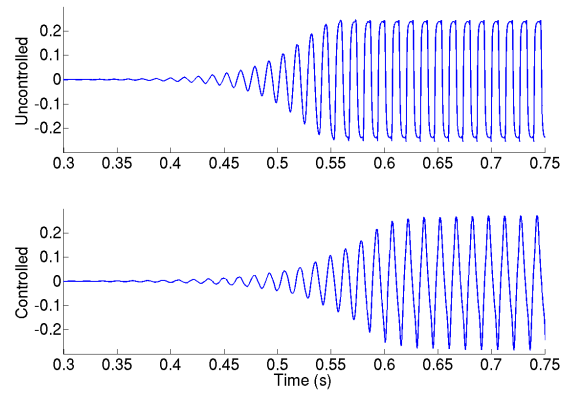


Figure 6. Attack transients of the uncontrolled (top) and controlled (bottom) self-sustained systems, with $\gamma = 0.3683$.

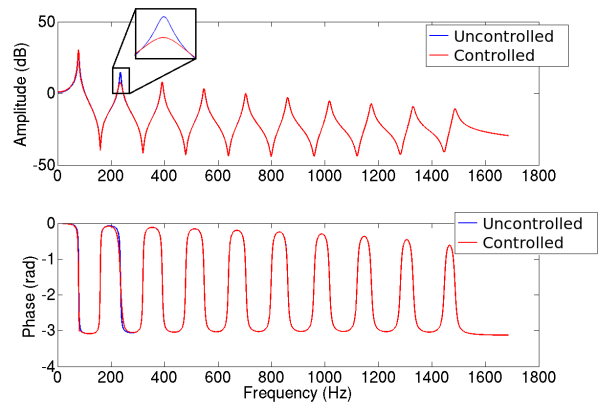


Figure 7. Top : Transfer functions of the uncontrolled (blue) and controlled (red) systems. The control aims to change the damping of the first resonance. It is increased 3 times. Bottom : Phases of the transfer functions of the uncontrolled (blue) and controlled (red) resonators.

Both cases are made blowing exactly in the same way in the tube. The transient is 0.06s longer in the controlled system, and the final amplitude is increased by 11%. The control changes the shape of the steady-state, from a square wave like signal to a sawtooth wave like signal.

3.2 Control of the Damping of the Second Resonance

The control is applied such that the damping of the second resonance is increased 3 times.

Figure 7 shows the transfer functions of the uncontrolled and controlled systems. With the control, the amplitude of the second resonance is decreased by 7dB. The first resonance is also affected by the control and is increased by 1.2dB.

Figure 8 shows the sound spectra of the steady states of the uncontrolled and controlled self-sustained systems, with $\gamma = 0.3533$. With the control applied, the number of harmonics in the sound decreases: seven harmonics when

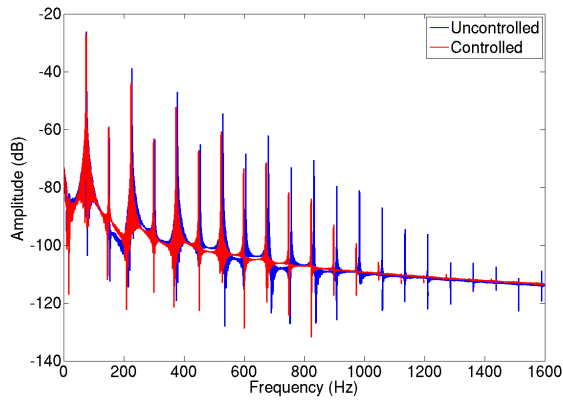


Figure 8. Sound spectra of the uncontrolled (blue) and controlled (red) self-sustained systems, with $\gamma = 0.3533$.

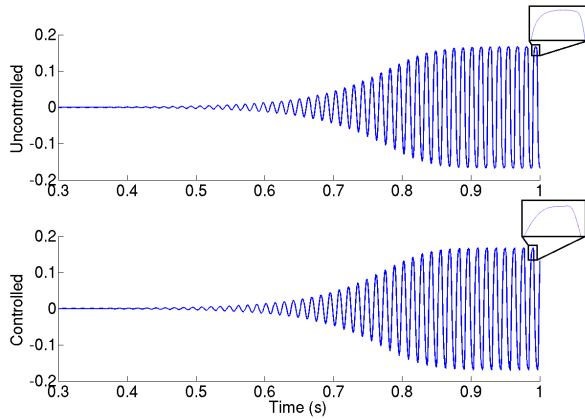


Figure 9. Attack transient of the uncontrolled (top) and controlled (bottom) self-sustained systems, with $\gamma = 0.3533$.

controlled, ten harmonics when uncontrolled. This comes with decreases of the amplitude of the resonances, from 0.6dB (first resonance) to 18.2dB (seventh resonance), and decreases of the frequency of the resonances, from 0.9Hz (first resonance) to 11.3Hz (seventh resonance).

Figure 9 shows the attack transients of the uncontrolled and controlled self-sustained systems, with $\gamma = 0.3533$. Both cases are made blowing exactly in the same way in the tube. The transients are longer than previously because the pressure in the musician's mouth is smaller. In the controlled system, the final shape of the signal is slightly different and its amplitude is increased by 1%.

3.3 Limits of the control of the frequency of the first resonance

The control is applied such that the frequency of the first resonance varies from 60Hz to 160Hz with a 1Hz step. For each frequency, γ varies from 0 (null pressure in the mouth of the musician) to 1 (the reed channel is closed in static regime) with a 0.002 step. Figure 10 shows which frequency is played (color) when the control moves the first resonance to these frequencies with these mouth pressures.

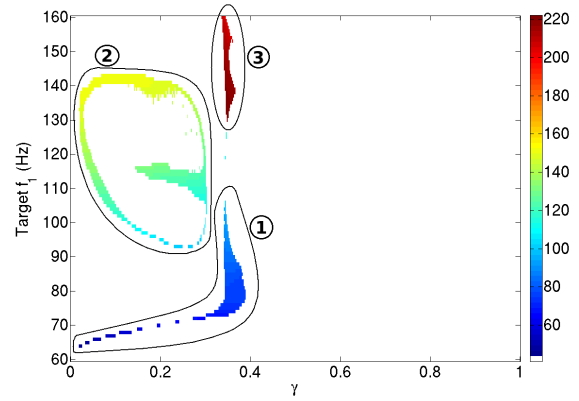


Figure 10. Map of the playing frequencies (color) regarding to the target frequency for the first resonance (ordinate) and to the pressure in the mouth of the musician (abscissa). White parts mean the model can not play.

In this section, to indicate a specific point, the $\{Target f_1; \gamma; Playing frequency\}$ formulation is chosen.

The map in figure 10 shows three parts (circled black).

In the first one, all the playing frequencies are lower than the frequency of the resonance. It gives the lowest playing frequencies, from $\{64; 0.018; 44\}$ to $\{106; 0.344; 100\}$. It also gives the highest mouth pressure with $\{78; 0.388; 75\}$, at the natural f_1 of the tube (78Hz). The playing frequencies have 45 cents ($\{78; 0.388; 75\}$) to 180 cents ($\{64; 0.018; 44\}$) differences with the frequency of the resonance.

In the second part, all the playing frequencies are higher than the frequency of the resonance. It shows stable solutions with low mouth pressures, between $\{133; 0.018; 149\}$ and $\{110; 0.302; 117\}$. The playing frequencies are between $\{93; 0.270; 100\}$ and $\{143; 0.068; 152\}$. The playing frequencies have 94 cents ($\{143; 0.162; 151\}$) to 212 cents ($\{115; 0.154; 130\}$) differences with the frequency of the resonance.

In the third part, all the playing frequencies are those of the second resonance, which is about 236Hz. As a consequence, this part shows the highest playing frequencies. The playing frequencies are between $\{130; 0.348; 222\}$ and $\{160; 0.344; 204\}$. It shows that the second resonance is influenced by the control of the first frequency, as its frequency also moves with large control. In this part, the stable solutions have a mouth pressure close to the main part of part 1, from $\{160; 0.336; 210\}$ to $\{138; 0.366; 217\}$. There is no stable solution beyond $\gamma = 0.388$, whatever the control and with all the other parameters made constant. Globally, the frequency increases when the mouth pressure decreases. In these simulations, every target frequencies between 64Hz and 160Hz have at least one stable solution. The playing frequencies on the first resonance (parts 1 and 2) are between 44Hz and 152Hz, a 21 semitones interval (an octave is 12 semitones).

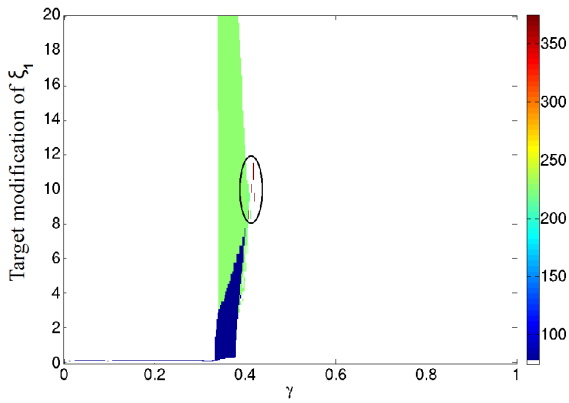


Figure 11. Map of the playing frequencies (color) regarding to the target modification of the damping of the first resonance (ordinate, $\xi = \text{value} \times \xi_1$) and to the pressure in the mouth of the musician (abscissa). White parts mean the model can not play.

3.4 Limits of the control of the damping of the first resonance

The control is applied such that the damping of the first resonance is :

- Decreased, from 0 to 1 times the natural damping (ξ_1) with a $0.05 \times$ step.
- Increased, from 1 to 5 times the natural damping with a $0.2 \times$ step.
- Increased, from 5 to 10 times the natural damping with a $0.5 \times$ step.
- Increased, from 10 to 20 times the natural damping with a $1 \times$ step.

For each damping, γ varies from 0 (null pressure in the mouth of the musician) to 1 (the reed channel is closed in static regime) with a 0.002 step. Figure 11 shows which frequency is played (color) when the control moves the first resonance to these dampings with these mouth pressures. In this section, to indicate a specific point, the $\{Target\ modification; \gamma; Playing\ frequency\}$ formulation is chosen.

The map in figure 11 shows three parts, the blue part, the green part and the red part (last is circled black).

In the first part, the playing frequency is the frequency of the first resonance. At low mouth pressures (from $\{0.1; 0.008; 78\}$ to $\{0.15; 0.344; 78\}$), the playing frequency is exactly the frequency of the first resonance. All the other stable solutions have mouth pressures close to $\gamma = 1/3$, with $0.344 < \gamma < 0.41$. With these mouth pressures, the playing frequencies varies from $\{7.5; 0.41; 75\}$ to $\{0.4; 0.344; 76\}$. The variations of the playing frequencies are almost null.

In the second part, the playing frequency is the frequency of the second resonance, when the damping of the first resonance is high enough. Between $\{3.2; 0.352; 227\}$ and $\{7.5; 0.414; 224\}$, the second and the first resonance are

played alternatively when the mouth pressure is growing. Beyond, from $\{8; 0.352; 228\}$ to $\{20; 0.394; 226\}$, only the second resonance is played. The variations of the playing frequency is of about 3Hz.

Part 3 shows few stable solutions (four in figure 11) where the playing frequency is the frequency of the third resonance, which is about 393Hz. These solutions are at $\{8.5; 0.418; 375\}$, $\{9.5; 0.430; 373\}$, $\{10; 0.424; 374\}$ and $\{11; 0.428; 374\}$.

Most of the stable solutions have a mouth pressure such that $0.344 < \gamma < 0.414$. When the damping is high enough, the playing frequency becomes the second resonance. All the modifications of the damping between $0.1 \times \xi_1$ and $20 \times \xi_1$ have at least one stable solution.

4. CONCLUSION AND PERSPECTIVES

A complete model of a simplified controlled self-sustained oscillating wind instrument has been proposed in order to observe through the simulation of the effects of modal active control. Effects of this control on the sound spectrum and attack transient of a simplified clarinet have been observed. Maps of the control limits of the frequency and damping of the first resonance have been done. The control enables changes in damping and frequency of the resonances of the instrument.

An investigation of the limits of the simulation, that is finding the maximum changes in damping and frequency of the other resonances that are stable, has to be done. A stability study for the adaptation of the control to a real instrument is necessary. A study of the optimal position of the sensor (microphone) and actuator (speaker) regarding to a specific control will be done, then the application of the control to a real simplified instrument with the same dimensions will be done. Finally, playing and perceptive tests with musicians will be done.

Acknowledgments

This work was carried out during the PhD of Thibaut Meurisse, funded by Agence Nationale de la Recherche (ANR IMAREV project) and Universite Pierre et Marie Curie (UPMC, Paris). We thank gratefully Baptiste Chomette for his help with the RFP algorithm, and Sami Karkar and Fabrice Silva for their help with their clarinet model. We thank the Newton Fellowship for funding the collaboration between France and UK.

5. REFERENCES

- [1] J. Hong et al, "Modeling, Identification, and Feedback Control of Noise in an Acoustic Duct," *IEEE Transactions on Control Systems Technology*, vol. 4, no. 3, May 1996.
- [2] A.Preumont, *Vibration Control of Active Structures, An Introduction - Third Edition* Springer, 2011.
- [3] E. Berdahl, J. Smith, "Feedback control of acoustic musical instruments: Collocated control using physical analogs", *J. Acoust. Soc. Am.*, vol. 131, pp.963-973, 2012.

- [4] H. Boutin, *Methodes de controle actif d'instruments de musique. Cas de la lame de xylophone et du violon*. Phd Thesis, Universite Pierre et Marie Curie - Paris VI, 2011.
- [5] S. Hanagud, S. Griffin, "Active Structural Control for a Smart Guitar," in *Proc. of the 4th European Conference On Smart Structures and Materials*, Harrogate, United Kingdom, July 6-8 1998.
- [6] M. E. McIntyre, R. T. Schumacher, J. Woodhouse, "On the oscillations of musical instruments", *J. Acoust. Soc. Am.*, vol. 74, no. 5, pp.1325-1345, November 1983.
- [7] R. T. Schumacher, "Self-Sustained Oscillations of the Clarinet: An Integral Equation Approach", *ACUSTICA*, vol. 40, pp.298-309, 1978.
- [8] F. Silva, J. Kergomard, C. Vergez, J. Gilbert, "Interaction of reed and acoustic resonator in clarinetlike systems", *J. Acoust. Soc. Am.*, vol. 124, no. 5, pp.3284-3295, November 2008.
- [9] S. Karkar, C. Vergez, B. Cochelin, "Oscillation threshold of a clarinet model: A numerical continuation approach", *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp.698-707, January 2012.
- [10] F. Silva, *Emergence des auto-oscillations dans un instrument de musique a anche simple*. Phd Thesis, Universite de Provence - Aix-Marseille I, 2009.
- [11] S. Karkar, *Methodes numeriques pour les systemes dynamiques non lineaires - Application aux instruments de musique auto-oscillants*. Phd Thesis, Universite de Provence - Aix-Marseille I, 2012.
- [12] K. Ogata, *Modern Control Engineering - Fourth Edition* Prentice Hall, 2002.
- [13] B. Chomette, *Contrôles modaux actif, semi-adaptatif et semi-actif de structures intelligentes embarquées, Application aux cartes électroniques*. Phd Thesis, Institut National des Sciences Appliquées de Lyon, 2008.
- [14] T. Meurisse, A. Mamou-Mani, R. Causse, D. Sharp, "Active control applied to wind instruments." in *Acoustics 2012*, Nantes, France, 23-27 April 2012.
- [15] P.A. Nelson, S.J. Elliot, *Active Control of Sound* Academic Press, 1992.
- [16] J. Kautsky, N.K. Nichols, "Robust Pole Assignment in Linear State Feedback", *Int. J. Control*, vol. 41, pp.1129-1155, 1985.
- [17] J. Kergomard, A. Chaigne, *Acoustique des instruments de musique* Belin, 2008.
- [18] T. Richardson, D. Formenti, "Parameter estimation from frequency response measurements using rational fraction polynomials" in *1eIMAC Conference*, 1982.
- [19] T. Meurisse, A. Mamou-Mani, R. Causse, D. Sharp, "Active control applied to simplified wind musical instrument", in *Proc. Int. Cong. on Acoustics 2013*, Montreal, Canada, 2-7 June 2013.

“In vivo” and “in vitro” characterization of single cane reeds

Alberto Munoz, Bruno Gazengel, Jean Pierre Dalmont

Laboratoire d’Acoustique de l’Université du Maine, UMR CNRS 6613,
Av O Messiaen, 72085 Le Mans Cedex 9, France
bruno.gazengel@univ-lemans.fr

ABSTRACT

The aim of this paper is to estimate the characteristic of the reed defined by the relation between the pressure drop across the reed channel and the displacement of the reed tip. Two differential pressure sensors are used to measure the pressure inside the mouth and the mouthpiece. A photointerruptor placed in the mouthpiece is calibrated and used in order to measure the reed tip displacement. *In vivo* measurements are performed when a musician plays a simplified clarinet. Using the measured characteristic *in vivo*, a phenomenological model is derived and enables to estimate some parameters which describe the reed behaviour. *In vitro* measurements are also performed using a vacuum pump to make the reed bend against the mouthpiece lay without any artificial lip. Using a physical model of the reed mechanics, parameters are also deduced to characterise the reed behaviour.

1. INTRODUCTION

Musicians’ experience shows that only a 30% of reeds are good reeds in a box, whereas 40% are mean quality reeds and 30% are considered as bad [1]. According to this, producers, distributors and musicians need a more accurate methodical approach to reed classification. The main issue in this context is developing a scientific approach to satisfy musicians’ request.

In the work developed in the *Laboratoire d’Acoustique de l’Université du Maine* (LAUM), the study of the quality of saxophone reeds carries out three research axes, one subjective axis and two objective axes. The role of each axis is to establish the parameters characterizing the reed and its behaviour. The correlations between these parameters should let us define proper indicators for the description of the reed’s quality. The subjective axis is compounded by a quality test realized by musician while playing the reed. Independent subjective quality indicators are searched. The objective axes are the reed characterization *in vivo* and the reed characterization *in vitro*. The *in vivo* parameters are playing parameters, measured while playing, and the *in vitro* parameters are physical parameters, which are measured by using a measurement bench, adaptable to a industrial bench redesigned for reed producers.

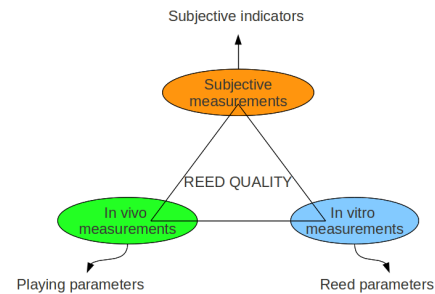


Figure 1. Organization of research to define objective indicators explaining subjective quality of reeds.

Today, work done on first axis has focused on three subjective descriptors to assess the reeds [2]:

- the Brightness of the sound produced with the reed,
- the Softness of the reed, which corresponds to the ease of producing a sound,
- the global perceived Quality of the reed.

Results obtained in the Psychoacoustics study show that the highest agreement between assessors (musicians) is obtained for the descriptor Softness (54.6% of variance on the first component). For Brightness (29.3%), the agreement is weaker, even though no assessor is very discordant. For Quality (29.2%), the agreement is the weakest. Results show also a strong correlation between descriptor Softness and Brightness. Reeds perceived as easy are also perceived as bright and reeds perceived as difficult are also perceived as dull.

Work have also been done on the playing parameters measurement and on the correlation between these parameters and the subjective descriptors. In these works, the mouth pressure and the acoustic pressure radiated at the saxophone bell are measured and playing parameters such Threshold Pressure and Spectral Centroid (SC) are deduced from these signals. Results obtained in [2] and [1] show that the Threshold Pressure deduced from the mouth pressure and the analysis of the sound radiated at the bell is well correlated with the reed Softness. Best correlation is obtained between the *in vivo* Threshold Pressure and the Softness. Moreover, when reeds are classified in three families (easy and bright, medium, difficult and dull), a prediction model using Partial Least-Squares Regression has 4 chances out of 5 to predict correctly the perceived quality

of the reeds from the *in vivo* measurements.

Different authors [3–8] propose to characterize the reed physical parameters *in vitro* and some [5, 8, 9] try to correlate the physical parameters with subjective quality.

In previous work, we focused on the measurement of the low level dynamic response of the reed excited by an acoustic field [1, 8, 10, 11]. Results obtained studying the Frequency Response Function of reeds show that this measurement technique can not be used for explaining subjective quality of reeds. On the one hand, the measurement is not repeatable in a long term whereas we assume that strong quality difference are repeatable in the long term (very difficult reeds remain very difficult some months after). On the other hand, we did not find any correlation between the *in vitro* parameters and *in vivo* parameters or subjective indicators. However, this technique can be used to estimate the mechanical parameters of reeds and can explain the viscoelastic behaviour as shown by Taillard [12].

In order to estimate playing parameters related to the reed behaviour and to be able to characterize the reed *in vitro*, the idea is to measure the reed displacement or the reed opening in both cases (*in vivo* and *in vitro*). Different authors propose *in vitro* techniques for measuring the reed opening using optical measurements. Dalmont [13] propose to measure the light intensity through the reed opening and Almeida [14–16] estimates the reed opening using a camera and a image processing. These authors estimate the volume flow velocity entering the mouthpiece using a diaphragm and two microphones.

Measurements performed *in vivo* generally do not enable to characterize directly the reed behaviour. Often, inverse modelling techniques are used to deduce the reed properties from acoustic measurement made on the resonator during the playing [17], [18], [19].

In this paper, we aim at measuring the reed displacement *in vivo* and *in vitro* in a direct way to try to estimate the reed Stiffness for quasi-static excitation (*in vitro*) and in real playing conditions (*in vivo*).

Section 2 presents the usual model used for describing the reed movement. In section 3, the instrumented mouthpiece and experimental system developed for measuring the reed displacement are presented. Section 4 presents measurements made *in vivo* in which we deduce some mechanical properties of the reed using a simplified model. Section 5 presents measurements of the reed opening performed *in vitro* in a quasi-static manner and the model used for estimating the mechanical equivalent parameters. Finally the results obtained *in vivo* and *in vitro* are compared in §??.

2. PHYSICAL MODEL OF SINGLE REED

The typical model describing the vibration of a reed is Single Degree Of Freedom (SDOF) model taking into account stiffness, damping and inertia. The reed tip displacement y is expressed as a function of the pressure drop $\Delta p = P_m - p$ across the reed channel where P_m is the

mouth pressure and p is the mouthpiece pressure (see figure 2). The reed displacement y is given by

$$\ddot{y}(t) + q_r \omega_r \dot{y}(t) + \omega_r^2 y(t) = \frac{-\Delta p(t)}{m_s}, \quad (1)$$

where $q_r \omega_r$ is the damping coefficient, ω_r is the resonance angular frequency, and m_s the mass of the first mode by unit of surface.

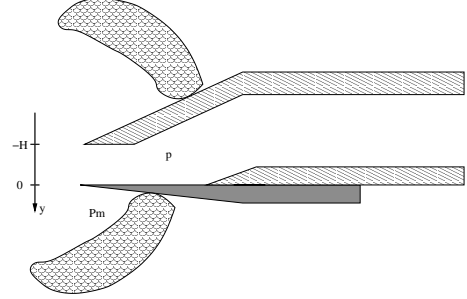


Figure 2. Reed and mouthpiece with physical quantities.

For theoretical simplicity, typically the most simple solution in low frequency of this equation is taken, which is the approximation of linear mechanical response for the reed in terms of the stiffness (Equation 2), defining the stiffness factor as $K_s = m_s \cdot \omega_r^2$:

$$y(t) = -\frac{\Delta p(t)}{K_s}. \quad (2)$$

However, previous studies [13] have shown that the reed characteristic is non linear so that this approximation does not describe the experimental results of the oscillating system with enough accuracy.

Some authors have proposed different models concerning this problem. Chatziioannou *et al* [18] propose a model including a contact force by means of a power law. This model uses 3 parameters. It keeps the stiffness factor K as a constant and adds a critical value of the reed displacement y_c above which the stiffness changes according to a power factor α . Using these assumptions, the equation of the reed movement is given by

$$Ky + K_c (\|y - y_c\|_c)^\alpha = \Delta p, \quad (3)$$

where the conditional norm $\|\dots\|_c$ has been defined as follows:

$$\|y - y_c\|_c = \begin{cases} y - y_c, & \text{if } y > y_c. \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

Dalmont [13], Almeida [20] and Taillard [12] show the effect of viscoelasticity on the reed. In [13], this effect is traduced by an hysteresis in the plane pressure - reed opening. In [20], the viscoelastic behaviour of the reed is characterized by estimating the relaxation time of the reed opening due to a sudden force applied on the reed. Finally, the viscoelastic effect can probably also be observed by analysing the resonance frequency of the reed excited by an acoustic field as described by Taillard [12].

In this work, we try to characterise the non linear behaviour and the viscoelastic behaviour of reeds using *in vivo* and *in vitro* experiments.

3. EXPERIMENTAL SET-UP

The aim of this section is to describe the experimental set up used for characterising the relation between the reed tip displacement y and the pressure drop Δp . We present first the instrumented mouthpiece. Then the principle used for measuring the reed tip displacement is explained.

3.1 Instrumented mouthpiece

The mouthpiece is equipped with two differential pressure sensors Endevco 8507-C2 measuring the mouth and the mouthpiece pressure and with a photointerruptor model SG2BC measuring the displacement of the reed tip (Figure 3).

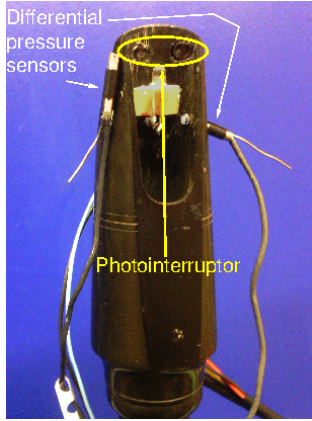


Figure 3. View of the instrumented mouthpiece equipped with two differential pressure sensors and with photointerruptors.

The three sensors output are connected to an acquisition board National instruments BNC-2110 using a sampling frequency $F_s = 50$ kHz.

3.2 Measurement of the reed tip displacement

3.2.1 Principle

A first study of the displacement of the reed for both cases *in vivo* and *in vitro* is proposed by placing IR reflective photointerrupters in a modified mouthpiece, facing the reed, in order to measure the distance between the sensors and the reed all along its displacement. The sensors are placed as shown in Figure 9. Each of the IR photointerrupters, model SG2BC, is composed by a diode, which emits the infrared light that is reflected on the displacing surface, and a transistor which measures the amount of light reflected, producing a current proportional to it.

In a first approximation, only the signal produced by one displacement sensor is analysed.

The phototransistor has to be used with quasi static and dynamic excitation for *in vitro* and *in vivo* experiments respectively. For this reason the calibration is performed in two regimes (static and dynamic calibration).

3.2.2 Static calibration

The static calibration consists in comparing the output voltage of the displacement sensor fixed on a micrometric screw

with the position of the screw. For this measurement, the sensor is placed in front of a reed for taking into account the reflective properties of the cane.

The relation between displacement and output voltage is shown on Figure 4.

The functioning point searched for making the experiment is the region in which the output voltage is almost a linear function of the induced displacement. The sensors are placed on the mouthpiece in order to respect this condition.

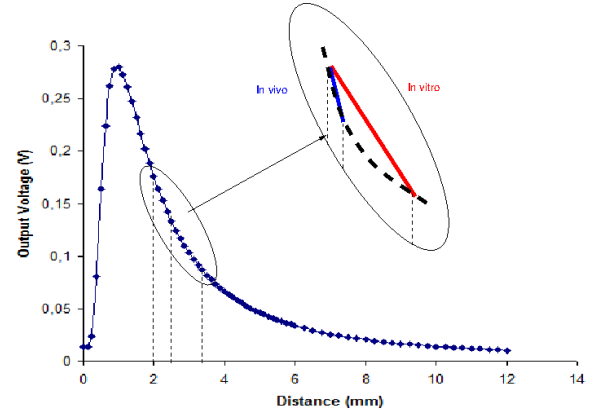


Figure 4. Photointerruptor experimental output voltage vs. distance. View of the different sensitivity estimation for *in vivo* and *in vitro* measurements.

The sketch of the double circuit of the diode-transistor system which has been employed is shown in Figure 5 with $R_1 = 1$ k Ω and $R_2 = 600$ Ω .

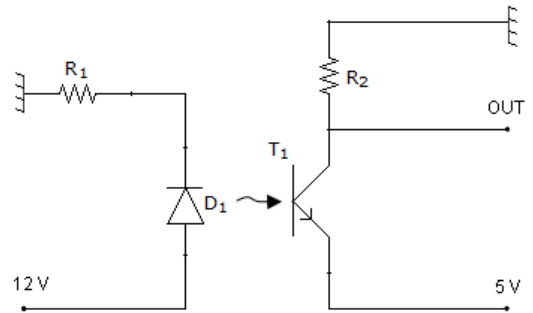


Figure 5. Sketch of the circuit for the reflective photointerruptors SG2BC. $R_1 = 1$ k Ω , $R_2 = 600$ Ω .

3.2.3 Dynamic calibration

The dynamic calibration consists in comparing the response of the displacement sensor (photointerruptor) with the response of a sensor of well known characteristics using a dynamic excitation. The reference sensor is an accelerometer and the excitation is provided by an electrodynamic shaker in which a sinusoidal oscillation is produced. The displacement sensor is placed in front of the moving surface (accelerometer glued on the shaker head) at the dis-

tance corresponding to its linear response (almost 2 mm). The Frequency Response Function (FRF) between the photointerruptor output voltage and accelerometer output voltage is obtained using a spectrum analyser.

For the calibration experiment, different magnitudes of amplitude for the sinusoidal oscillation were considered, obtaining quite independent measurements, with the exception of large amplitudes (10^{-3} m).

The measured FRF shows that the response of the photointerruptor is almost flat in amplitude and phase in the frequency range 20 Hz - 1 kHz.

3.2.4 Conclusions

In order to adapt the sensitivity to the mean position of the reed, we consider two different values of the sensitivity for *in vitro* and *in vivo* measurements with a variation of 60% the obtained slope (Figure 4). For future applications, we will consider a model of the sensitivity in order to avoid such an approximation.

The dynamical calibration has shown that the photointerruptor provides good results for the dynamical measurements in the frequency range [0-1] kHz. The sensitivity estimated by dynamic excitation is of same order of magnitude than the sensitivity estimated *in vivo*.

Finally, the final set up is compact enough to be used for both *in vivo* and *in vitro* measurements.

4. IN VIVO ANALYSIS

The aim of the experiment performed *in vivo* is to observe the relation between pressure drop Δp and the displacement of the reed tip y . This experiment should enable to estimate equivalent parameters that could be related to the physical model while the cane reed is put on an instrument played by a real musician.

For the sake of simplicity, the experimental system is made of a cylindrical tube (length about 60 cm) used as a resonator connected with a tenor saxophone mouthpiece. For each experiment, a single note is played during a few seconds producing a signal with fundamental frequency of 130 Hz.

4.1 Signal processing

In a first step, the three signals are low pass filtered using a second order filter (Butterworth) with a cut off frequency of 1 kHz. Although the filter phase response is not linear, the relation between pressure and displacement is not affected by the filter as verified by using noisy square signals applied to this filter.

Then, the three signals are analysed using the principle described in [1] for note detection, Threshold Pressure estimation and spectrum analysis. This estimation enables to deduce the value of the Threshold Pressure and to calculate the fundamental frequency Stiffness defined as $K_{f1} = \frac{\Delta p(f_1)}{y(f_1)}$ where $\Delta p(f_1)$ and $y(f_1)$ are respectively the amplitudes of the pressure drop and reed tip displacement at

the fundamental frequency obtained thanks to the Fourier Transform of both signals.

Assuming that the phase response of the pressure sensors are zero and knowing that the phase response of the displacement sensor is less than 5 degrees in the frequency range of interest, it is possible to study the relation between pressure drop Δp and displacement y .

Using the stationary part of the signal estimated as described in [1], we detect each signal period to represent the displacement $y_k(t)$ as a function of the pressure drop $\Delta p_k(t)$ for period k as shown in Figure 6.

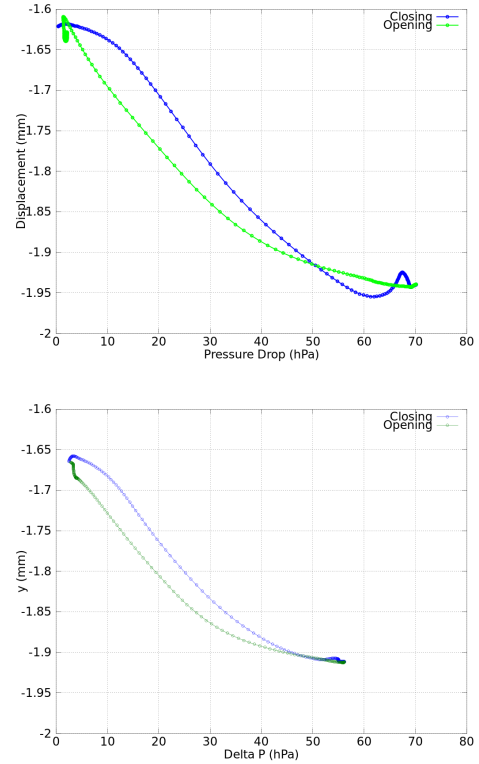


Figure 6. Example of experimental characteristic $\Delta p - y$ for one period of the stationary part of the signal. Top : reed A (evaluated as difficult), bottom : reed B (evaluated as easy).

4.2 Model of the reed for *in vivo* analysis

Figure 6 shows that the relation between displacement and pressure drop is not linear and shows hysteretic effects. During the closing, we assume that the characteristic can be approximated with a 3^{rd} order polynomial. During the opening, we assume it can be approximated with a 2^{nd} order polynomial. Using these assumptions, we can deduce the following parameters (Figure 7) for each period k :

- Closing phase:
 - Equivalent Compliance,
 - Inflection Pressure,
 - a third parameter could be estimated.
- Opening phase:

- Equivalent Compliance,
- Final Compliance,
- Curvature.
- The losses are estimated by calculating the area inside the curve.

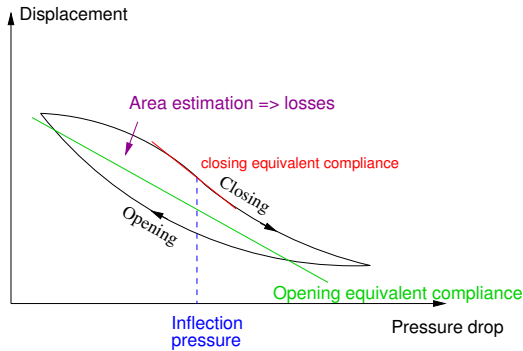


Figure 7. Principle of equivalent reed parameters estimation *in vivo*.

An example of parameters estimated using this approach is shown in figure 8 as a function of the period number. In order to obtain a single value for each parameter, the mean value of this parameter is calculated over the different periods.

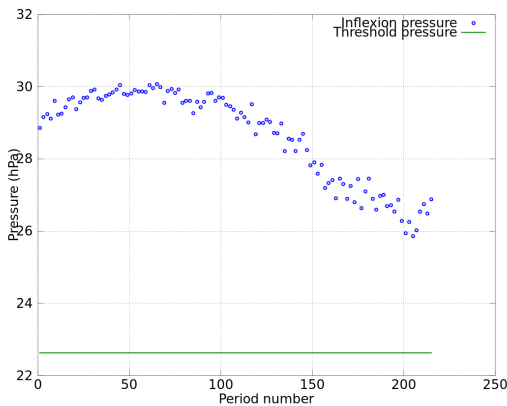


Figure 8. Example of the inflection pressure estimated using the analysis of the characteristic $\Delta p - y$ as a function of the period number.

5. IN VITRO EXPERIMENTS

The aim of the *in vitro* experiments is to observe the relation between the reed displacement and the pressure drop in quasi static regime. The reed is mounted without any artificial lip and the pressure drop is controlled manually by an operator which opens and closes a valve.

5.1 Experimental system

The experimental set-up is shown in Figure 9. The source of the difference of pressure is a vacuum pump which cre-

ates a depression at the end of a duct connected to the mouthpiece. The difference of pressure between the mouthpiece and the atmosphere is measured by placing a differential pressure sensor Endevco 8507-C2 in the mouthpiece similarly to *in vivo* experiments. The displacement of the cane reed is measured with one of the reflective photointerruptors SG2BC.

The excitation of the reed is artificially generated by opening and closing a valve, producing the displacement of the reed from the state of total opening to the closing state.

This system does not enable to obtain a stable position of the reed at any position as described in [13] even when using a diaphragm. This phenomenon seems to be due to the internal characteristic of the vacuum pump. Due to this effect, the operator needs to close the valve suddenly in order to close the reed against the lay of the mouthpiece. The opening and closing phase duration is about 0.1 second. Each measurement is taken for 8-10 beats (one beat corresponds to one closing and one opening) of the reed, in a time of 10 seconds.

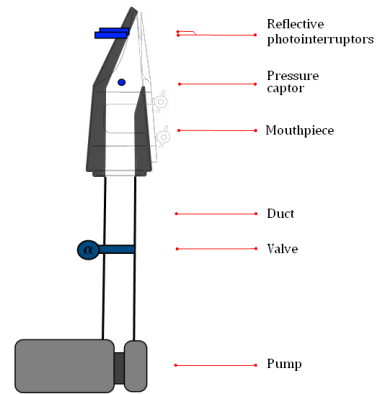


Figure 9. Schematic view of the *in vitro* experimental system.

The order of magnitude of the pressure drop differs from the “*in vivo*” experiment and the maximum value is about 400 hPa (figure 10) which creates a reed deformation inside the mouthpiece after the reed has closed.

In order to study only the behaviour of the reed between the opening state and the closing state without taking into account the deformation, the displacement of the reed is studied only from the states of near opening and near closing using two threshold values ($\simeq 0.1$ mm and $\simeq 0.8$ mm, as shown in figure 11).

As for *in vivo* experiments, a low-pass filter Butterworth was applied to signals to reduce noise (cut of frequency of 1 kHz), taken in this case with a sampling frequency $F_s = 1$ MHz.

5.2 Model of the reed for *in vitro* analysis

The characteristic $\Delta p - y$ obtained for one reed and 11 beats is shown in Figure 12. This figure shows two phases:

- during the closing phase, the reed behaviour is non linear and the curve shows two parts. Above a critical pressure, the stiffness increases

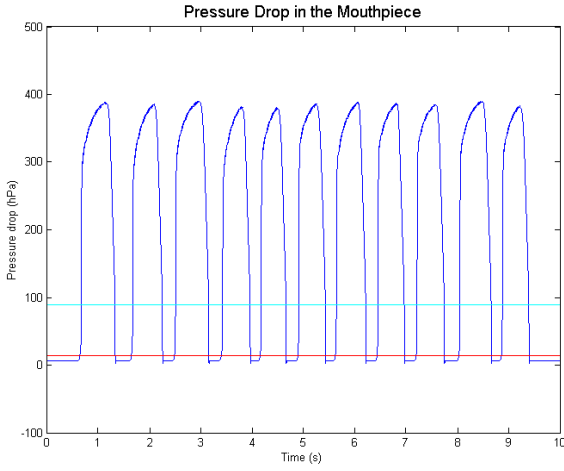


Figure 10. Measured pressure drop vs. time for 11 beats (opening and closing).

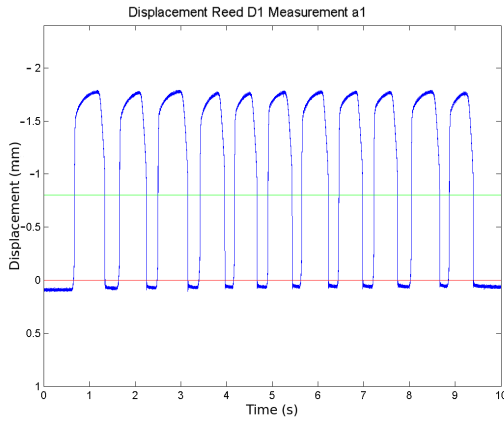


Figure 11. Measured displacement vs. time for 11 beats (opening and closing).

- during the opening phase, the reed behaviour is linear and its stiffness is almost the same than the stiffness observed below the critical pressure.

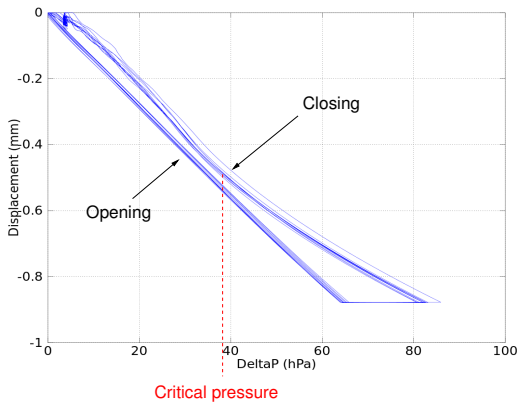


Figure 12. Displacement vs. pressure drop for 11 opening and closing phases done by an operator using a valve.

In order to estimate the reed parameters, we propose to use a simplified version of the non linear model described in equation 3. In a first order, we consider $\alpha = 1$ and we are simplifying this model by taking the following approach:

$$\Delta P = \begin{cases} -K_1 \cdot y & , \text{if } \Delta p < P_c. \\ -K_2 \cdot y & , \text{if } \Delta p > P_c. \end{cases} \quad (5)$$

This critical pressure P_c is defined as the pressure corresponding to the point of maximal curvature of the measurement $y(\Delta P)$.

Figure 12 shows that the model given in equation 5 leads to an estimation of $K_{1,close}$, $K_{2,close}$ and P_c during the closing. During the opening the reed shows a linear behaviour which can be traduced by the opening stiffness K_{open} (Figure 13).

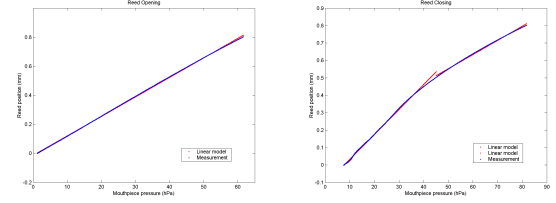


Figure 13. Stiffness factor for the opening (left) and the closing (right) of a single beat of one reed.

Assuming the two models, the obtained physical parameters are:

- Closing phase:
 - first Stiffness $K_{1,close}$ in non linear modeling (NL),
 - second Stiffness $K_{2,close}$ in non linear modeling (NL),
- Opening phase:
 - Stiffness K_{open} in linear modeling (L),
- The losses are estimated by calculating the area inside the curves.

For the measurement comprising 10 beats, the estimated parameters show a variance for each beat as shown for example in Figure 14.

For each reed and each measurement (11 beats), a mean value of each parameter is calculated. For each reed, four measurements are taken in short-term intervals. A final mean of the four measured factors is calculated and the uncertainty u_p on parameter p is obtained by $u_p = 2.25 \frac{\sigma_p}{\sqrt{N}}$ where σ_p is the standard deviation on the N measurement with $N = 4$.

6. CONCLUSIONS

An experimental device (instrumented mouthpiece) has been developed for measuring the reed displacement y and the pressure drop Δp across the reed channel. This device is

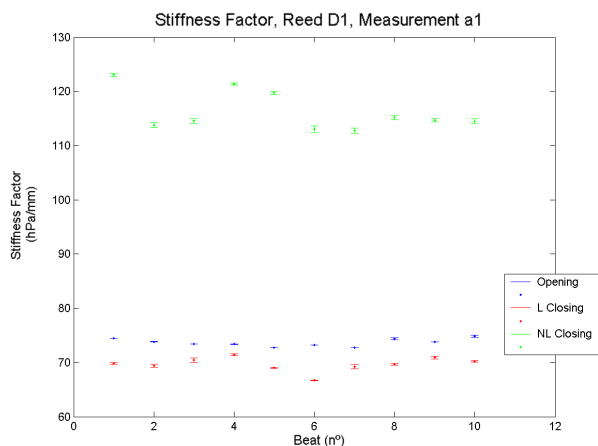


Figure 14. Stiffness factors for 10 beats of one reed.

used during *in vivo* experiments and enables to show that the characteristic curve $\Delta p - y$ is non linear and reveals an hysteretic behaviour of the reed. The results obtained *in vivo* do not show any linear behaviour for the reed, which can be explained by the fact that the reed is coupled with the musician's lip.

Experiments made *in vitro* with this instrumented mouth-piece use a vacuum pump to produce a reed movement in a quasi-static way. Results show that the reed closing is non linear and the opening is linear. hysteric effects are also observed.

First measurements performed on four reed considered as subjectively different (two difficult and two easy reeds) show that some indicators obtained "*in vivo*" and "*in vitro*" seem to explain the perceptive difference between reeds.

For future studies, a calibration function for the photointerruptor will be developed to compensate for its linear response. New displacement sensors will be also studied to try to get a linear sensibility. Moreover, a measurement of the displacement of the reed in different points will be developed to measure assymetric movements.

7. REFERENCES

- [1] B. Gazengel and J. Petiot, "Objective and subjective characterization of saxophone reeds," *submitted to Journal of New Music Research*, 2013.
- [2] J. Petiot, P. Kersaudy, G. Scavone, S. McAdams, and B. Gazengel, "Modeling of the subjective quality of saxophone reeds," in *ICA 2013, 2-7 June 2013*, 2013.
- [3] P. Kolesik, A. Mills, and M. Sedgley, "Anatomical characteristics affecting the musical performance of clarinet reeds made from arundo donax l. (gramineae)," *Annals of Botany*, vol. 81, no. 1, pp. 151–155, 1998.
- [4] E. Obataya and M. Norimoto, "Acoustic properties of a reed (arundo donax l.) used for the vibrating plate of a clarinet," *The Journal of the Acoustical Society of America*, vol. 106, p. 1106, 1999.
- [5] F. Pinard, B. Laine, and H. Vach, "Musical quality assessment of clarinet reeds using optical holography," *The Journal of the Acoustical Society of America*, vol. 113, p. 1736, 2003.
- [6] S. C. Mukhopadhyay, G. S. Gupta, J. D. Woolley, and S. N. Demidenko, "Saxophone reed inspection employing planar electromagnetic sensors," *Instrumentation and Measurement, IEEE Transactions on*, vol. 56, no. 6, pp. 2492–2503, 2007.
- [7] P. Picart, J. Leval, F. Piquet, J. Boileau, T. Guimezanes, and J. Dalmont, "Study of the mechanical behaviour of a clarinet reed under forced and auto-oscillations with digital fresnel holography," *Strain*, vol. 46, no. 1, pp. 89–100, 2009.
- [8] B. Gazengel, J.-F. Petiot, and M. Soltes, "Objective and subjective characterization of saxophone reeds," in *Proceedings of Acoustics 2012*, 23-27 April 2012,.
- [9] M. L. Facchinetti, X. Boutillon, and A. Constantinescu, "Numerical and experimental modal analysis of the reed and pipe of a clarinet," *The Journal of the Acoustical Society of America*, vol. 113, p. 2874, 2003.
- [10] B. Gazengel, T. Guimezanes, J.-P. Dalmont, J. B. Doc, S. Fagart, and Y. Léveillé, "Experimental investigation of the influence of the mechanical characteristics of the lip on the vibrations of the single reed," in *Proceedings of the International Symposium on Musical Acoustics, Barcelona, Spain*, 2007.
- [11] B. Gazengel and J. Dalmont, "Mechanical response characterization of saxophone reeds," in *6th Forum Acusticum*, 26 June-1 July 2011.
- [12] P. Taillard, F. Laloë, M. Gross, J.P., Dalmont, and J. Kergomard, "Measurements of resonance frequencies of clarinet reeds and simulations," *arXiv preprint arXiv:1202.2114*, 2012.
- [13] J. Dalmont, J., Gilbert, S., and Ollivier, "Nonlinear characteristics of single-reed instruments: Quasistatic volume flow and reed opening measurements," *The Journal of the Acoustical Society of America*, vol. 114, p. 2253, 2003.
- [14] A. Almeida, C. Vergez, R. Caussé, and X. Rodet, "Physical study of double-reed instruments for application to sound-synthesis," in *Proceedings of the International Symposium on Musical Acoustics*, 2002, pp. 215–220.
- [15] C. Vergez, A. Almeida, R. Causse, and X. Rodet, "Toward a simple physical model of double-reed musical instruments: influence of aero-dynamical losses in the embouchure on the coupling between the reed and the bore of the resonator," *Acta Acustica united with Acustica*, vol. 89, no. 6, pp. 964–973, 2003.
- [16] A. Almeida, C. Vergez, and R. Caussé, "Quasi-static non-linear characteristics of double-reed instruments," *arXiv preprint physics/0607011*, 2006.

- [17] X. Boutillon and V. Gibiat, "Evaluation of the acoustical stiffness of saxophone reeds under playing conditions by using the reactive power approach," *The Journal of the Acoustical Society of America*, vol. 100, p. 1178, 1996.
- [18] V. Chatzioannou and M. van Walstijn, "Estimation of clarinet reed parameters by inverse modelling," *Acta Acustica united with Acustica*, vol. 98, no. 4, pp. 629–639, 2012.
- [19] T. Smyth and J. S. Abel, "Toward an estimation of the clarinet reed pulse from instrument performance," *The Journal of the Acoustical Society of America*, vol. 131, p. 4799, 2012.
- [20] A. Almeida, C. Vergez, R., and Causse, "Experimental investigation of reed instrument functioning through image analysis of reed opening," *Acta acustica united with acustica*, vol. 93, no. 4, pp. 645–658, 2007.

A Digital Bagpipe Chanter System to Assist in One-to-One Piping Tuition

Duncan W. H. Menzies

Centre for Digital Music
Queen Mary University of London
London, UK

d.w.h.menzies@eecs.qmul.ac.uk

Andrew P. McPherson

Centre for Digital Music
Queen Mary University of London
London, UK

andrewm@eecs.qmul.ac.uk

ABSTRACT

This paper describes an electronic bagpipe chanter interface and software system, developed to assist in one-to-one Highland piping tuition. The chanter employs infrared reflectance sensors to detect the continuous movements of the player's fingers, and incorporates an air pressure sensor in place of the chanter reed, allowing it to be connected to a traditional acoustic set of pipes. The software is intended to assist the instructor in communicating feedback to the student by providing facilities for recording, playback, visualisation and comparison of teacher and pupil performances. A user study of the system was carried out with an experienced piping instructor and seven students. The sessions yielded encouraging and constructive feedback from both students and instructor, and produced promising avenues for further work.

1. INTRODUCTION

The Great Highland Bagpipe (GHB) is widely regarded, at least among pipers, as an instrument with a high barrier to entry. The Highland piping tradition requires the aspiring player to memorise a diverse array of distinct and formally defined ornamentation techniques before attempting all but the simplest of tunes; a process that can often take six to twelve months of regular and disciplined practice. Historically, bagpipe music was passed on through the instructor singing to the student in a precise musical language known as *canntaireachd*. Indeed, piping notation is a comparatively recent development, having been introduced in the early 19th century [1]. The use of sheet music in bagpipe lessons is now reasonably common. Nonetheless, GHB music is generally devoid of any phrase markings or other high level performance instructions of the kind that might be seen in classical music notation.

The GHB provides no facility for dynamic control, and produces a constant, uninterrupted sound, preventing the use of silences or timbral changes for the purposes of emphasis or articulation. Variations in rhythmic phrasing are thus an integral aspect of expressive bagpipe performance, and one of the primary means by which proficient pipers

can convey their own interpretations of the otherwise largely inflexible traditional repertoire.

Communicating such subtle temporal deviations can be a challenging task for piping tutors. While singing or playing a passage for the student to repeat is undoubtedly effective, it is often necessary for the teacher to verbally describe their intentions. This can lead to the use of somewhat abstract language such as “push out the first beat” and “the G gracenote takes you to the E doubling on the beat”, which can be difficult to understand, even for students with significant experience of other musical instruments.

The aim of this work is to develop teaching tools which are specifically tailored to the requirements of the piping community, with the goal of assisting and accelerating the learning process in the context of one-to-one lessons. This paper presents a digital GHB chanter interface and accompanying software system which enables the recording, playback, visualisation and comparison of teacher and pupil performances. This is intended to help the instructor illustrate and convey feedback to the student.

2. RELATED WORK

2.1 Electronic Bagpipes

Several brands of electronic Highland bagpipes are commercially available, of which the DegerPipes¹, TechnoPipes² and Redpipes³ are most prominent. These use single capacitive touch-switches in place of the finger-holes, which are binary in nature; the “holes” are always either fully open or closed. This does not accurately reflect the finger-holes of an acoustic chanter, which can be gradually covered and uncovered to slide between notes.

There have been several attempts within the academic community to develop alternatives to this discrete sensor strategy. The FrankenPipe [2] uses photoresistors mounted inside the holes of an acoustic GHB chanter. This provides a wide analogue range for each hole, and has the advantage of retaining the physical feel of a traditional chanter. The EpipE [3] is a uilleann bagpipe chanter interface, which extends the capacitive sensing approach to include an array of sixteen small binary touch-switches for each hole.

While the Redpipes and EpipE have the capability to measure the pressure exerted on the bag by the player's arm (e.g. using force-sensitive resistors), the authors are not

Copyright: ©2013 Duncan W. H. Menzies et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ <http://www.deger.com/>

² <http://www.fagerstrom.com/technopipes/>

³ <http://redpipes.eu/>

aware of any existing electronic chanter which can be connected to a standard set of bagpipes and controlled directly using air pressure. The interface presented in this work achieves this using an air pressure sensor similar to those employed in experimental wind controllers such as the CyberWhistle [4] and *The Pipe* [5].

2.2 Technology in the Context of Music Tuition

The use and development of technological tools for musical education is an active field of research. A significant proportion of existing work in this area is concerned with piano pedagogy using MIDI input from a digital keyboard, due at least in part to the MIDI protocol providing a simple means of capturing multiple aspects of a performance. One such project is the Piano Tutor [6], which combines score-following software and performance evaluation algorithms with extensive multimedia feedback in order to “create a natural dialogue with the student”. The *pianoFORTE* system [7] produces visualisations of tempo, articulation and dynamics of a performance in the form of an annotated musical score. In addressing the development of tools to assist in one-to-one instrumental instruction, the authors assert that the aim is not to automate the teacher, but to facilitate the “difficult communication process” through which the instructor attempts to describe the subtleties of expressive interpretation beyond simply playing the correct notes.

The *MIDIator* [8] program takes MIDI input to allow the user to compare separate renditions of the same piece by producing graphs to illustrate variations in tempo, note velocity and duration. In addition to MIDI, the SYSSOMO system [9] uses raw audio, video and motion data from accelerometer and gyroscope sensors to capture a comprehensive record of a pianist’s movements. A score following algorithm is employed to align and superimpose two performances with different tempi, enabling direct visual comparison between the playing of instructor and student.

The *i-Maestro* tool [10] records audio, video and VICON motion capture data of musicians playing bowed string instruments. This information can be played back and displayed in a variety of formats to help the tutor “identify, illustrate and explain certain issues involved with performance”. The Digital Violin Tutor [11], intended primarily as a solo practice tool to provide feedback in the absence of an instructor, employs a transcription algorithm to visualise and compare the student’s playing with an existing score, or earlier recording made by the teacher.

A study by the Office for Standards in Education (Ofsted) in which inspectors visited 52 schools around the UK highlights several ways in which technology can “enable attainment”, “enhance progress” and “increase pupils’ motivation” in music classroom contexts [12]. It is again noted that the tools should not take over the role of teacher, but should instead be employed to help clarify conceptual information for the student. For example, in lessons concerning critical listening, tutors could reduce or enhance the level of certain parts of a multitrack audio piece in order to help pupils focus on musical features which were previously imperceptible.

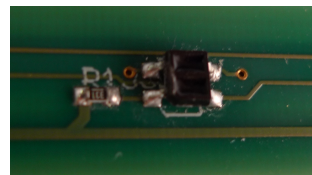


Figure 1. Close-up of infrared reflectance sensor.



Figure 2. Complete chanter and PCB.

3. HARDWARE DEVELOPMENTS

3.1 Physical Construction and Sensing Strategy

The purpose of the electronic chanter hardware is to detect the continuous movements of the player’s fingers quickly and accurately, and to transmit this data to the host computer via USB. The interface described in this paper extends a prototype first presented in [13], which employs infrared (IR) LED and photodiode pairs for each hole, between which a constant IR beam exists. The player covers the “hole” by interrupting this beam with a finger.

This strategy was successful in providing a continuous analogue reading for each hole. However, the physical construction of the original chanter led to a somewhat unnatural playing experience, primarily as a result of being built using strip-board and through-hole components. The spacing between the sensors was dictated by practical layout constraints, and the requirement that the player’s fingers sit between the IR emitter and detector prevented the board from being housed in a cylindrical shell. While appropriate as a tool to investigate the ornament recognition concepts that were the focus of the previous work, this interface was deemed unsuitable for use in studying the existing playing technique of piping students in a lesson.

One of the primary concerns when developing the improved hardware was therefore to make the physical playing experience as similar as possible to a traditional GHB. A custom printed circuit board (PCB) was designed, employing an integrated IR reflectance sensor for each hole (Figure 1). The distances between the sensors reflect the hole spacing of an acoustic chanter.

Each sensor is comprised of an IR LED and phototransistor in a single package, both directed upwards. When an object comes within range of the sensor, the IR radiation from the LED is reflected back and detected by the phototransistor. This allows the proximity of the player’s finger to the sensor to be measured with a high degree of precision. Moreover, this sensing strategy allows the board to be mounted inside a cylindrical casing with real holes, providing a more realistic playing experience than contact dependent approaches such as capacitive touch pads. The completed chanter and PCB are shown in Figure 2.

3.2 Microcontroller Processing

The PCB incorporates an ARM 32-bit Cortex-M3 microcontroller which gathers and processes the raw sensor data before transmitting it to the computer. Timing synchronisation is handled by the onboard clock. The board sends one complete 20 byte message via USB every millisecond, providing accurate temporal information about the performance. Each message is comprised of a timestamp and nine sensor values in a packed binary representation.

For the interface to be effective in accurately measuring the player's finger movements, it is important that the sensor readings are as stable and reliable as possible. Each of the IR sensors is therefore read eight times during every millisecond period and an average taken to reduce the effect of inaccuracies caused by momentary fluctuations. Moreover, since the sensing strategy uses optical reflectance, it is necessary to account for variations in ambient light. By measuring the output of each sensor with the LED off directly after each reading, an indication of the current background conditions is obtained. This measurement is subtracted from the original sensor reading, ensuring that the final value is robust to environmental interference.

3.3 Air Pressure Sensor

In addition to the intricate fingering technique needed to reproduce a melody on the chanter, the GHB requires a steady flow of air through the chanter and drone reeds. This involves applying a constant pressure to the bag with the arm, and a significant degree of physical exertion to keep the bag filled with air. These essential aspects of bagpipe playing cannot be addressed using a standard practice chanter (a quieter single pipe instrument with no bag, sounding roughly an octave below the GHB), and thus regular practice sessions on a full set of pipes are traditionally required to maintain the necessary endurance. However, the high sound intensity levels involved can render this impractical for many pipers (e.g. those living in urban areas).

Therefore, a technological system that allows the user to work on the breathing and arm pressure elements of GHB technique at any volume (or wearing headphones) could provide significant benefit to the piping community. To achieve this, the hardware developed in this work incorporates a MPXV5010 pressure sensor at the top of the PCB, and a small hole for the air to escape as it would through a conventional chanter reed. By closing off the drones of a standard set of pipes using stoppers and inserting the electronic interface into the bag in place of an acoustic chanter, the player can control the instrument using exactly the same physical interactions as with a traditional GHB.

This provides a complete and realistic playing experience, and allows the user to practice all aspects of Highland piping technique without any acoustic sound being produced. Moreover, the pressure at which the drones and chanter sounds are activated can be specified and modified in the software, enabling the player to adjust the strength of the virtual "reeds" and progressively develop stamina.

4. TUITION SOFTWARE SYSTEM

4.1 Communication and Audio Output Software

The aim of the work presented in [13] was to produce an algorithm for automatic recognition of Highland piping ornamentation. This was intended primarily as a solo practice tool to provide novice pipers with immediate feedback on their technique in the absence of an experienced tutor, in order to avoid the introduction of bad habits between lessons. By contrast, the focus of this paper is not on style-specific machine intelligence, but rather the development of an analytical tool to assist piping instructors in communicating their feedback to the student during a lesson.

The software framework consists of three components which communicate via the Open Sound Control (OSC) protocol [14]. A simple command line utility receives and interprets the incoming serial data from the chanter. The unpacked sensor readings are transmitted to two separate programs: the tuition software; and a standalone Super-Collider [15] application that produces the audio output. Rather than employing a wavetable synthesis approach as in commercially available digital chanters, this system instead uses sampled GHB recordings. Given the stationary nature of bagpipe sounds, and the lack of expressive parameters in interacting with the traditional instrument, this method provides a highly convincing sound. Moreover, the system can be easily extended to include samples from other varieties of bagpipe such as the Scottish smallpipes.

4.2 Functionality of Tuition Software

Experienced Highland pipers frequently employ creative rhythmic phrasing around the strong beats of the bar to add expression to their playing. While an important aspect of advanced performance, which many proficient players may understand intuitively, these subtle temporal variations can be difficult to describe. The purpose of this program is therefore to provide an illustrative tool to aid the tutor in describing and explaining their comments on a student's technique. This requires the ability to capture the performance and to represent it in a clear and intuitive format.

The system allows the instructor to record a piece of any length, tempo and time signature to a metronome track. This performance can then be displayed on screen as shown in Figure 3. The visualisation uses the traditional staff system to illustrate pitch (maintaining the piping convention of omitting the accidental symbols on the C \sharp and F \sharp) and barlines are shown. Note duration is portrayed using a proportional notation similar to the familiar piano roll format. This representation allows the nuances of a player's rhythmic phrasing, which would be obscured by standard classical notation, to be clearly and explicitly depicted.

Using the instructor's recording as a guide, the student can attempt to play the same piece. While recording, a playhead scrolls through the staves to illustrate the current point in the sequence (continuing onto multiple pages if necessary), and the results of the pupil's playing can be shown contemporaneously or hidden as preferred. Once complete, the two performances can be displayed either individually or overlaid in different colours (Figure 4), al-

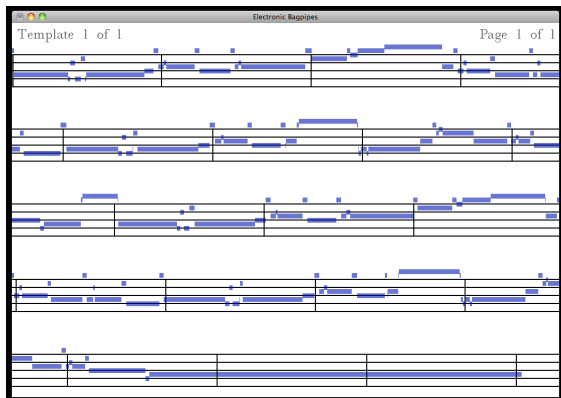


Figure 3. Visualisation of instructor's performance.

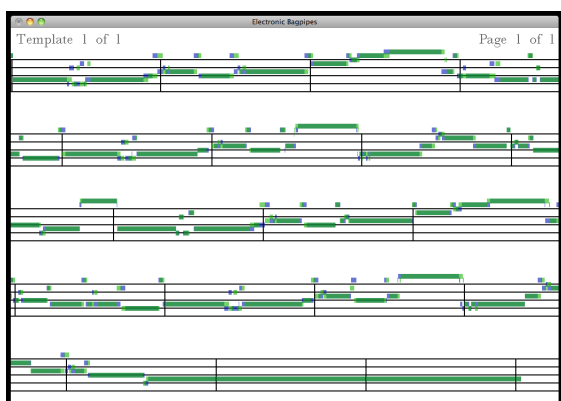


Figure 4. Instructor and student performances overlaid.

lowing subtle variations in timing to be identified and examined visually. Since aural training is a crucial aspect of musical learning [16], the recordings can also be played back from any point, separately or together, and at varying playback speeds by clicking on the screen. This enables direct and repeatable comparison between the two, promoting critical listening on the part of the student and providing the instructor with an additional tool with which to explain their feedback. The system also includes the facility to log performances, such that they can be reopened in future sessions to gauge ongoing progress.

4.3 Illustrative Example of Tuition Software

Figure 5(a) shows the Highland piping notation for high A to low A with a *birl* ornament on the first beat of the bar. It is standard practice in GHB music that the stems of all melody notes point down, while embellishments are written as demisemiquavers with smaller note heads and stems pointing upwards. It is important to note that ornaments are not assigned any durational value in the score; a bar of $\frac{4}{4}$ will contain four full crotchet melody notes regardless of the number of embellishments. This can cause significant confusion as to how and where an ornament should be played, particularly among piping students who already have some experience with other musical instruments.

While there are some general guidelines as to where certain embellishments should be performed with regard to the beat, this is one of comparatively few aspects of Highland piping which is open to interpretation by the player.

Instructors will therefore seek to guide their students towards particular phrasing characteristics depending on the desired expressive effect.

Figure 5 illustrates how the visualisation software can be used by the tutor to help explain the nuances of such feedback. A straightforward interpretation of the notation from Figure 5(a) would involve the birl ornament being played directly on the beat, as shown in Figure 5(b). A more experienced player may instead wish to perform the embellishment fractionally earlier, in order to give the piece a more “lively” edge. This is clearly visible in Figure 5(c).

Such concepts are not easily described, and piping ornamentation is executed so quickly that even practical demonstration by the instructor is not always sufficient to fully clarify the distinction between the two renditions. The software system presented in this work provides a simple and unambiguous visual representation of subtle temporal variations in order to assist the tutor in conveying this central aspect of Highland piping technique to students.

5. USER STUDY

5.1 Location and Participants

An initial user study of the system was carried out at a private boarding school in the North East of Scotland. The study took place under the supervision of the school's piping instructor, a highly proficient piper with around thirty years experience of playing and twenty years of teaching, thirteen of which had been spent at the school. At present there are forty-seven piping students at the school, aged between eight and eighteen years. Based on his detailed knowledge of the pupils' playing, the instructor selected seven students to participate in the study. The participants were aged between thirteen and seventeen years, and their playing experience ranged from six months to eight years.

5.2 Purpose and Structure of Study

The study took place over a period of four days, the first of which was spent with the instructor only, in order to gather and address his initial comments on the system prior to using it with the pupils. A short interview was also carried out to learn more about his approach to teaching, and some of the particular challenges faced by piping instructors. In response to this discussion, some minor adjustments to the sensitivity of the finger position sensors were made in order to make the playing experience as similar as possible to an acoustic bagpipe chanter.

Each of the students had one session (between 30-60 minutes) with the digital chanter as part of their normal one-to-one lesson time. In each case the student was given some time to get used to the interface. The instructor would then record a tune while the pupil listened, following which the visualisation and playback mechanisms were demonstrated. The student was then instructed to play the same piece. Once the student had finished recording, the instructor would use the visualisation and playback functions to illustrate his observations about the pupil's performance. This process was typically repeated several times per lesson, often with different tunes.

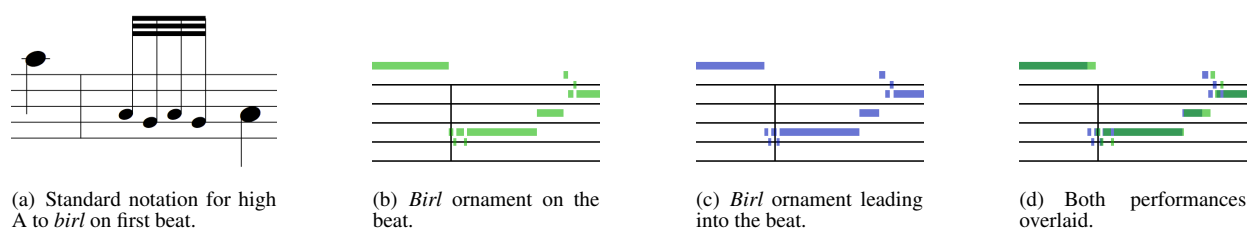


Figure 5. Example of how the visual display can be used to illustrate subtle variations in rhythmic phrasing.

5.3 Observations and Outcomes

5.3.1 Instructor Feedback

The developments made during this work with regard to how bagpipe teaching could be best supported using technology were largely informed by the author's personal experience as a piping student. During the introductory interview, the instructor indicated his agreement with many of these assumptions, saying of rhythmic variation in GHB performance that "pipe music is so different... the way we express our tunes isn't the way it's written". In particular, he described his experience of teaching the former Director of Music at the school, who had struggled with the phrasing of many ornaments, despite having a wealth of musical knowledge and proficiency with other instruments.

During the lessons, there were several instances in which the instructor identified specific ornaments which regularly caused the student to lose track of the beat as a result of incorrect phrasing. In such cases, it was possible to locate these points using the display and compare the two recordings both visually and aurally, allowing the pupil to analyse their own playing with the instructor's comments (e.g. "you're labouring the *throw on D* [ornament]") in mind. At the conclusion of the study, the instructor described the system as "a great idea" and "such an interesting piece of kit", saying that he could definitely envisage it being used regularly in his lessons. Furthermore, he felt that the proportional notation provided an intuitive means of visualising rhythmic phrasing, and stated that even he had found it helpful in illustrating exactly how he played certain embellishments relative to the beat.

In addition to its use in lessons, the instructor suggested that the system could also prove to be a useful tool for solo practice. During lessons, he regularly asks his pupils to identify for themselves how their performance could be improved before providing feedback, so as to promote critical listening to their own playing when practicing alone. By using the system to identify problem areas and repeatedly comparing their recorded performance to the instructor's template, the student could keep track of their progress and avoid introducing bad habits between lessons. Moreover, the instructor indicated that one of the major challenges in teaching the bagpipes is to maintain the students' levels of enthusiasm. To this end, he described the system as "a fantastic thing to get people enthused".

The instructor also offered a number of suggestions for how the system might be improved. Some of these concerned minor refinements to enhance the physical playing experience, and were amended during the study. Criticisms pertaining more generally to the high-level affordances of

Likert statement	Mean response
I found the physical feel of the digital chanter was realistic to play.	4.2 / 5
I found the sound quality realistic.	4.0 / 5
I found the system easy to use.	4.4 / 5
I found the display easy to understand.	4.2 / 5
I found the system fun to use.	4.8 / 5
I think the system would be useful as a practice tool.	4.6 / 5
I would use the digital chanter system in my lessons and practice.	4.4 / 5

Table 1. Average student responses to survey questions.

the system will be addressed in future work. The instructor felt that by providing the facility for tutor and student to record simultaneously using separate electronic chanters, a more meaningful illustration of the differences in rhythmic phrasing might be obtained. He also suggested that the software could be used to indicate instances of *false fingering*, which refers to the practice of playing a note with the top hand (e.g. high A), without correctly executing the corresponding bottom hand fingering. While considered extremely bad practice in traditional Highland piping, this can be a difficult habit to diagnose and correct, particularly when playing quickly, as the resulting note is generally very close to the desired pitch.

5.3.2 Student Feedback

Following the sessions, students were given a short survey consisting of seven Likert-type questions and a box for additional comments, which was completed by five of the seven participants. Possible answers ranged from 1 ("strongly disagree") to 5 ("strongly agree"). Table 1 shows the mean response for each question. The numerical results indicate that the system was well received by the pupils. The lowest score was 4.0 for the perceived authenticity of the GHB sound, which may well be improved if the audio were to be reproduced through headphones or a loudspeaker system of reasonable quality rather than the built-in speakers on a laptop. The result of 4.6 in response to the statement "I think the system would be useful as a practice tool" is particularly encouraging, and the 4.8 reaction to "I found the system fun to use" is aligned with the instructor's assertion that it could help generate and maintain enthusiasm for practicing. The students also suggested several possible improvements to the system, which included a zoom feature to focus on specific sections, and making the operation of the software more intuitive for the user.

5.3.3 “Bagpipe Hero”

One significant addition to the system during the course of the study was the inclusion of a more game-like interface for the student to record along to the instructor’s performance, loosely inspired by the popular “Guitar Hero” series. This development was prompted by the instructor’s suggestion that providing a score for accuracy might help the students gauge their progress. In *Bagpipe Hero* configuration, the template notes approach a fixed marker on the staff which indicates the current note on the digital chanter.

The reaction from the students was decidedly positive; all stated that they would be more inclined to practice in their own time if the Bagpipe Hero system was set up in the school. In particular, when the instructor suggested that if a Bagpipe Hero leader board was set up each time a new tune was introduced to the pipe band, one pupil agreed that the motivation to practice would increase because “you’d want to beat everyone.” Other feedback included “This thing is cool”, and “it’s a lot of fun; I want one!”.

An interesting observation made by the instructor regarding the Bagpipe Hero system was that several pupils seemed to emulate the template performance significantly more accurately than with the original visualisation (playhead moving through stationary notation). One participant felt that it “helped quite a lot because you know how long to hold each note on”. Another student, a relative beginner who had been playing for around six months, struggled greatly when recording to the metronome, and rarely held the beat for more than a few bars before rushing into the next phrase. Subsequent investigation of the original recording indicated less than 11% accuracy. On the first attempt with Bagpipe Hero mode, the student achieved 65% for the same piece, which the instructor described as “an unbelievable difference”.

5.4 Conclusions

Feedback from this preliminary user study of the digital chanter system has been both positive and constructive. The instructor felt it had significant potential to be a valuable teaching tool in one-to-one piping lessons, and could also prove useful for solo practice. The study also yielded some promising avenues for further work (e.g. the capacity to highlight false fingering) and some important criticisms regarding the user interface design which will be addressed in the near future. The inclusion of some variant of the ornament recognition algorithm presented in [13] might also prove interesting, particularly for individual practice.

Acknowledgments

The authors would like to thank the instructor and students at the school for participating in the user study, and for their valuable feedback. This work was funded by the Engineering and Physical Sciences Research Council (EPSRC) as part of the Centre for Doctoral Training in Media and Arts Technology at Queen Mary University of London.

6. REFERENCES

- [1] J. Dickson, *The Highland Bagpipe: Music, History, Tradition*. Ashgate Publishing, Ltd., 2009.
- [2] T. Kirk and C. Leider, “The FrankenPipe: A Novel Bagpipe Controller,” in *Proc. NIME*, NY, USA, 2007.
- [3] C. Cannon, S. Hughes, and S. Ó Modhráin, “EpipE: Exploration of the Uilleann Pipes as a Potential Controller for Computer-based Music,” in *Proc. NIME*, Montreal, Canada, May 2003.
- [4] D. Menzies and D. Howard, “The CyberWhistle, An Instrument For Live Performance,” in *Colloquium on Musical Informatics XII*, Gorizia, Italy, Sept 1998.
- [5] G. P. Scavone, “THE PIPE: Explorations with Breath Control,” in *Proc. NIME*, Montreal, Canada, 2003.
- [6] R. B. Dannenberg, M. Sanchez, A. Joseph, R. Joseph, R. Saul, and P. Capell, “Results from the Piano Tutor Project,” in *Proc. Fourth Biennial Arts and Technology Symposium*, CT, USA, Mar 1993.
- [7] S. W. Smoliar, J. A. Waterworth, and P. R. Kellock, “pianoFORTE: A System for Piano Education Beyond Notation Literacy,” in *Proc. ACM MM*, San Francisco, CA, USA, Nov 1995.
- [8] S. Shirmohammadi, A. Khanafar, and G. Comeau, “MIDIATOR: A Tool for Analysing Students’ Piano Performance,” *Revue de Recherche en Éducation Musicale*, vol. 24, pp. 35–48, 2006.
- [9] A. Hadjakos, E. Aitenbichler, and M. Mühlhäuser, “SYSSOMO: A Pedagogical Tool for Analyzing Movement Variants Between Different Pianists,” in *Proc. Enactive Int. Conf.*, Pisa, Italy, Nov 2008.
- [10] K. Ng, T. Weyde, O. Larkin, K. Neubarth, T. Koerselman, and B. Ong, “3D Augmented Mirror: A Multimodal Interface for String Instrument Learning and Teaching with Gesture Support,” in *Proc. Multimodal Interfaces*, Nagoya, Japan, Nov 2007.
- [11] J. Yin, Y. Wang, and D. Hsu, “Digital Violin Tutor: An Integrated System for Beginning Violin Learners,” in *Proc. ACM MM*, Singapore, Nov 2005.
- [12] J. Mills and A. Murray, “Music technology inspected: good teaching in Key Stage 3,” *British Journal of Music Education*, vol. 17, no. 2, pp. 129–156, 2000.
- [13] D. W. H. Menzies and A. P. McPherson, “An Electronic Bagpipe Chanter for Automatic Recognition of Highland Piping Ornamentation,” in *Proc. NIME*, Ann Arbor, MI, USA, May 2012.
- [14] M. Wright and A. Freed, “Open Sound Control: A New Protocol for Communicating with Sound Synthesizers,” in *Proc. ICMC*, Thessaloniki, Greece, Sept 1997.
- [15] J. McCartney, “Rethinking the Computer Music Language: SuperCollider,” *Computer Music Journal*, vol. 26, no. 4, pp. 61–68, Winter 2002.
- [16] A. Williamon, *Musical Excellence: Strategies and Techniques to Enhance Performance*. Oxford University Press, Oxford, 2004.

An experimental study of temperature variations inside a clarinet

Daniel Noreland

Department of Computing Science, Umeå University, 901 87 Umeå, Sweden
noreland@cs.umu.se

ABSTRACT

A method for measuring the air column temperature of woodwind instruments using a medical infrared thermometer is devised and applied on a clarinet. It is found that the temperature varies roughly linearly along the bore, but with a considerable temperature swing according to playing conditions. If the temperature is approximated by an average along the bore, computations erroneously predict the octave to be compressed by 8.5 musical cents and 5.7 cents in the fundamental and second registers, respectively. The importance of considering humidity in accurate woodwind modelling is also pointed out.

1. INTRODUCTION

Every wind instrument player is aware of the influence of temperature effects in wind instruments. Tuning with the rest of the orchestra is not useful until the instrument has reached its "working temperature", and a cold instrument is not even "in tune with itself". Physical models that account for temperature effects with a high degree of accuracy have been long established [1], and it would be fair to say that this part of physics is not a major delimiting factor in modelling of wind instruments. A few articles [2–4] treat the effects of temperature gradients along the instrument and devise how to deal with them in a mathematical model. Knowing the temperature variation along the bore of an instrument, it is easy to account for the effect for instance in a transmission line model [5], either by introducing expressions like

$$\begin{aligned} c &= 347.23(1 + 0.00166\Delta T), \\ \rho &= 1.1769(1 - 0.00335\Delta T), \end{aligned} \quad (1)$$

where c is the speed of sound, ρ is the density and ΔT is the temperature deviation from 26.85 °C, or simply using average values of c and ρ along the bore as an approximation. That said, surprisingly little has been written specifically about what temperature one would expect to find along the bore of different wind instruments. An interest in this question was awakened in the course of work on a paper [6] on woodwind geometry optimisation — the success of tonehole optimisation depends ultimately on knowledge of the different conditions in the instrument. At the time

of writing the paper, the question about effects due to temperature gradients was only given a cursory treatment. The present paper is intended to cast some more light on the issue.

One reason for the scarcity in the literature might be that questions about the very air column temperature adds little to the understanding of the functioning of the instrument in general terms. Another reason might be that accurate measurement with a small time constant have been difficult to do in practice. Most temperature sensors are either intrusive due to their size, inaccurate, or suffer from a large time constant.

Gilbert et al. [4] use an infrared (IR) camera in order to register the temperature variation over the outer surface along a trombone. While this temperature is probably representative of the temperature of the inner surface, it is not identical with the local air column temperature; the air cools off on its passage through the instrument, but if there were no temperature difference between the air and the surface, no cooling would take place. There is, however, a large difference in airflow past the inner and the outer surfaces, so it is safe to assume that the surface temperature of the thin brass in a brasswind is more representative of the interior temperature than that of the surrounding air. Most brasswinds have significant cylindrical sections, and a more advanced model than a simple length correction that accounts for temperature gradients is probably unwarranted. It is only in the expanding sections that a more detailed model might be justified. In the case of woodwinds, the situation is more complicated. The rather thick wooden body of an instrument makes surface measurements of temperature infeasible as representatives of the inner temperature. Furthermore, the average speed of sound or air density in the bore is only a zero order approximation of what can be expected to be an essentially linear variation along the tonehole lattice. At any rate, only a careful measurement can assess the accuracy of such an approximation.

2. EXPERIMENTAL SET UP

The intention of the present paper was to measure the air column temperature inside a clarinet with an accuracy of less than 0.5 °C, and with a time constant not exceeding a few seconds. This would make possible to monitor the warming up procedure and, more interesting, to show whether an appreciable cooling takes place during pauses in the music or even as the fingering changes. A standard B♭ clarinet (Keilwerth) with a wooden body was chosen as the study object. As a means of measuring the temper-

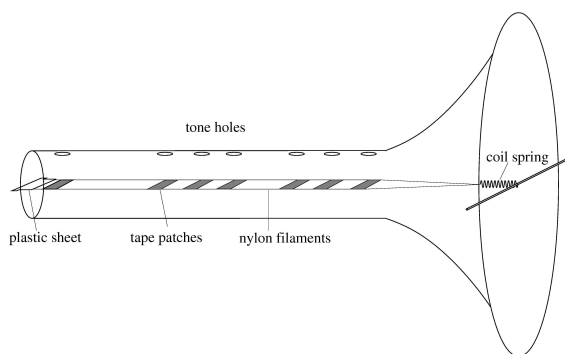


Figure 1. Schematic representation of the measurement set up. Two strands of nylon filament holding the measurement patches were taut between a plastic holder and a dowel at the bell exit.



Figure 2. The measurement patches mounted on the nylon filaments. Measurement positions are indicated by arrows.

ature, an IR sensing medical ear thermometer was used. The thermometer has a probe with a diameter of 8 mm and can be inserted projecting radially into the bore at some of the toneholes. In order to capture the air column temperature rather than the inner wall temperature, which may or may not be similar to the former, small measurement surfaces were placed floating coaxially underneath the fingerholes. This was arranged by stretching two nylon filaments along the bore between the barrel and the mouth of the bell. Patches of electrical insulation tape¹ were placed on the appropriate positions along the filaments. The arrangement is shown in Fig. 1 and 2. In order to provide a measurement point close to the mouthpiece, a hole was drilled in the barrel. This hole was normally kept closed by a plug and opened only for probe insertion. Subsequent positions were the holes of the upper joint covered by the left hand thumb, middle finger and ring finger, and the holes of the lower joint covered by the right hand index finger, middle finger, and ring finger. Two measurement points were also placed at the bottom F3 hole and the bell exit, respectively. In order to access the measurement patch at the F3 hole, it was necessary to unscrew the key.

3. EXPERIMENTS

Temperature measurements in wind instruments are difficult since they depend both on ambient temperature and the playing conditions. The ambient temperature during the experiments was comfort room temperature. It is not equally obvious how to choose representative playing conditions. Factors such as which register is played, at what

t (min)	0	5	10	20	25	30
T (°C)	20.8	22.5	22.9	23.4	24.3	24.2

Table 1. Outer surface temperature measured adjacent to the left ring finger hole during warm up.

dynamic level etc. vary considerably according to the music. Nevertheless, it was postulated that a reasonable playing "average" is represented by varied playing throughout the fundamental, second and third registers.

In the experiments presented, an able amateur musician played various classical passages. The introduction of the probe into the toneholes interferes somewhat with playing but can be done while playing continuously, and quickly enough (around one second including capturing of the temperature) to mimic common musical passages.

3.1 Calibration

Fundamental to IR based temperature measurements is knowledge or control of the emissivity of the measured surface. The tape used has an emissivity close to one, and also very close to that of the bore surface; Measurements on two adjacent spots of which one was covered with tape showed a difference of less than 0.1 °C, which is the resolution of the thermometer. According to the specification, the accuracy of the IR thermometer was ± 0.3 °C. The accuracy of the used thermometer together with the measurement patches was verified against a laboratory thermometer on whose sensor was placed a patch of the tape used.

3.2 Warm up period

In a first experiment, the instrument was played from ambient temperature (19.9 °C) until a steady state was reached. The temperature envelope measured at the ring finger hole of the upper joint is shown in Tab. 1. At least 30 minutes must be allowed before the instrument has reached a temperature equilibrium.

3.3 Measurements during ordinary playing conditions

As the clarinet had reached working temperature, two sets of internal measurements were recorded by an assistant repeatedly inserting the thermometer into the hole in the barrel, the seven finger holes and at the bell exit. First, a series of measurements of the inner bore temperature were made. At each position, 10 measurements were recorded. After the first set, the measurement patches were quickly inserted, and the process was repeated. The curves presented in Fig. 3 presents the results of the measurements. For comparison, the temperature measured on the outer surface of the clarinet body is also shown. In spite of the irregularities, the temperature curve for the air column is fairly well described by a simple first order approximation

$$T(x) \approx 31.0 - 15.8x, \quad (2)$$

where x is the distance from the upper rim of the barrel. Except at at the middle finger hole of the upper joint, the expression is accurate within 0.4 °C.

¹ ISOMA PT 3900, 0.18 mm thick PVC tape.

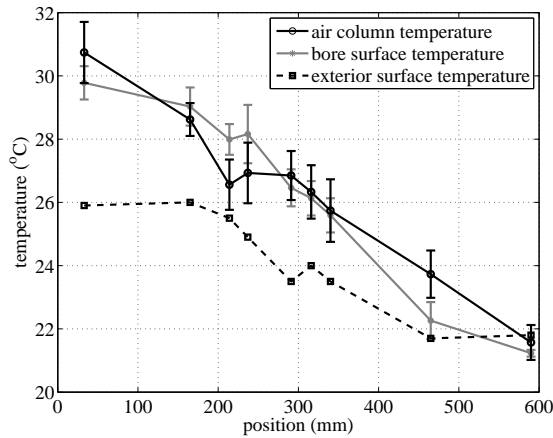


Figure 3. Temperature variation as a function of the distance from the upper rim of the barrel. The error bars show the standard deviations of the measurements. Ambient temperature: 19.9 °C.

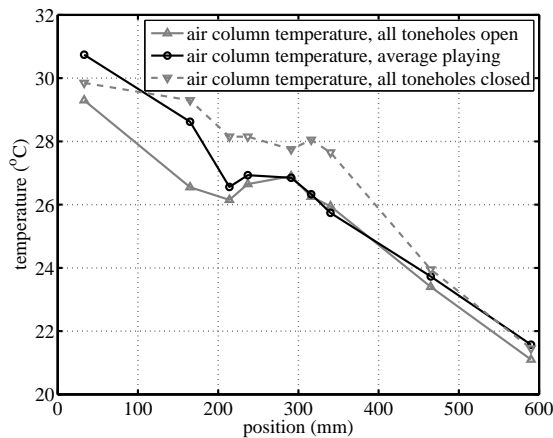


Figure 4. Temperature variation for the two extreme cases all toneholes open and all toneholes closed, respectively. Ambient temperature: 20.3 °C.

3.4 Effect of fingering

There is obviously a large variation in the number of toneholes open during musical performance. In order to estimate bounds for the temperature variation attributable to different fingerings, measurements were made playing the clarinet with all finger holes open (G4) and all holes closed (E3), respectively. Fig. 4 shows the results for continuously played *mf* notes. As expected, a faster temperature drop is seen with all holes open than with all holes closed. The difference diminishes further down the bore. It is however possible that a prolonged period of playing might have pronounced the difference between the situations due to additional downstream cooling or warming of the instrument. Except at the first two measurement positions, the temperature variation for average playing condition is very close to that seen with all holes open.

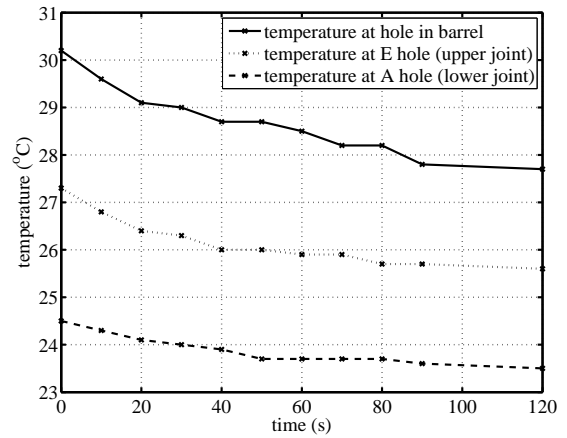


Figure 5. Cooling of air column during a pause in playing. Ambient temperature: 20.0 °C.

3.5 Cooling during a pause

In a last experiment, the air column temperature was monitored during a pause in playing. Fig. 5 shows the temperature evolution at three positions (hole in barrel, middle finger hole of upper joint and ring finger hole of lower joint) along the bore from the moment playing was stopped and the clarinet was put aside, the finger holes facing upwards, on a wooden table covered by a cloth. An appreciable cooling takes place over the course of one minute, but the temperature showed to be regained within seconds as playing was resumed. The cooling does not follow a simple exponential decay. After a rapid initial decrease, the cooling rate decreases. Conjecturally, this process is governed by the gradually decreasing thermal buoyancy forces on the air inside the bore as the warm air from the player is ventilated out and partly replaced by ambient air.

4. ACOUSTIC EFFECT OF A TEMPERATURE GRADIENT

The ultimate question regards to what extent the temperature gradient along the bore affects pitch, and how accurate results can be expected using a simple temperature average. This has been investigated using parts of the software employed by Noreland et al. [6] for woodwind optimisation. The program uses a transmission line analogy to compute the input impedance spectra for different fingerings of an instrument defined by its geometry. Allowing for specification of the temperature profile, various scenarios can be explored. The investigation was done for the geometry of a Buffet Crampon clarinet given by Debut et al [7]. If playing frequency is identified with the first input impedance peak for the fingering at hand, the octave interval of the notes E3 and E4 in the Chalumeau register is underestimated by 8.5 cents if the temperature curve is approximated by the average temperature along the bore from the reed to the bell end, rather than using the linear approximation of Eq. (2). For B4–B5 in the second register, the corresponding error is 5.7 cents.

5. DISCUSSION

Temperature variations along the bore occur mainly due to heat transfer at the air/bore interface and due to convective effects governed by bore ventilation. The relative importance of the effects can be expected to vary considerably according to the choice of material for the instrument, the tonehole geometry and the playing conditions. The results presented in this paper are likely not representative for clarinets made of plastic or metal. As expected, the measurements show a considerable variance of the air column temperature depending on playing conditions – somewhat larger than is the case on the inner surface of the bore that has a larger thermal inertia than the air column and the measurement patches. The variation showed to be strongly correlated to the playing level, since the air flow is the determining factor of the temperature at a given position.

The measurements show temperatures considerably lower than that of the human body temperature. At the barrel, the temperature has decreased to around 31 °C, depending on the playing level.

A number of observations can be made from Fig. 3. There seems to be a quite small difference in temperature between the bore surface and the air column. The difference in temperature between the inner and the outer surfaces is considerable, however. Interestingly, the measurements indicate that the surface temperature near the middle and ring finger holes of the upper joint is *higher* than in the air column. Although care was taken to perpetually clear out water from the bore using a brush, tiny water droplets or a water film may have time to build up and alter the emissivity whence the absolute readings of temperature on the bore surface are not conclusive. The error is estimated to be less than 0.5 °C, though. However, the presence of a water film reduces the emissivity, which would render a reading lower than the actual. This gives further support for the observation of the locally higher bore surface temperature. The musician holding the instrument transfers heat to the body of the instrument, but this is not the main explanation, as can be seen from the curve showing the temperature on the outer surface. It is reasonable to assume that the local dew point in the bore essentially equals the temperatures on the measurement patches and that condensation mainly takes place on the cooler bore surface. The assumption is supported by the observation that no condensation droplets were visible on the measurement patches, with one exception; There was a slight build up of water droplets on the patch adjacent to the left thumb hole. This hole is fitted with a brass tube projecting into the bore, and the air flow inside the bore creates a drop in the static pressure that sucks air into the thumb hole. This can easily be demonstrated qualitatively by forming a soap film on the rim of the brass tube and blowing in the mouthpiece. The air entering the bore is likely responsible for the decrease in air column temperature near the thumb hole. The soap film experiment indicated that air is less prone to enter finger holes without a brass tube, but more cannot be said in this matter without further study.

The curves of Fig. 4 and 5 indicate that, while it is important to keep the instrument warm during breaks etc.,

cooling conditioned by the fingering and short pauses is probably of subordinate importance.

The values in Sec. 4 for the pitch error due to the use of a temperature average in calculations are small, although significant. An average may be sufficient in some cases, but for precise modelling, a better description of the temperature gradient is motivated. However, it remains questionable if a more detailed model than a linear approximation like Eq. (2) is justified given the considerable temperature variation during normal playing conditions.

A question closely connected to the present study is the effect and variation of the content of CO₂ [8] and H₂O(g) in the bore. These gases are known to have a significant effect on the speed of sound. Assuming a 100% relative humidity (RH) in the air column, it is possible to include humidity effects in expressions for the local c and ρ as functions of the temperature only. Cramer [9] gives a model for c as a function of temperature, pressure, humidity and CO₂ concentration. Accordingly, the difference in the speed of sound between the two situations 0% RH and 100% RH, respectively, introduces a shift (speed of sound increases with humidity) corresponding to 7.1 musical cents at 22 °C, 9.6 cents at 27 °C and 12.7 cents at 32 °C. While a simple average of around 10 cents throughout a clarinet bore probably is sufficient in many cases, the effect of humidity is on the whole large enough to warrant consideration and should not be neglected in an accurate model for a wind instrument.

6. REFERENCES

- [1] D. Keefe, "Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximations for isothermal and nonisothermal boundary conditions," *J. Acoust. Soc. Am.*, vol. 75, pp. 58–62, 1984.
- [2] A. Benade, "Thermal perturbations in woodwind bores," *J. Acoust. Soc. Am.*, vol. 35, p. 1901, 1963.
- [3] J. Coltman, "Acoustical analysis of the boehm flute," *J. Acoust. Soc. Am.*, vol. 65, pp. 499–506, 1979.
- [4] J. Gilbert, L. L. Ruiz, and S. Goujeon, "Influence of the temperature on the tuning of a wind instrument (in french)," in *Congrès Français d'Acoustique*, Tours, 2006, pp. 599–602.
- [5] D. Keefe, "Woodwind air column models," *J. Acoust. Soc. Am.*, vol. 88, pp. 35–51, 1990.
- [6] D. Noreland, J. Kergomard, F. Laloë, C. Vergez, P. Guillemain, and A. Guilloteau, "The logical clarinet: numerical optimization of the geometry of woodwind instruments," Tech. Rep. arXiv:1209.3637, Mar 2013.
- [7] V. Debut, J. Kergomard, and F. Laloë, "Analysis and optimisation of the tuning of the twelfths for a clarinet resonator," *Applied acoustics*, vol. 66, pp. 365–409, 2005.
- [8] L. Fuks, "Prediction and measurements of exhaled air effects in the pitch of wind instruments," in *Proc. of IoA*, vol. 19, no. 5, 1997, pp. 373–378.

- [9] O. Cramer, "The variation of the specific heat ratio and the speed of sound in air with temperature, pressure, humidity, and CO₂ concentration," *J. Acoust. Soc. Am.*, vol. 93, pp. 2510–2516, 1993.

STUDY OF THE PERCEIVED QUALITY OF SAXOPHONE REEDS BY A PANEL OF MUSICIANS

Jean-François Petiot

Pierric Kersaudy

LUNAM Université, Ecole Centrale de Nantes
CIRMMT, Schulich School of Music, McGill

University

Petiot@irccyn.ec-nantes.fr

Pierric.Kersaudy@eleves.ec-nantes.fr

Gary Scavone

Stephen McAdams

CIRMMT, Schulich School of Music, McGill
University

gary@music.mcgill.ca

smc@music.mcgill.ca

ABSTRACT

The subjective quality of cane reeds used on saxophones or clarinets may be very different from one reed to another even though the reeds have the same shape and strength. The aim of this work is to study the differences in the subjective quality of reeds, assessed by a panel of musicians. The work focuses mainly on the agreement of the panel of musicians, the reliability of the evaluations and the discrimination power of the panel. A subjective study, involving 10 skilled musicians, was conducted on a set of 20 reeds of the same strength. Three descriptors were assessed: *Brightness*, *Softness*, and *Global quality*. The ratings of the musicians were analyzed using sensory data analysis methods to estimate the agreement between them and the main consensual differences between the reeds.

Results show that for *Softness* and *Brightness*, the agreement between the musicians is important and that significant differences between the reeds can be observed. For *Global quality*, the inter-individual differences are more important. The performance of the panel in providing reliable assessments opens the potential for an objectification of the perceived quality.

1. INTRODUCTION

For a saxophone player, the quality of a reed (a piece of cane that the player places against the mouthpiece) is fundamental and has big consequences on the quality of the sound produced by the instrument. The experience of saxophone players roughly shows that in a box of reeds, 30% are of good quality, 40% are of medium quality and 30% are of bad quality. The only indicator a musician can see on a box of reeds is the strength, which is usually measured by the maker by submitting a static force on a particular location from the tip. The reeds are then classified according to the strength measured. But this strength is not representative of the perceived quality of the reed. According to musicians, there are many differences among the reeds in a given box. But it is still difficult to understand which physical or chemical properties govern the perceived quality. The control of reed quality remains an important problem for reeds makers, because of the important variability of this natural material (*arundo donax*) and of the huge number of influencing factors. A thorough study of the perceived quality of reeds, and more generally of musical instruments, necessitates two

categories of measurements on a set of products: subjective assessments (given by musicians or listeners) [1] and objective measurements (chemical or physical), made on a set of instruments [2]. The principle is next to uncover (with statistical methods) a model for predicting subjective dimensions from the objective measurements.

In [3], optical measurements were used to assess the vibrational modes of clarinet reeds, which had been correlated with the quality of the reeds as judged by musicians. The authors suggested different patterns of vibrations that should be representative of good reeds.

In [4], B. Gazengel and J.P. Dalmont proposed two categories of physical measurements to explain the behavior of a tenor saxophone reed (in vivo during playing, and in vitro with a testing bench measuring the mechanical frequency response). Additional studies using these measurements showed that the perceived strength of a reed can be explained by the estimated threshold pressure in the musician's mouth, and that the perceived brightness correlates with the high-frequency content of the sounds [5, 6]. But these results were based on a small set of reeds (12) and used only one musician to assess their quality. They were limited to simple correlations between subjective variables and objective measurements and need to be confirmed.

The main difficulty in the study of the perceived quality of musical instruments is to get subjective assessments from musicians that are reliable and representative enough of the subtle interaction between the musician and the instrument. Many uncontrolled factors may influence this complex interaction. The subjective ratings of a "subject" may be non-reproducible, context-dependent, semantically ambiguous, and dependant on cultural and training aspects of the musician. To get representative data, it is necessary to find an acceptable trade-off between realistic playing conditions and artificial assessments of stimuli that could be oversimplified and then too caricatural. And to trust the data, it is necessary to control the assessments with repetitions and with several independent assessors. In this context, experimental protocols and data analysis techniques developed in sensory analysis can be very useful [7]. A number of statistical analysis methods are proposed to assess the evaluations of subjects and the panel's performance in descriptive analysis tasks [8].

In a previous paper [9], we defined a predictive model of tenor saxophone reed quality with PLS regression.

This model was based on a set of 20 reeds and a panel of 10 musicians.

This paper is the continuation of that work. It is centered particularly on the study of the performances of the panel of musicians. We propose to evaluate the inter-individual differences and to assess the reliability of the subjective assessments.

The paper is organized as follows: Section 2 presents the details of the experiment carried out with a set of reeds and a panel of musicians for the subjective study. Section 3 is dedicated to the presentation of the results of the subjective study. The agreement between the different assessments is presented. The last section presents the general conclusions and discusses the contribution of this study.

2. MATERIAL AND METHOD

2.1 Reed samples

The set of 20 reeds for tenor saxophone all had the same cut, strength and brand (Classic Vandoren, Strength 2.5). There was no preliminary selection of the reeds; they all came from 4 commercial boxes of 5 reeds each. The objective here is to estimate the perceived differences in 20 “similar” reeds.

Ten musicians participated in the subjective tests. They were all skilled saxophonists (students or professionals, with more than 10 years of practice). For the sake of consistency, all subjects used the same mouthpiece during the study (Vandoren V16 T7 Ebonite), however they were asked to play on their own tenor saxophone. These subjective tests took place at CIRMMT (Center for Interdisciplinary Research in Music Media and Technology) in Montreal, Quebec, Canada in May 2012.

2.2 Subjective evaluation of the reeds

In subjective tests, different semantic dimensions are generally defined to assess the differences between products [10]. For saxophone reeds, interviews of saxophonists have shown that the most frequent dimensions relate to “ease of emission”, “quality of sound”, or “homogeneity”. We proposed three subjective descriptors to assess the reeds:

- The *Brightness* of the sound produced with the reed,
- The *Softness* of the reed, which corresponds to the ease of producing a sound,
- The *Global quality* of the reed.

The test was divided into 3 phases: a training phase, an evaluation phase, and the filling out of a questionnaire concerning the mouthpiece, reed, saxophone and musical style the musicians usually play, as well as their past experience.

The training phase was proposed to help the subjects understand the meaning of the two descriptors *Softness* and *Brightness* and to verify their use of the scale. “Anchor reeds”, located at the extremes of the *Softness* scale, were proposed, and recorded sounds with different *brightnesses* were proposed. The method is inspired from the training phase described in [11]. Finally, subjects were asked to rate 3 quite different reeds on the interface,

to train them in the use of the scales and to verify their discrimination.

The evaluation phase used a graphical interface to assess the reeds. The musician was asked to play each reed and to assess each descriptor on an unstructured continuous scale (example in figure 1).

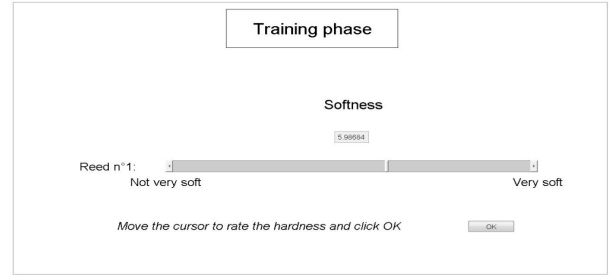


Figure 1. Continuous scale for the assessment of *Softness*

The reeds were presented to the subject in an order following a Williams Latin square in order to control the order and carry over effects. Given that we have 20 reeds and 10 subjects, the presentation plan was perfectly balanced. The assessments were repeated two times in two independent blocks. For each of the 10 subjects, the subjective data consists of 2 arrays of quantitative values (one per repetition). The arrays have 20 rows (one per reed) and 3 columns (one per descriptor).

The sensory panel consisted of $J=10$ assessors who judged $I=20$ products during $K=2$ sessions using $M=3$ attributes. The assessment of product i by assessor j during session k according to descriptor m is denoted Y_{ijk}^m .

3. RESULTS AND DISCUSSION

3.1 Individual performances of the assessors

This section focuses on the individual performances of the assessors, to whether the results of some subjects should be discarded. We use in this section the principles of the GRAPES method [12], which has been developed to assess the performances of a panel of experts in sensory analysis. It provides graphical representations of assessors' performances. We will focus on the different uses of the scale, the reliability of the subjects, their repeatability and their discrimination capacity.

3.1.1 Use of the scale

Two quantities can be computed to compare the use of scales by assessors. $LOCATION_j$ is the average of the scores given by assessor j (equation 1); $SPAN_j$ is the average standard deviation of a score given by assessor j within a session (equation 2). It represents the average magnitude used by the assessor to discriminate the products.

$$LOCATION_j = Y_{.j} \quad (1)$$

$$SPAN_j = \frac{1}{K} \sum_k \left[\frac{\sum_i (Y_{ijk} - Y_{.jk})^2}{(I-1)} \right]^{1/2} \quad (2)$$

N.B. We use a synthetic notation for the representation of the mean: considering the evaluation Y_{ijk} (see section

2.2), the notation $Y_{j.}$ means the mean of evaluations Y_{ijk} over the indices i (product) and k (session).

Figure 2 presents $SPAN_j$ vs $LOCATION_j$ for the different descriptors for subjects S1 to S10.

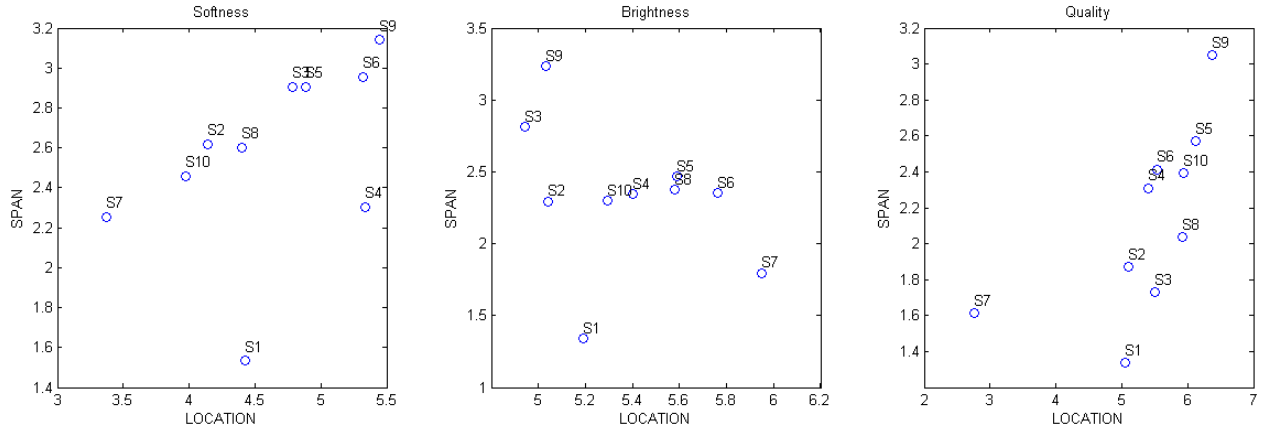


Figure 2. Plot of $SPAN_j$ vs $LOCATION_j$ for each subject and each descriptor.

The results show that subject S1 uses a small range for all the assessment (the SPAN is very small) and subject S7 globally dislikes all the reeds and assesses them as not soft ($LOCATION$ is low for this subject).

3.1.2 Reliability of the subjects and influence of the session

Two coefficients can be computed to assess the performance of each subject for each descriptor concerning their reliability and the influence of the different repetitions.

The unreliability ratio, labeled $UNRELIABILITY_j$, represents the measurement error of the subject, relative to the average magnitude used for the ratings. It is given by equation (3):

$$UNRELIABILITY_j = \frac{\left[\frac{1}{(I-1)(K-1)} \sum_{i,k} (Y_{ijk} - Y_{ij.} - Y_{.jk} + Y_{j.})^2 \right]^{1/2}}{SPAN_j} \quad (3)$$

The $DRIFT_MOOD_j$ (equation 4) is the between-sessions error relative to the average magnitude used for the ratings (expressed in SPAN units). It represents the deviation of the ratings of the subject across the sessions.

$$DRIFT_MOOD_j = \frac{\left[\frac{1}{K-1} \sum_k (Y_{.jk} - Y_{j.})^2 \right]^{1/2}}{SPAN_j} \quad (4)$$

Figure 3 represents, for each descriptor, the performance of the subjects according to $DRIFT_MOOD$ and $UNRELIABILITY$.

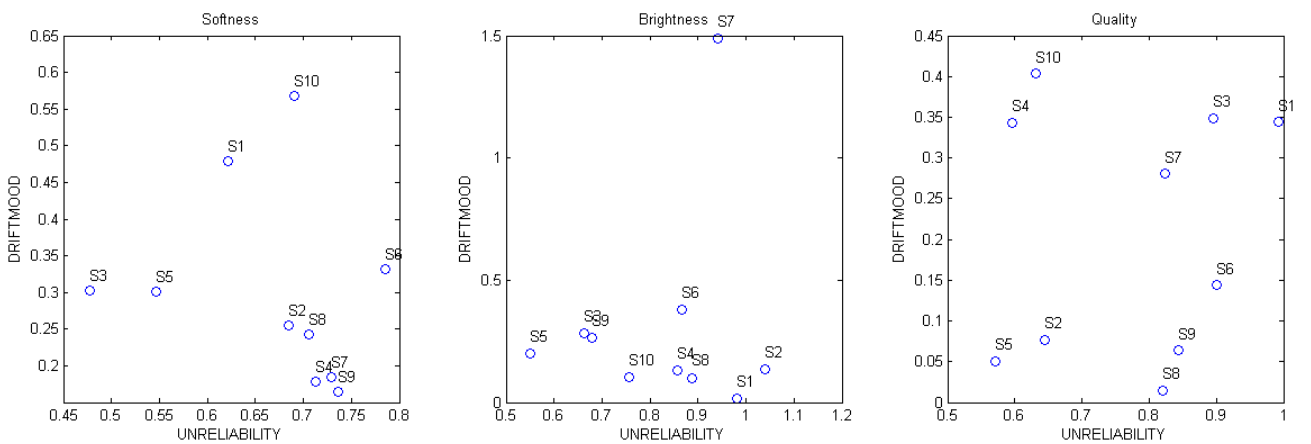


Figure 3. Plot of $DRIFT_MOOD_j$ vs $UNRELIABILITY_j$ for each subject and each descriptor.

For *Softness*, S6 is the least reliable and S3 and S5 are the most reliable. S10 deviates the most between the 2 sessions (high $DRIFT_MOOD$). For *Brightness*, S2 is the least reliable and S5 is the most reliable. S7 deviates the

most between the 2 sessions. For *Quality*, S1 is the least reliable and S5 is the most reliable.

We can conclude that S5 is a particularly reliable subject. We can also see that the worst value of unreliability for *Softness* is lower than most of the values for *Brightness*.

This means that most subjects (S6, S4, S8, S1, S2, S7) are less reliable for *Brightness* than for *Softness*. This result is in accordance with the feedback of the subjects during the tests, who indicated having more difficulty assessing *Brightness* than *Softness*.

These graphs are interesting to verify the quality of the individual assessments in order to detect possible unreliability or misunderstanding in the ratings. In our panel, no subject is particularly identified as unreliable in the assessment.

3.2 Global performance of the panel

3.2.1 Agreement between the assessors

The agreement between the assessors in their evaluation of the reeds can be estimated by consonance analysis, a method based on a principal component analysis (PCA) of the assessments. A description of this method can be found in [13]. To study the agreement for each descriptor (independent of the sessions), the repetitions are merged

vertically (repetitions are considered as different products). A standardized PCA is performed on the matrix $Y^m(2 I \times J)$ (equation 5):

$$Y^m = \begin{bmatrix} Y_1^m \\ Y_2^m \end{bmatrix} \quad (5)$$

A perfectly consensual panel would consist of assessors who rate the reeds in the same way. In this case, the first component of PCA would account for a very large variance. The more the panel is consensual, the more the arrows of the assessors point in the same direction. The percentage of the variance explained by the first principal component is considered as an indicator of the consonance of the panel. The results of the PCA of the matrices Y^m are given in figure 4 for each descriptor. In this PCA, the variables are the assessors (S1 to S10) and the individuals are the reeds.

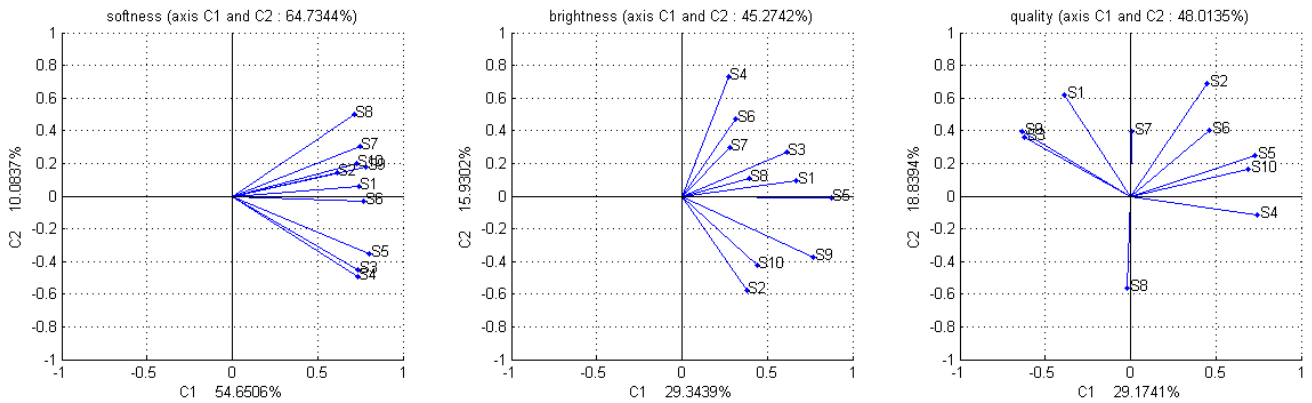


Figure 4. Consonance analysis for each descriptor: plot of the first two factors of the PCA (plane of the variables)

To evaluate more precisely the strength of the consensus for each descriptor, we can use indicators such as the Consonance C defined by equation 6 [13]:

$$C = \frac{\lambda_1}{\sum_{r=2}^J \lambda_r} \quad (6)$$

where J is the components number in the PCA (here the number of assessors), and λ_r is the r^{th} eigenvalue of the covariance matrix associated with the r^{th} component in the PCA. So this indicator emphasizes the weight of the first principal component and considers the higher dimensions as error or noise. It can be compared to a signal/noise ratio. We can also use the percentage of the total variance explained by the first principal component as an indicator to estimate the consonance of the panel. The consonance ratio C and the variance accounted for by the first factor are given in Table 1.

Descriptor	Consonance C	% Variance first PC
<i>Softness</i>	1.2	54.6%
<i>Brightness</i>	0.4	29.3%
<i>Global quality</i>	0.4	29.2%

Table 1. Results of consonance analysis for the panel of subjects.

The highest agreement is obtained for the descriptor *Softness*. The opinions of the assessors are convergent and the agreement is strong. For *Brightness*, the agreement is weaker, even though no assessor is very discordant.

For *Quality*, the agreement is the weakest. This is rather normal, given that *quality* is strongly related to the preference of the saxophonist, and that the tastes of the musician can be very diverse. Subjects S1, S3, and S9 are rather opposite to the rest of the panel; subject S8 is independent of the general trend according to preference. Given this result, we will have to analyze the *global quality* separately from the two other descriptors and for different groups of subjects.

3.2.2 Discrimination power of the panel

A general method to estimate the discrimination power and reproducibility of a panel of assessors is the Analysis of Variance (ANOVA). It is used in sensory analysis to study the differences between products and, more generally, to test the statistical significance of qualitative factors [14].

The assessment of the product i by assessor j during session k is denoted Y_{ijk} ($i=1$ to I , number of products, $j=1$ to J , number of assessors, $k=1$ to 2, number of sessions). A

model for the whole panel (equation 7) is proposed, taking into account the reed effect α_i , the session effect γ_k , and the reed*session interaction $\alpha\gamma_{ik}$:

$$Y_{ijk} = \mu + \alpha_i + \gamma_k + (\alpha\gamma)_{ik} + \epsilon_{ijk} \quad (7)$$

In this model, we don't introduce the subject effect because we consider that we don't have enough degrees of freedom to estimate correctly the contribution of the subject effect, the reed effect, the session effect and the associated interactions in the same model. As a matter of fact, the reed effect determines the discriminant power of the panel, and the reed*session interaction determines the repeatability of the panel. Consequently, the subject becomes a random variable in the model and gives us more analysis power. An ANOVA model is fit for each descriptor. The results of the ANOVA for the whole panel are given in Table 2.

Source of variation	p-value		
	<i>Softness</i>	<i>Brightness</i>	<i>Quality</i>
Reed	<0.001	<0.001	0.028
Session	<0.001	0.005	0.34 (n.s.)
Reed*Session	0.21 (n.s.)	0.88 (n.s.)	0.96 (n.s.)

Table 2. Results of ANOVAs for the three descriptors (p-value)

The reed effect is significant for all the descriptors ($p < 0.05$), which signifies that the panel discriminated the reeds well. The reed*session interaction is not significant for all the descriptors ($p > 0.05$), which means that there is no significant disagreement in the panel from one session to another. The session effect is significant for *Softness* and *Brightness*. It is a sign of a slight change in the use of the scale between the two sessions. Given that the reed effect is significant, we consider that the panel of assessors is discriminant/repeatable enough to aggregate the data in a consensual evaluation, representative of the reeds.

3.3 Subjective characterization of the reeds

3.3.1 Descriptive analysis

The mean value and the standard deviation of the assessments have been computed for each descriptor. The mean values are represented in figure 5 for *Brightness* and figure 6 for *Softness*.

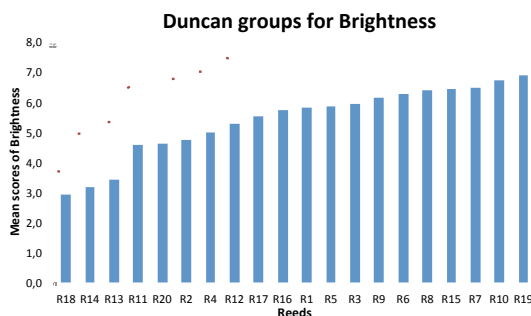


Figure 5. Mean value of *Brightness* and Duncan groups (multiple comparison test – $p = 5\%$)

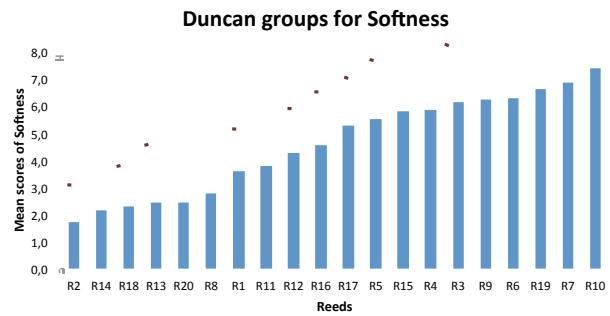


Figure 6. Mean value of *Softness* and Duncan groups (multiple comparison test – $p = 5\%$)

Significant differences between the reeds are evaluated by a Duncan multiple comparison test. Depending on the attributes, the Duncan multiple comparison test enables discrimination between 7 (*Brightness*) and 9 (*Softness*) non-overlapping groups of reeds. The Duncan groups (5% level) are represented by the pieces standing under the same horizontal. Figures 5 and 6 detail the differences between reeds that are significant for each attribute. The test confirms that the discrimination between the reeds is better for *Softness* than for *Brightness*. The average position of the reeds (R1 to R20) is given in Figure 7.

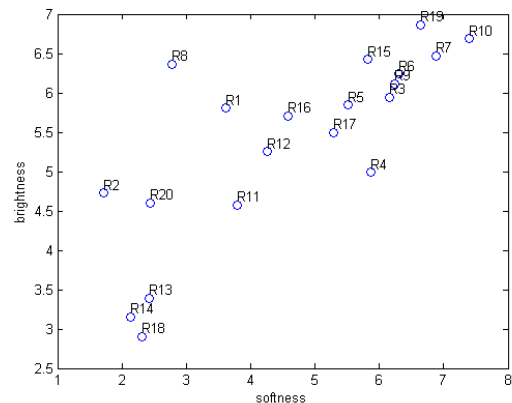


Figure 7. Position of the reeds according to *Softness* and *Brightness* (average configuration)

R10, R7, R19 are the most soft and bright reeds; R14, R18, R13 are the least soft and bright reeds. There is also a correlation between the two descriptors *Brightness* and *Softness*: a bright reed is also generally soft.

3.3.2 Analysis of the global quality

We showed in section 3.2 that the agreement between the assessors for the attribute *Quality* was relatively weak, and that discordant subjects should be considered. For these reasons, the subjects were partitioned according to quality. Let us consider the assessments of quality in the matrix Y^3 of dimension $(2I \times J)$, which considers the repetition as additional variables (variable = reed*session). A cluster analysis with Hierarchical Ascendant Classification has been made on the matrix Y^3 . We performed the cluster analysis on the row data (not centered nor

reduced) because we consider that the verbal anchoring of the scale gives a meaning to the scores and the mean. The distance used for the HAC is the Euclidian distance and the linkage rule is the Ward criterion (variance criterion). The dendrogram of the classification is presented in figure 8 (grouping of the subjects).

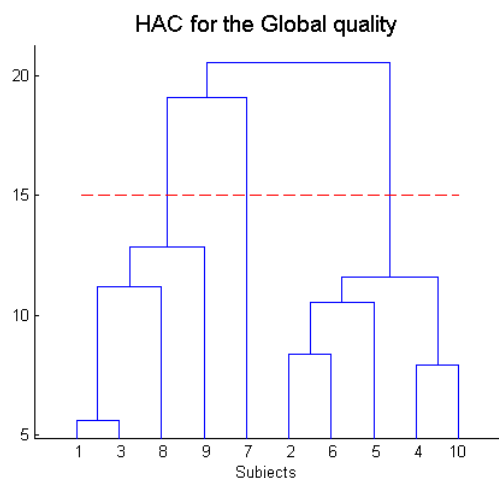


Figure 8: Dendrogram of the HAC according to the global quality ratings for the mean of the 2 sessions

3 clusters can be formed:

- Group1: S1 S3 S8 S9.
- Group2: S2 S6 S5 S4 S10.
- Group3: S7.

The average scores of reed quality for the two main groups 1 and 2 are given in figure 9.

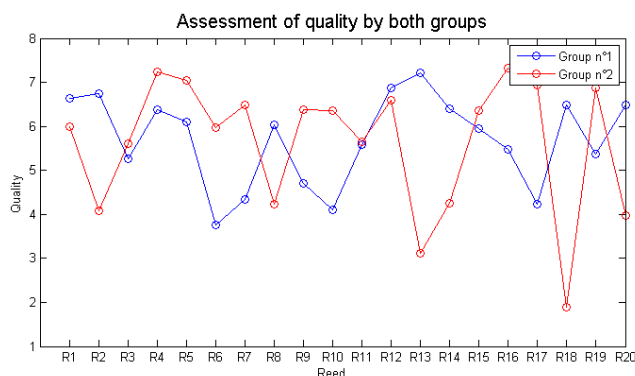


Figure 9: quality scores for the 2 different groups

Group 1 and 2 have mainly conflicting opinions on reeds R13 and R18 (most segmenting reeds). Group 1 (typical subject S3) appreciates R13 and R18, whereas Group 2 (typical subject S10) dislikes them.

We tried to characterize both groups with external information concerning the subjects, obtained from the questionnaires, but no feature of the musicians seems to clearly characterize the groups. However it seems that most of the musicians in group 1 play hard reeds and most of the musicians in group 2 play soft reeds. But we can't generalize this because of the small number of musicians we had. This seems logical, because the biggest differences we can see between the two groups are on the softest reeds or on the hardest reeds. For example we can see big differences for the reeds R2, R13 and R18, which are

perceived as the hardest reeds, and we also see big differences for the reeds R10 and R17, which are perceived as soft reeds.

4. CONCLUSIONS

This paper presented an analysis of the subjective assessments of a set of 20 saxophone reeds. Three descriptors were assessed by a panel of 10 musicians: *Softness*, *Brightness* and *Global Quality*.

The results show that the agreement between the subjects is more important for *Softness* than for *Brightness*. For these two descriptors, with the proposed task, the musicians were able to provide discriminant assessments and significant differences between the reeds are observed.

Differences between the musicians concerning the perceived quality necessitated the definition of subgroups of musicians. These differences are normal and due to the differences in personal tastes of the musician.

Future work will consist in using machine learning technique to model the subjective assessments by objective measurements.

Acknowledgments

The authors would like to thank the 10 musicians of the Schulich School of Music, McGill University, Montreal for their participation in the subjective tests, as well as Bruno Gazengel from LAUM for his advices.

5. REFERENCES

- [1] Pratt R.L., Bowsher J.M. "The subjective assessment of trombone quality". *Journal of Sound and Vibration* 57, 425-435 (1978).
- [2] Pratt R.L., Bowsher J.M. "The objective assessment of trombone quality". *Journal of Sound and Vibration* 65, 521-547 (1979).
- [3] F. Pinard, B. Laine, and H. Vach. Musical quality assessment of clarinet reeds using optical holography. *The Journal of the Acoustical Society of America*, 113:1736, 2003.
- [4] B. Gazengel and J. Dalmont, "Mechanical response characterization of saxophone reeds," in proceedings of Forum Acusticum, Aalborg, June-July 2011.
- [5] B. Gazengel, J.-F. Petiot and E. Brasseur, "Vers la définition d'indicateurs de qualité d'anches de saxophone," in proceedings of 10ème Congrès Français d'Acoustique, Lyon, April 2010
- [6] B. Gazengel, J.-F. Petiot and M. Soltes, "Objective and subjective characterization of saxophone reeds," in *proceedings of Acoustics 2012*, Nantes, april 2012.
- [7] Marjorie C. King, John Hall, and Margaret A. Cli . A comparison of methods for evaluating the performance

of a trained sensory panel. *Journal of Sensory Studies*, 16(6):567-582, 2001.

- [8] Zacharov N., Lorho G. What are the requirements of a listening panel for evaluating spatial audio quality? *Spatial audio & sensory evaluation techniques*, Guildford, UK, 2006, April 6-7.
- [9] Petiot J-F., Kersaudy P., Scavone G., McAdams S., Gazengel B. Modeling of the subjective quality of saxophone reeds. *Proceedings of ICA 2013*, June 2013, Montreal, Quebec, CANADA.
- [10] A. Nykänen, O. Johansson, J. Lundberg, J. Berg. Modelling Perceptual Dimensions of Saxophone Sounds. *Acta Acustica united with Acustica*, Volume 95, Number 3, May/June 2009, pp. 539-549 (11).
- [11] S. Droit-Volet, W. Meck and T. Penney, "Sensory modality and time perception in children and adults," *Behavioural Processes*, vol. 74, pp. 244-250, 2007.
- [12] P. Schlich, "GRAPES: A method and a SAS program for graphical representation of assessor performances," *Journal of sensory studies*, vol. 9, pp. 157-169, 1994.
- [13] G. Dijksterhuis, "Assessing Panel Consonance," *Food Quality and Preference*, vol. 6, pp. 7-14, 1995.
- [14] T. Couronne. "A study of Assessors' Performance Using Graphical Methods". *Food, Quality and Preference* Vol. 8, No. 5/6, pp. 359-365, 1997.

Numerical analysis of the mean flow effect on the sound directivity pattern of cylindrical ducts

Yong Shi

Centre for Interdisciplinary Research
in Music Media and Technology,
McGill University,
Canada
yong.shi2@mail.mcgill.ca

Andrey da Silva

Centro de Engenharia da Mobilidade,
Federal University of Santa Catarina,
Brazil
andrey.rs@ufsc.br

Gary Scavone

Centre for Interdisciplinary Research
in Music Media and Technology,
McGill University,
Canada
gary@music.mcgill.ca

ABSTRACT

This paper presents a numerical investigation of the sound directivity pattern of normal modes radiated from the open end of a cylindrical pipe. A good agreement is found between the numerical results and the analytical predictions of the directivity pattern for an unflanged pipe in the absence and in the presence of a low Mach-number mean flow. The investigations are conducted by using an axisymmetric two-dimensional lattice Boltzmann model. The numerical model is first validated by comparing its directivity with the established analytical model by Levine and Schwinger [1] for the case of zero mean flow. Then the numerical results under the condition of mean flow with two different Mach numbers are compared with the analytical model by Gabard and Astley [2] and recent experimental observations by Gorazd et al [3]. The effects of the so-called zone of relative silence are observed in the results even for very low Mach number ($M = 0.036$).

1. INTRODUCTION

The directivity characteristics of the acoustic wave radiating from the open end of cylindrical ducts have been investigated by many researchers over the last century. Much attention has been focused on engineering cases, such as exhaust pipes, ventilation systems, air-conditioning systems, etc. For musical instrument makers and researchers of woodwind musical instruments, the direction-dependent sound radiation characteristics of the external sound field are also of great interest. This problem can be investigated by analytical, experimental, and numerical approaches.

Levine and Schwinger [1] proposed an analytical model for the dominant mode propagation of sound out of a semi-infinite circular duct in the absence of flow. Their solution is based on the Wiener-Hopf technique, and gives rigorous and explicit results including the reflection coefficient R , the length correction l/a and the angular distribution of the emitted sound radiation described as the power gain function $G(\theta)$.

In the so-called outlet problems, the sound propagation out of a pipe carrying a non-zero mean flow is more complex due to the interaction between the sound field and the fluid field. Assuming a uniform flow, a thin vortex sheet separating the jet and the surrounding fluid and a full Kutta condition¹ at the edges of the open end, Munt [4] proposed an expression for the far-field sound radiation for the range of $0 \leq ka \leq 1.5$ and $M < 0.3$, which also uses the Wiener-Hopf technique and can be seen as an extension of Levine and Schwinger's model. This solution was elaborated by Rienstra [5] who introduced a complex parameter to take into account the effects of unsteady shedding of vorticity in the vicinity of the trailing edge, with particular attention to the energy balance of the sound field and the fluid field. Savkar [6] also presented an approximate model for the sound radiation from a semi-infinite circular duct by using a Wiener-Hopf analysis solved by an approximate method. Based on the work of both Munt and Rienstra, Gabard and Astley [2] presented an extended model that includes a center body for the cases of annular pipe and proposed an explicit numerical procedure for evaluating the solutions for higher frequencies (ka in the range of $0 - 60$ and higher Mach numbers (in the range of $0 - 0.8$). The pressure directivity of any single mode or all modes combined together can be derived from each of the analytical models cited above.

The analytical models are useful for problems involving simple geometries and as benchmark solutions for numerical simulations and experimental measurements. In realistic situations, however, the geometric characteristics of outlet systems, such as tailpipes and woodwind instruments, are not limited to simple geometries such as the unflanged circular duct. For cases of complex geometries, numerical approaches are more suitable.

In a recent example, Hornikx et al. [7] presented a numerical solution for calculating the radiation sound field emanating from an automotive exhaust pipe above a rigid surface. Other examples of numerical methods for problems of sound radiation can be found in [8], [9] and [10], to name just a few.

For the particular problems to be investigated in this paper, a numerical approach in the time domain called the lattice Boltzmann method (LBM) is used. The LBM is very different from the traditional continuum-based techniques

¹ The vortex layer leaves the edge of the cylinder with zero gradient.

in that it directly simulates the propagation and collision involving the space-time evolution of the fluid particles in a mesoscopic level rather than the pressure and velocity in a macroscopic level. Here the particles are not real atoms or molecules but rather velocity distribution functions representing the probability of finding a population of fluid molecules in a certain phase space.

The dynamic behavior of the populations are governed by a space-temporal discretization of the Boltzmann equation, known as the lattice Boltzmann equation (LBE). The macroscopic fluid properties, such as density, momentum, internal energy, and energy flux, can be recovered from the moments of the velocity distribution functions. Since the Boltzmann equation describes the physical phenomena of a fluid at a lower level and a smaller time scale, the Navier-Stokes equations can be fully recovered from the Boltzmann equation. The numerical solution of the Boltzmann equation is relatively simple compared to the Navier-Stokes equations, and can be implemented in a parallel computation scheme. This is advantageous for simulations of problems featuring complicated boundary conditions and multiphase interfaces.

The LBM has been extensively used in hydrodynamic problems, but only a few researchers have used the LBM in the field of wind musical instruments and musical acoustics. Skordos [11] first simulated the interaction between fluid flow and the acoustic waves within a flute-like instrument at different blowing speeds by using a two-dimensional lattice Boltzmann model.

In a series of simulations, da Silva et al [12–15] investigated the sound radiation at the open end of axisymmetric cylindrical pipes in terms of reflection coefficient, length correction and directivity factor. In [12], the sound radiation of the unflanged cylinder pipe immersed in a stagnant fluid is simulated. The simulation results in terms of $|R|$, l/a and radiation directivity agree well with the analytical model for inviscid wave propagation by Levine and Schwinger [1]. In [14], they investigated the influence of a cold subsonic mean flow of low Mach number in the stagnant fluid, as well as the effects of circular horns of different radii of curvature at the open end of the cylindrical pipe. Their simulation correctly demonstrates that a free laminar jet is formed downstream from the open end due to discharge of the upstream mean flow into the radiation area. For the case of an unflanged pipe, the simulation data in terms of $|R|$ and l/a agrees well with the analytical solution proposed by Munt [16] and the experimental data obtained by Allam and Abom [17], where the maximum value of the reflection coefficient is greater than one due to the energy exchange between the vortex sheet and the acoustic field. For the case of circular horns, their simulation data is in accordance with the experiments carried out by Peters et al. [18]. In a similar simulation [15], they investigated a more realistic case of the effects of mean flow corresponding to different dynamic playing levels for a cylindrical pipe terminated by a catenoidal horn. The series investigations conducted by da Silva et al. demonstrate that the lattice Boltzmann technique is a reliable numerical tool for investigating musical acoustics problems that in-

volve a radiating waveguide and a low Mach number fluid flow.

The objective of this paper is to investigate the pressure directivity of sound waves as they propagate into the far field from the open end of a cylindrical duct issuing a subsonic cold mean flow. The investigations are conducted by using the lattice Boltzmann method to represent a two-dimensional radiation domain. This paper is structured as follows: Section 2 describes the numerical technique used in the study. In Section 3, the simulation results are compared to the analytical model by Levine and Schwinger and the experimental observations by Gorazd et al. for the case of no flow, and to the analytical model by Gabard and Astley and the experimental observations by Gorazd et al. for the case of low Mach number mean flow. The phenomenon of the zone of relative silence is discussed for two different Mach numbers. Finally, Section 4 provides a discussion of the results and suggestions for further investigations.

2. NUMERICAL PROCEDURE

The purpose of the LBM scheme presented here is to reproduce the sound radiation at the open end of a cylindrical duct in the presence of a low Mach subsonic cold mean flow. General descriptions of the lattice Boltzmann method can be found in books by Succi [19] and Gladrow [20].

The LBM scheme is described by an axisymmetric cylinder structure immersed in a fluid domain surrounded by open boundaries, as illustrated in Fig. 1. The fluid domain defined by an axisymmetric half-plane is represented by a rectangular D2Q9 structure [21] of 1000 by 500 lattice cells. The left, top and right boundaries of the radiation domain are implemented by absorbing boundary conditions prescribed with a zero velocity, as proposed by Kam et al. [22]. The boundary representing the axis of symmetry of the system, along the bottom, is implemented by a free-slip boundary condition. The axisymmetric nature of the system is recovered by using an axisymmetric source term [23, 24].

The length and the radius of the cylindrical waveguide is $L = 469.5$ and $a = 10$ in lattice cells, respectively. The walls of the waveguide are represented by a boundary of zero thickness [25, 26]. The outer walls are treated by a simple bounce-back scheme [19] for which the viscous boundary phenomena are represented with second-order accuracy, while the inner walls are treated using a free-slip scheme in order to reduce the viscous boundary layer effects that result in a transfer of momentum by the tangential motion of particles along the walls. To ensure the numerical stability and to make the viscosity as small as possible, the relaxation time is set to $\tau = 0.5714$, which is equivalent to a dimensionless kinematic viscosity of $\nu = 0.0238$.² The undisturbed fluid density was set as $\rho_0 = 1.0 \text{ kg/m}^3$ for convenience.

The system is excited by a source signal that consists of a linear chirp signal running from $ka = 0.1$ to $ka = 3.8$ (up

² The physical kinematic viscosity is related to the dimensionless kinematic viscosity by $\nu^* = \frac{\nu c_s^* D_x}{c_s}$, where c_s^* is the physical speed of the sound, c_s is the speed of sound in lattice unit, D_x is the ratio of the physical length and the lattice length.

to the cut-on frequency of the first symmetric non-planar axial mode) superimposed on a DC offset representing the non-zero mean flow. The excitation is implemented by a source buffer at the left end of the pipe using absorbing boundary conditions with a non-zero target velocity prescribed by the source signal. Before the acoustic source is superimposed, there should be enough initialization time to allow the fluid in the whole domain to accelerate from stagnation to a stationary state. The minimum initialization steps can be estimated by

$$N_t \sim N_{t0} + L_x / (M * c_s),$$

where $N_{t0} = 4000$ is the acceleration time for the source buffer with thickness equivalent to 60 cells [14], $c_s = 1/\sqrt{3}$ is the speed of sound in lattice units, M is the Mach number of the non-zero mean flow and $L_x = 1000$ is the maximum traveling distance of the plane sound wave in the radiation domain in this system. For example, the minimum initialization steps corresponding to the Mach number of $M = 0.036$ is $N_t = 5.21 \times 10^4$. The highest Mach number used in this paper is $M = 0.15$, which makes the flow slightly compressible. However, the numerical model is still valid because the slightly unsteady compressible form of the Navier-Stokes equations can be fully recovered from the isothermal form of the Boltzmann equation by performing the Chapman-Enskog expansion, as described in [21] and [20].

The time histories of fluid density are probed at 75 points evenly distributed around the semi-circle (corresponding to angle increments of 2 degrees), with the center point at the outlet of the duct in the range of $\theta = 0^\circ$ to $\theta = 150^\circ$. The measuring distance is $d = 250$ cells from the outlet. The acoustic pressure p is calculated by

$$p(\theta, t) = (\rho(\theta, t) - \rho_0)c_s^2, \quad (1)$$

where $\rho(\theta, t)$ is the spontaneous fluid density and ρ_0 is the equilibrium density. For the case of zero mean flow, ρ_0 is nearly a constant and usually has the value of 1. For non-zero mean flow, however, ρ_0 in the vicinity of the probing points fluctuates over time and the fluctuating density can not be calculated by simply subtracting 1 from the spontaneous fluid density. For such a case, a zero-phase DC-blocking filter can be used to remove the offset caused by the flow.

Once the time history of acoustic pressures has been obtained, the pressure directivity can be calculated by

$$G(\theta, f) = \frac{P(\theta, f)}{P_h}, \quad (2)$$

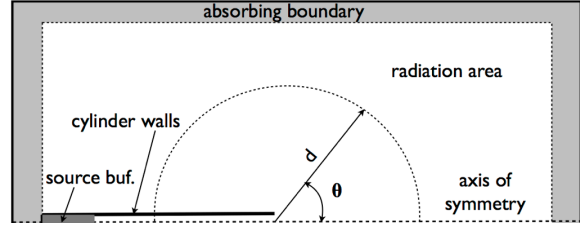
where $P(\theta, f)$ is obtained by performing a DFT on the time history of sound pressure $p(\theta, t)$ measured at the same distance d , $P_h = \sqrt{\sum P^2(\theta)/N}$ is the square root of the averaged value of $P^2(\theta, f)$ over all the measured angles.

3. RESULTS

3.1 Directivity in the absence of mean flow

The LBM scheme in the absence of mean flow is first validated by comparing its results with the established analytical model proposed by Levine and Schwinger [1] in

Figure 1. LBM scheme



the form of relative pressure directivity. For six different frequencies expressed in terms of the Helmholtz number ($ka = 0.48, 1, 2, 2.5, 3, 3.5$) that are below the cut-on frequencies of higher-order modes, the numerical simulations are in good agreement with the analytical results, as shown in Fig. 2. The tiny ripples found for $ka = 0.48$ and $ka = 1$ in the numerical results can be explained by the fact that $G(\theta)$ should be measured in the far-field condition, which is not fully satisfied for low frequencies given the size of the lattice (1000×500 cells) and the measuring radius (250 cells) used in this paper (due to computation time limits). Not surprisingly, the results for higher frequencies ($ka \geq 2$) are smooth and the ripples are barely observed. To evaluate the far-field condition in this simulation, we measured the acoustic impedance Z as a function of ka at a distance $d = 250$ and angle $\phi = 0$ from the outlet of the pipe. As depicted in Fig. 3(a), the amplitude of the impedance Z quickly converges to the characteristic impedance of the medium, $Z_c = \rho_0 c_s^2$, for values of $ka \geq 1$. A similar phenomenon can be found for the phase between the acoustic pressure and particle velocity, ϕ , which gradually converges to zero for $ka \geq 2$, as depicted in Fig. 3(b). The results suggest that the far-field condition is not fully satisfied for $ka < 1$, while for $ka \geq 1$, the acoustic impedance Z of the spherical wave propagating into the radiation domain approximates that of a plane wave.

From Fig. 2(f), we can observe smoothing of directivity characteristics of numerical results in the vicinity of $\theta = 100^\circ$ compared to that of the analytical results for the high frequency $ka = 3.5$. That might be attributed to the issue that, in the numerical simulation, there may be some transfer of energy from the exciting chirp signal to higher-order modes, while for the case of the analytical model, no higher modes are involved and the energies are exclusively coming from the dominant plane mode. A similar phenomenon was reported in a recent experimental measurement conducted by Gorazd et al. [3], where the curves presenting the directivity characteristics of the experimental results (excited by broadband noise) around $\theta = 100^\circ$ and for higher frequencies ($ka \geq 2.96$) are smoothed comparing to those analytical results obtained for a single-frequency exciting signal.

In the next step, the numerical and analytical results are compared with the experimental results by Gorazd et al. [3] in the form of relative pressure directivity. The experiment involved an unflanged radiating waveguide set-up with free field conditions in an anechoic chamber, which is comparable to the solid pipe wall and absorbing boundary con-

dition used in the numerical simulation. All three results (numerical, analytical and experimental) have been normalized to the same dB level, as depicted in Fig. 4. For the two lower frequencies of $ka = 0.74$ and 1.48 and for angles within the range of $0^\circ < \theta < 90^\circ$, the three results are in good agreement with each other, despite the fact that the measurements are carried out using 1/3 octave broadband noise and the calculation of numerical and analytical results are based on a single frequency. As the angle increases, the measurements are still in good agreement with the analytical results, though the numerical results have discrepancies less than 3 dB compared to the analytical results. For the higher frequency of $ka = 2.96$, the numerical results are in good agreement with both the analytical results and the measurements for angles within the range of $0^\circ < \theta < 75^\circ$. As the angle increases from 75° to 150° , both the measurements and the numerical results deviate from the analytical results, but in opposite ways. Compared to the analytical results, the highest discrepancies are found at the largest angle of $\theta = 150^\circ$, which is $+3.8$ dB for the measurements and -2.6 dB for the numerical results, respectively.

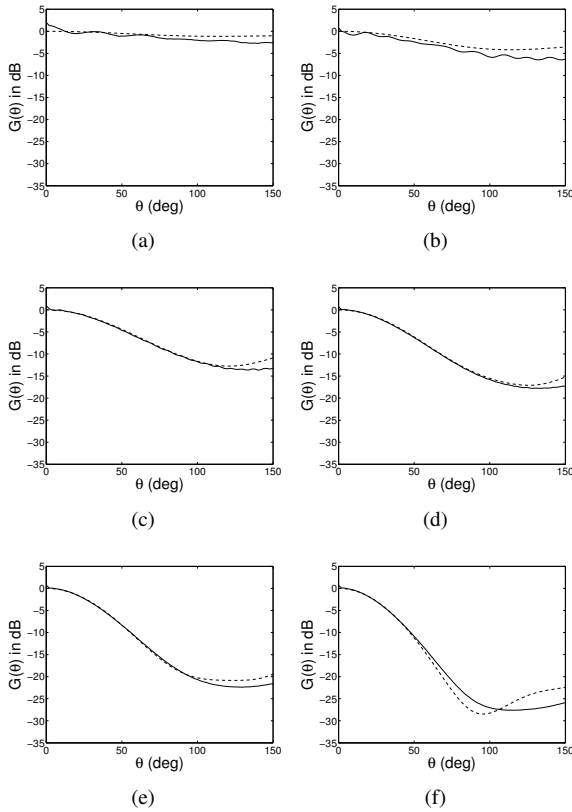


Figure 2. Comparison between numerical (solid) and analytical predictions [1] (---) of the acoustic pressure directivity as a function of the angle in the absence of a mean flow: (a) $ka = 0.48$, (b) $ka = 1$, (c) $ka = 2$, (d) $ka = 2.5$, (e) $ka = 3$, (f) $ka = 3.5$.

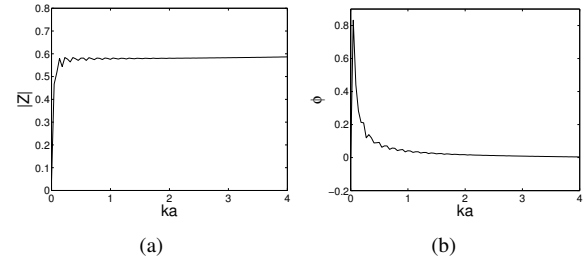


Figure 3. Evaluation of the far-field condition in terms of acoustic impedance in the radiation domain: (a) amplitude of acoustic impedance, (b) phase of acoustic impedance.

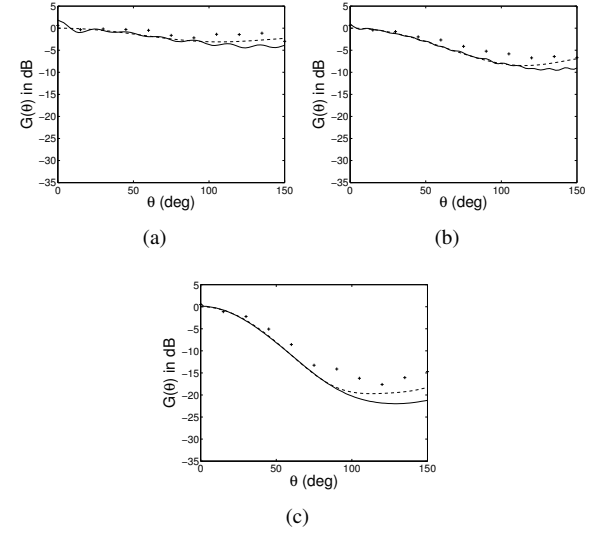


Figure 4. Comparison between numerical (solid), analytical predictions [1] (---) and experimental measurements [3] (+++++) of the acoustic pressure directivity as a function of the angle in the absence of a mean flow: (a) $ka = 0.74$, (b) $ka = 1.48$, (c) $ka = 2.96$.

3.2 Directivity in the presence of mean flow

For the case of a cold mean flow with a low Mach number ($M = 0.036$), the numerical results are compared with the theoretical prediction given by Gabard and Astley [2] as well as the recent experimental results obtained by Gorazd et al. [3] in the form of normalized pressure directivity, as depicted in Fig. 5. All three results (numerical, analytical and experimental) have been represented in the form of pressure directivity and normalized to the same dB level.

In general, the results are in good agreement for angles in the range $0^\circ < \theta < 60^\circ$. Discrepancies between the numerical and analytical results become more obvious as the angle increases and the maximum differences are found to be at $\theta = 150^\circ$, i.e., -3.11 dB for $ka = 0.74$, -2.22 dB for $ka = 1.48$ and -2.3 dB for $ka = 2.96$, respectively. For all three frequencies and for most angles, the analytical solution is located between the numerical and the experimental results.

For the case of a cold mean flow with a higher Mach number ($M = 0.15$), the numerical results are compared with the theoretical prediction only, since no experimental re-

sults are available from Gorazd et al. for $M = 0.15$. The comparisons are depicted in Fig. 6. In general, good agreement is found for angles in the range $30^\circ < \theta < 150^\circ$. For most angles, the discrepancy from the theory is less than 3dB. The deviation of the simulation from the theoretical results is mainly found in the region of angles less than 30° . The smoothing of the curve representing the numerical results versus the analytical results in the region $90^\circ < \theta < 120^\circ$ for the high frequency of $ka = 3.77$, as depicted in Fig. 6(d), might be due to the transfer of energy from the exciting chirp signal to higher-order modes, as discussed before.

An important feature of the directivity characteristics in the presence of a non-zero mean flow concerns the so-called “zone of relative silence”, where the sound wave in the vicinity of the axis is subject to additional attenuation. The result from the theoretical analysis of Savkar [6] and Munt [4] suggests that, for high frequencies and large Mach numbers, the zone of relative silence is so obvious that a cusp can be observed at $\theta = \theta_s$ in the directivity pattern. Assuming that the medium outside the duct is stagnant and the speed of sound remains constant, the zone of relative silence is defined by [6]

$$\theta_s = \cos^{-1} \left(\frac{1}{1 + M} \right), \quad (3)$$

where M is the Mach number inside the duct.

Even for the low Mach number $M = 0.036$, the zone of relative silence ($\theta_s = 15.15^\circ$) can be observed in both the experiments and the numerical results for $ka = 2.96$, as depicted in Fig. 5(c). For the case of higher Mach number $M = 0.15$, the zone of relative silence ($\theta_s = 29.59^\circ$) is more obviously observed in the numerical results for all four frequencies ($ka = 0.74, 1.48, 2.96$ and 3.77).

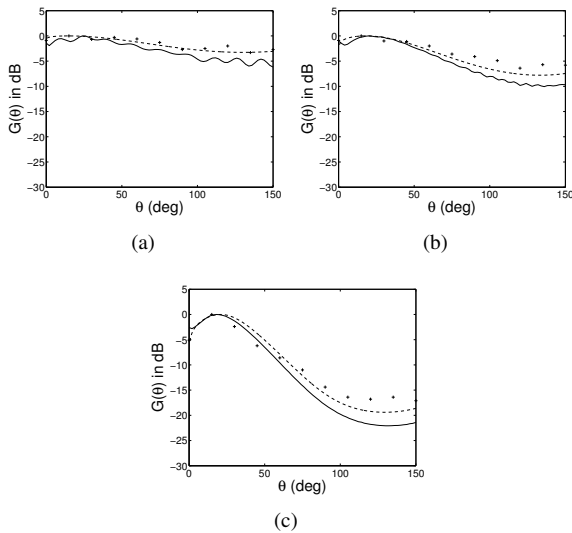


Figure 5. Comparison between numerical (solid), analytical [2] (---) predictions and experimental measurements [3] (+ + +) of the acoustic pressure directivity as a function of the angle in the presence of a mean flow at Mach = 0.036: (a) $ka = 0.74$, (b) $ka = 1.48$, (c) $ka = 2.96$.

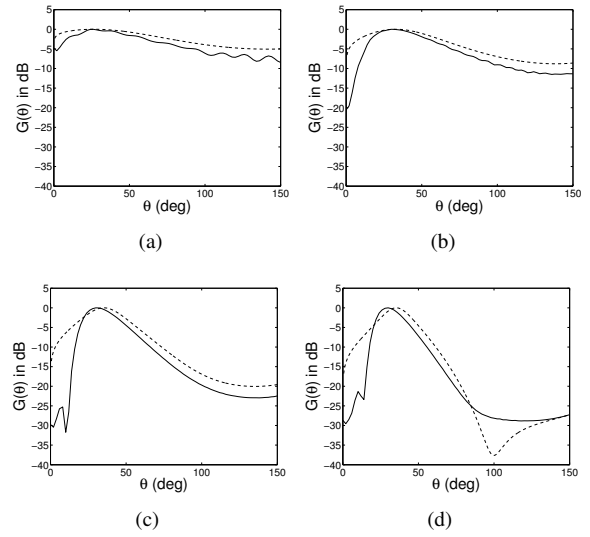


Figure 6. Comparison between numerical (solid) and analytical predictions [2] (---) of the acoustic pressure directivity as a function of the angle in the presence of a mean flow at Mach = 0.15: (a) $ka = 0.74$, (b) $ka = 1.48$, (c) $ka = 2.96$, (d) $ka = 3.77$.

4. CONCLUSIONS

In this paper, we presented a numerical technique based on an axisymmetric two-dimensional lattice Boltzmann scheme to predict the directivity pattern associated with the sound radiation at the open end of cylindrical ducts issuing a low Mach number cold subsonic jet into a stagnant fluid region.

The LBM scheme was first validated by comparing its results with the analytical model of Levine and Schwinger and experimental results of Gorazd et al. for the case of no flow. Then for the case of non-zero mean flow, the numerical results were compared with the theoretical prediction given by Gabard and Astley for Mach number $M = 0.036$ and $M = 0.15$ as well as experimental results obtained by Gorazd et al. for Mach number $M = 0.036$. Very good agreement was found with theoretical and experimental results for the case of no flow and the lower Mach number of $M = 0.036$. For the relatively higher Mach number of $M = 0.15$, the numerical result agrees very well with the theoretical prediction for angles greater than 30° , though significant discrepancies are observed for angles less than 30° . The effects of the so-called zone of relative silence are clearly observed in the results of non-zero mean flow even for very low Mach number ($M=0.036$). This is interesting for the studies of the sound radiation of woodwind instruments, which normally exhibit a very low Mach number flow.

The aforementioned discrepancies for the case of $\theta < 30^\circ$ and $M = 0.15$ are not well explained yet. For further investigations conducted by either experimental measurements or numerical simulations, some facts might be considered. The theoretical model assumes an infinitely thin vortex sheet separating the jet and the neighboring quiescent fluid, which is not true for far field in real situations as well as the numerical simulations presented here. In ad-

dition, it was found in the numerical simulation presented here that the directivity pattern in directions close to the axis is very sensitive to the probing distance.

Acknowledgments

The authors would like to thank Dr. Gwenael Gabard at University of Southampton, UK, for providing the numerical calculations of the theoretical model, and Mr. Lukasz Gorazd and Dr. Anna Snakowska at AGH University of Science and Technology, Poland, for providing the experimental data. We also wish to acknowledge funding from the Fonds québécois de la recherche sur la nature et les technologies (FQRNT), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Centre for Interdisciplinary Research in Music Media and Technology.

5. REFERENCES

- [1] H. Levine and J. Schwinger, "On the radiation of sound from an unflanged circular pipe," *Physical Review*, vol. 73, no. 4, pp. 383–406, 1948.
- [2] G. Gabard and R. J. Astley, "Theoretical model for sound radiation from annular jet pipes: far- and near-field solutions," *J. Fluid Mech.*, vol. 549, pp. 315–341, 2006.
- [3] L. Gorazd, J. Jurkiewicz, and A. Snakowska, "Experimental verification of the theoretical model of sound radiation from an unflanged duct with low mean flow," *Archives of Acoustics*, vol. 37, no. 2, pp. 227–236, 2012.
- [4] R. M. Munt, "The interaction of sound with a subsonic jet issuing from a semi-infinite cylindrical pipe," *J. Fluid Mech.*, vol. 83, no. 4, pp. 609–640, 1977.
- [5] S. Rienstra, "Acoustic radiation from a semi-infinite annular duct in a uniform subsonic mean flow," *J. of Sound and Vibration*, vol. 94, pp. 267–288, 1984.
- [6] S. D. Savkar, "Radiation of cylindrical duct acoustic modes with flow mismatch," *J. of Sound and Vibration*, vol. 42, no. 3, pp. 363–386, 1975.
- [7] M. Hornikx, W. De Roeck, and W. Desmet, "Simplified exhaust pipe noise radiation modelling using the fourier pstd method," in *International Conference on Noise and Vibration Engineering. Leuven, Belgium*, 2010, pp. 20–22.
- [8] X. Chen, X. Zhang, C. Morfey, and P. Nelson, "A numerical method for computation of sound radiation from an unflanged duct," *J. of Sound and Vibration*, vol. 270, pp. 573–586, 2004.
- [9] C. Rumsey, R. Biedron, and F. Farassat, "Ducted-fan engine acoustic predictions using a navierstokes code," *J. of Sound and Vibration*, vol. 213, no. 4, pp. 643–664, 1998.
- [10] Y. Ozyoruk and L. Long, "Computation of sound radiating from engine inlet," *American Institute of Aeronautics and Astronautics Journal*, vol. 34, no. 5, pp. 894–901, 1996.
- [11] P. A. Skordos, "Modeling flue pipes: subsonic flow, lattice boltzmann, and parallel distributed computers," Ph.D. dissertation, Massachusetts Institute of Technology, 1995.
- [12] A. R. da Silva and G. P. Scavone, "Lattice Boltzmann simulations of the acoustic radiation from waveguides," *J. Phys. A: Math. Theor.*, vol. 40, pp. 397–408, 2007.
- [13] A. R. da Silva, "Numerical studies of aeroacoustic aspects of wind instruments," Ph.D. dissertation, McGill University, Montreal, Canada, 2008.
- [14] A. R. da Silva, G. P. Scavone, and A. Lefebvre, "Sound reflection at the open end of axisymmetric ducts issuing a subsonic mean flow: A numerical study," *J. of Sound and Vibration*, vol. 327, pp. 507–528, 2009.
- [15] A. R. da Silva, G. P. Scavone, and A. Lenzi, "Numerical investigation of the mean flow effect on the acoustic reflection at the open end of clarinet-like instruments," *ACTA Acustica United with Acustica*, vol. 96, pp. 959–966, 2010.
- [16] R. M. Munt, "Acoustic transmission properties of a jet pipe with subsonic jet flow: I. the cold jet reflection coefficient," *J. of Sound and Vibration*, vol. 142, no. 3, pp. 413–436, 1990.
- [17] S. Allam and M. Abom, "Investigation of damping and radiation using full plane wave decomposition in ducts," *J. of Sound and Vibration*, vol. 292, no. 3-5, p. 519534, 2002.
- [18] M. C. A. M. Peters, A. Hirschberg, A. J. Reijnen, and A. P. J. Wijnands, "Damping and reflection coefficient measurements for an open pipe at low mach and low helmholtz numbers," *J. Fluid Mech.*, vol. 256, p. 499534, 1993.
- [19] S. Succi, *The lattice Boltzmann equation for fluid dynamics and beyond*. Oxford University Press, 2001.
- [20] D. A. Wolf-Gladrow, *Lattice Gas Cellular Automata and Lattice Boltzmann Models: An Introduction. Lecture Notes in Mathematics*. Springer, Berlin / Heidelberg, 2004.
- [21] Y. Qian, D. d'Humieres, and P. Lallemand, "Lattice bkg models for navier-stokes equation," *Europhysics Letters*, vol. 17, no. 6, pp. 479–484, 1992.
- [22] E. W. S. Kam, R. M. C. So, and R. C. K. Leung, "Non-reflecting boundary conditions for one-step lbm simulation of aeroacoustics," in *27th AIAA Aeroacoustics Conference, 8-10 May 2006, Cambridge, Massachusetts*, 2006.

- [23] I. Halliday, L. A. Hammond, C. M. Care, K. Good, and A. Stevens, "Lattice boltzmann equation hydrodynamics," *Physical Review E*, vol. 64, no. 011208, pp. 1–8, 2001.
- [24] T. Lee, H. Huang, and C. Shu, "An axisymmetric incompressible lattice bgk model for simulation of the pulsatile flow in a circular pipe," *International Journal for Numerical Methods in Fluids*, vol. 49, no. 1, pp. 99–116, 2005.
- [25] M. Bouzidi, M. Firdaouss, and P. Lallemand, "Momentum transfer of a Boltzmann-lattice fluid with boundaries," *Physics of Fluids*, vol. 13, no. 11, p. 3452, 2001.
- [26] P. Lallemand and L. S. Luo, "Lattice Boltzmann method for moving boundaries," *J. of Computational Physics*, vol. 184, pp. 406–421, 2003.

Is the jet-drive flute model able to produce modulated sounds like Flautas de Chinos ?

Soizic Terrien

Laboratoire de Mécanique et
d'Acoustique - CNRS UPR 7051
Aix-Marseille Université,
Marseille, France

terrien@lma.cnrs-mrs.fr

Christophe Vergez

Laboratoire de Mécanique et
d'Acoustique - CNRS UPR 7051
Aix-Marseille Université,
Marseille, France

vergez@lma.cnrs-mrs.fr

Patricio de la Cuadra

CITA, Centro de Investigación
en Tecnologías de Audio, Pontifi-
cia Universidad Católica de Chile,
Santiago, Chile

pcuadra@uc.cl

Benoît Fabre

LAM-UPMC Univ. Paris 6,
UMR 7190, Institut Jean
le Rond d'Alembert,
Paris, France

benoit.fabre@upmc.fr

ABSTRACT

Flautas de chinos - prehispanic chilean flutes played during ritual celebrations in central Chile - are known to produce very particular beating sounds, the so-called *sonido rajado*. Some previous works have focused on the spectral analysis of these sounds, and on the input impedance of the complex resonator. However, the beating sounds origin remains to be investigated.

Throughout this paper, a comparison is provided between the characteristics of both the sound produced by *flautas de chinos* and a synthesis sound obtained through time-domain simulation of the jet-drive model for flute-like instruments. Jet-drive model appears to be able to produce quasiperiodic sounds similar to *sonido rajado*. Finally, the analysis of the system dynamics through numerical continuation methods allows to explore the production mechanism of these quasiperiodic regimes.

1. INTRODUCTION

The *Flautas de Chinos* are prehispanic chilean flutes still performed nowadays in ritual celebrations. The specificity of these instruments lies both in the production of very particular sounds (the so-called *sonido rajado*, literally "torn sounds") and in the complex form of their resonator (highlighted in the schematic representation of the instrument provided in figure 1), formed by two pipes in series with different diameters. Early works by Wright and Campbell [1] have focused on the characteristics of the produced sounds (in terms of fundamental frequency, waveform and spectrum). Blanc *et al.* [2] then attempted to describe and model the complex resonator.

However, the production mechanism of *sonido rajado* remains unexplained. Moreover, to the authors knowledge, the behaviour of such instruments has never been investigated through the study of a physical model of sound production.

Copyright: ©2013 Soizic Terrien *et al.* This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

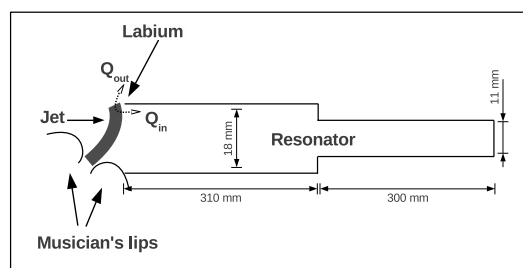


Figure 1. Schematic representation of both the resonator of the studied *flauta de chinos* and the excitation mechanism in flute-like instruments.

We do not develop in this paper an accurate, specific physical model of *flautas de chinos*, but we propose to evaluate the ability of the classical jet-drive model for flute-like instruments to produce regimes similar to *sonido rajado*. Moreover, such a study allows us to investigate the generation mechanism of those particular regimes.

The second section focuses on the description of the sound produced by *flautas de chinos*. In section 3, we recall the state-of-the-art model for flute-like instruments, whose capacity to produce modulated sounds is evaluated in section 4 using time-domain simulation. Finally, we examine in section 5 the origin of such regimes.

2. ANALYSIS OF SONIDO RAJADO

The "sonido rajado" produced by *Flautas de Chinos* seems to be the combination of a very loud sound, a rich harmonic structure, and an outstanding beating component [1, 2]. These beats can evoke the particular features of some quasiperiodic regimes, also produced for example in recorders or transverse flutes [3, 4]. The sound produced by a musician playing a particular *flauta de chinos* named *puntera* -whose dimensions are specified in figure 1- has been recorded through a microphone located outside the instrument, and is represented in figure 2. One can note that this radiated sound shows strong envelope modulations, whose modulation frequency can be roughly approximate between 13 and 15 Hz.

To understand the production mechanism of these particular sounds, the knowledge of the regime nature is an essential information. However, the study of spectrum and

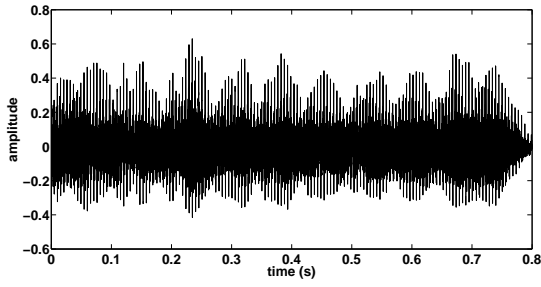


Figure 2. Radiated sound produced by a *flauta de chinos* "puntera".

waveform may be insufficient to distinguish, for example, a quasiperiodic regime from a chaotic regime. The computation of the Poincaré section (see for example [5, 6]) is a commonly used method that allows to reveal the nature of the regime. However, in the present case, both the small length of the recorded signal and its nonstationarity prevent the use of this tool.

Return times can be computed to estimate periodicity at different time scales in embedded time series. The so-called return times are the times for which the time series returns to its location in the pseudo phase space [7]. When this is done by choosing any point of the time series as a reference point, an histogram can be computed, which represents the occurrence of the different return times. Strong peaks in the histogram indicate periodicity. According to Slater's theorem [7], for a noise-free signal, a quasiperiodic regime is characterized by the presence of three peaks in the histogram, corresponding to three "return times". Two of them are respectively related to the "base frequency" and to the "modulation frequency", whereas the third is the sum of the two others. In experimental time series, noise is inevitable. Its influence has been studied on return times by Zou [7] who concludes that other peaks are likely to occur, but the sum relationship still holds.

The histogram of the recorded sound, displayed in figure 3, shows two principal peaks, respectively for the return times $t_1 = 0.0054$ s and $t_2 = 0.0767$ s. It highlights the presence, in the recorded sound, of two incommensurate frequencies: the "base frequency" $f_1 = 185$ Hz and the "modulation frequency" $f_2 = 13.04$ Hz. All other return times of the histogram are either multiples of t_1 or combinations of the form $t_2 + p \cdot t_1$, with p an integer. The histogram structure thus seems to satisfy Slater's theorem. Indeed, in the case of a chaotic regime, other peaks should appear in the histogram, related to return times which are neither multiples nor linear combinations of t_1 and t_2 .

Therefore, the computation of return times allows us to conclude about the quasiperiodic nature of the *sonido rajado* produced by *flautas de chinos*.

3. MODEL OF FLUTE-LIKE INSTRUMENT

3.1 General mechanism of sound production

Although instruments such as transverse flutes, flue organ pipes or the prehispanic chilean flutes studied here may

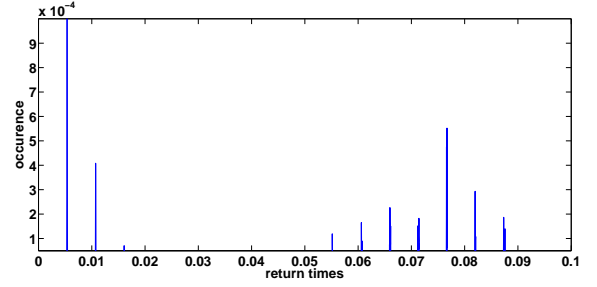


Figure 3. Histogram representing the return times of the sound produced by the "puntera" *flauta de chinos*.

appear different from each other, they involve the same general mechanism of sound production. When the musician blows into the instrument, a naturally unstable jet is created at a channel exit. This channel is formed by the player's lips for transverse flutes and *flautas de chinos*, whereas it is a part of the instrument for recorders and flue organ pipes (see figure 1). This jet oscillates around an edge called "labium". The resulting interaction produces acoustic energy in the resonator, constituted by the air column contained in the instrument. In turn, the acoustic field thus created in the instrument perturbs the jet at the channel exit. This perturbation being convected and amplified from the channel exit toward the labium, it sustains the oscillation of the jet around the labium, and thus allows the emergence of self-sustained oscillations.

This functioning is detailed in [8], whose main elements are recalled in sections 3.2 and 3.3.

3.2 Exciter

3.2.1 Hydrodynamics of the unstable jet

Once a permanent operating regime of the instrument is reached, the perturbation of the jet (the so-called receptivity) is provided by the acoustic field, at the channel exit. According to the empirical model proposed by de la Cuadra [9], the receptivity is represented as a transverse displacement $\eta(0, t)$ of the jet at the channel exit:

$$\eta(0, t) = \frac{h}{U_j} v_{ac}(t), \quad (1)$$

where h is the channel height (*i.e.* the opening height of the musician's lips), v_{ac} is the acoustic velocity in the pipe, and U_j is the jet central velocity (directly related to the fact that the musician blows hard or not).

As the jet is naturally unstable, this perturbation is amplified while convecting along the jet, from the channel exit to the labium. Based on the work of Rayleigh concerning instability of an infinite jet [10], this phenomenon can be described in a simplified way as an exponential amplification of the perturbation with respect to the convection distance x :

$$\eta(x) = \eta(0)e^{\alpha_i x}, \quad (2)$$

Based on experimental works by de la Cuadra [11], an approximation of the amplification parameter α_i through the

empirical expression:

$$\alpha_i \approx 0.4/h, \quad (3)$$

seems reasonable regarding the range of the Strouhal number $Str = \frac{fh}{U_j}$ considered throughout the paper.

Finally, we can express the jet transversal deflection at the labium:

$$\eta(W, t) = \eta(0, t - \tau)e^{\alpha_i W} = \frac{h}{U_j} v_{ac}(t - \tau)e^{\alpha_i W}, \quad (4)$$

where W is the distance between the channel exit and the labium (corresponding to the length of the jet, highlighted in figure 1) and τ is the convection delay of the perturbation along the jet. Calling c_p the convection velocity of this perturbation, τ is given by: $\tau = \frac{W}{c_p}$. Both theoretical and experimental works [9, 10, 12] have shown that in the general case $0.3U_j \leq c_p \leq 0.5U_j$, which leads to:

$$\tau \approx \frac{W}{0.4U_j} \quad (5)$$

3.2.2 Nonlinear jet-labium interaction

The interaction between the perturbed jet described above and the labium (see figure 1) makes the jet oscillate from one side to another of the labium. This oscillation causes a flow injection alternatively inside and outside the instrument. Based on the jet-drive model proposed by Coltman and followed by Verge [13, 14], these two (localised) flow injection in phase opposition are modelised as a dipolar pressure source:

$$\Delta p = -\frac{\rho \delta_d}{WH} \cdot \frac{dQ_{in}}{dt}, \quad (6)$$

In this equation, ρ is the air density, H the width of the jet and we note, as proposed by Verge [15], δ_d the effective distance between the two flow injection points. Based on both theoretical works and empirical estimation of parameters, Verge [15] proposed the approximation $\delta_d \approx \frac{4}{\pi} \sqrt{(2hW)}$, which is adopted here. Q_{in} represents the flow injection in the pipe [8], given by:

$$Q_{in} = H \int_{-\infty}^{y_0 - \eta(t)} U(y) dy, \quad (7)$$

In this expression, y_0 is the offset between the labium and the jet centerline, and $U(y)$ the velocity profile of the jet, representing by a Bickley profile:

$$U(y) = U_j \operatorname{sech}^2\left(\frac{y}{b}\right), \quad (8)$$

where b is the half width of the jet, related (under certain assumptions detailed for example in [16]) to the height of the channel exit through $b = 2h/5$.

By injecting equations (7) and (8) in equation (6), we finally obtain the expression of the pressure source that excites the resonator:

$$\Delta p_{src}(t) = \frac{\rho \delta_d b U_j}{W} \cdot \frac{d}{dt} \left[\tanh\left(\frac{\eta(t) - y_0}{b}\right) \right]. \quad (9)$$

3.2.3 Nonlinear losses

Between the channel exit and the labium of flue instruments, the presence of an important transversal flow induced by the acoustic field in the pipe can cause, for high acoustic velocities, vortex shedding at the labium [17]. This phenomenon is modeled as nonlinear losses, and so represented by an additional nonlinear term Δp_{los} in equation (9):

$$\begin{aligned} \Delta p(t) &= \Delta p_{src}(t) + \Delta p_{los}(t) \\ &= \frac{\rho \delta_d b U_j}{W} \frac{d}{dt} \left[\tanh\left(\frac{\eta(t) - y_0}{b}\right) \right] \\ &\quad - \frac{\rho}{2} \left(\frac{v_{ac}(t)}{\alpha_{vc}} \right)^2 \operatorname{sgn}(v_{ac}(t)) \end{aligned} \quad (10)$$

where $\alpha_{vc} \approx 0.6$ is a *vena contracta* factor.

3.3 Resonator

The acoustical response of the air column contained in the pipe, excited by the pressure source described above, is represented through the input admittance $Y_{in} = V_{ac}/\Delta P$ of this resonator. V_{ac} and ΔP are respectively the frequency-domain expressions of the acoustic velocity at the pipe inlet and the pressure source. The use of a modal decomposition of Y_{in} is interesting as it allows an independant control of the different resonance mode characteristics. In the frequency domain, Y_{in} is thus represented as a sum of resonance modes:

$$Y_{in} = \sum_m \frac{a_m j \omega}{\omega_m^2 - \omega^2 + j \omega \frac{\omega_m}{Q_m}}, \quad (11)$$

where a_m , ω_m and Q_m are respectively the modal amplitude, the resonance pulsation and the quality factor of the m^{th} resonance mode.

As the particularity of the *flautas de chinos* lies in the shape of their resonator, we use a modal decomposition of the measured admittance of the "puntera" flute whose sound is studied in section 2. The measure was realised thanks to an impedance sensor for wind instruments [18]. The corresponding modal parameters (displayed in table 1) of the first five resonance modes are estimated by fitting the admittance calculated using equation (11) to the measured one. One can note that the agreement between the original admittance and the fitted one, which are both displayed in figure 4, can be improved by increasing the number of resonance modes taken into account in the fit process. However, it considerably increases the computation cost for the foregoing different resolution methods.

3.4 Numerical resolution of the model

The complexity of the model, due to its neutral nature and to the presence of nonlinear terms, imposes the use of numerical methods for solving the model. The neutral nature is related to the presence of a delayed derivative term, as discussed in [19]. In this study, we compare the results provided by two complementary methods: a classical time-domain simulation scheme, and a calculation method of

mode number	a_m	ω_m	Q_m
1	11.39	1156.7	26
2	7.05	2342.8	34.4
3	9.55	4796.4	50.7
4	8.12	5943.4	52.9
5	12.93	8418.9	58

Table 1. Fitted modal coefficients corresponding to the measured admittance of the *puntera* flute.

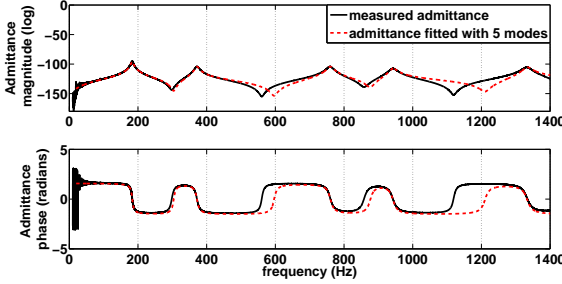


Figure 4. Measured and fitted input admittance of a *puntera* prehispanic flute. Only the five first resonance modes taken into account in the model are represented.

periodic solutions coupled with a numerical continuation algorithm. Both are presented below and used in sections 4 and 5.

3.4.1 Time-domain simulations

Time-domain simulations are performed in Matlab/Simulink, with a classical fourth-order Runge-Kutta method. Such an approach is interesting as it allows to access many kind of stable stationary regimes (for example static, periodic, quasiperiodic or chaotic) and to transients between these regimes. However, results are very sensitive to initial conditions and to numerical parameters such as the sampling frequency. Moreover, unstable solutions cannot be computed: a slight disturbance (noise or even round-off errors, for example) is enough to make the system move away from these solutions. Therefore, it can be difficult to access, with such a method, to a global knowledge of the different oscillation regimes, essential to understand the system dynamics.

3.4.2 Orthogonal collocation and numerical continuation

Orthogonal collocation method allows to compute periodic solutions of dynamical systems. Its principle is based on the discretization of a single (unknown) period of the solution x on n intervals. On each interval, the solution is approximated by a polynomial of degree d . Projecting the model equations on this set of representation points leads to an algebraic system, whose unknowns are the value of x at each discrete point, and the value of the period T .

This method, which allows to determine a solution x_0 for a set of model parameters λ_0 , is coupled to a numerical continuation algorithm. Starting from x_0 , it computes, using a Newton-Raphson prediction-correction method, the

neighboring solution corresponding to a set of parameter $\lambda_0 + \Delta\lambda$ [20]. Proceeding by successive iterations, one finally access to a complete *branch* of periodic solutions.

The computation of the Floquet multipliers allows subsequently to determine the stability properties of each point of the branch [5]. Indeed, according to Floquet theory, a periodic solution (*i.e.* a point of the branch) is stable until all its Floquet multipliers lies in the unit circle. When a periodic solution loses its stability, the resulting regime observed after this bifurcation point depends on which way the Floquet multipliers leave the unit circle at the bifurcation point. For example, a quasiperiodic regime may appear if two complex conjugate Floquet multipliers leave the unit circle [5].

Giving access to the bifurcation diagram, which represents in an ideal case all the periodic and static solutions of the studied system, this method provides a more global knowledge of the system dynamics (coexistence of multiple solutions, unstable solutions ...), and thus permits an easier interpretation of phenomenon such as hysteresis or emergence of non periodic regimes (see for example [19]).

The neutral nature of the model prevents the use of classical numerical continuation software (such as, for example, AUTO [21] or Manlab [22]). We use here DDE-Biftool [23], a software specifically developed for numerical continuation in delay differential equations, and its extension for neutral systems [24].

4. TIME-DOMAIN SIMULATION: CHARACTERISTICS OF THE SYNTHESIS SOUND

Although we study here a very generic model of flute-like instruments, time-domain simulations lead to non periodic regimes which can perceptually recall the *sonido rajado* produced by *flautas de chinos*. Because of the few studies about parameter estimation in *flautas de chinos*, some of the model parameters are chosen arbitrarily close to those currently uses in the case of a model of an alto recorder [19], except for the excitation window area WH and the labium offset y_0 which are here slightly higher than in the case of an alto recorder. As these parameters are fixed by the player in the case of *flautas de chinos*, the chosen values ($W = 1\text{cm}$, $H = 6\text{mm}$ and $y_0 = 0.2\text{ mm}$) seem reasonable. Parameter values used throughout the paper are provided in table 2.

A first simulation is achieved using a constant jet velocity $U_j = 39\text{ m/s}$, which would correspond to a case where the musician blows in a steady way. This value, which would be equivalent, using the stationary Bernoulli equation, to a mouth pressure of about 915 Pa, is of the same order of magnitude as the values measured in the player's mouth for the *puntera* flute (between 886 Pa and 1546 Pa).

The computed acoustic velocity at the pipe inlet v_{ac} , represented in figure 5 as a function of the time, shows, as the real sound (see figure 2), envelope modulations. A Fourier analysis of the signal envelop provides the modulation frequency $f_{mod} = 8.6\text{ Hz}$. One can note that this value is of the same order of magnitude as the modulation frequency of the real sound, $f_2 = 13.04\text{ Hz}$ (see section 2), and coherent with the observations of Wright and

Parameter	Numerical value (S.I.)
α_i	$0.4/h = 400$
δ_d	$\frac{4}{\pi} \sqrt{(2hW)} = 0.0057$
b	$2h/5 = 0.0004$
y_0	0.0002
h	0.001
W	0.01
ρ	1.2
c_p	$0.4U_j$
α_{vc}	0.6

Table 2. Parameter values used for numerical resolutions of the model.

Campbell [1]. In the case of time-domain simulation, this frequency strongly depends on the jet velocity U_j , and a slight shift of this parameter is enough to achieve the experimental value of 13.04 Hz. However, the fact that the modulation frequency is related to the jet velocity is a significant difference compared to the real instruments behaviour. Indeed, as observed by Wright and Campbell [1], the modulation frequency shows only little variations with the pressure in the musician's mouth.

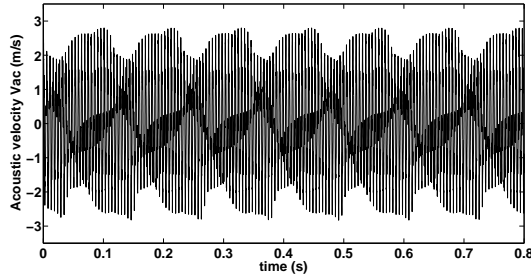


Figure 5. Acoustic velocity computed through a time-domain simulation of the model for a constant jet velocity $U_j = 39$ m/s, represented as a function of the time. The sampling frequency used for computation is 22x44.1 kHz.

Although the physical model is not specific to the instrument studied here, and that some parameters can only be roughly estimated, the implementation of a measured admittance allows us to compare qualitatively, in figure 6, the spectrum of the synthesis sound with that of the real sound. Such a comparison shows that both the fundamental and modulation frequencies (particularly visible around 370 Hz) of the synthesized sound are of the same order of magnitude as those of the real sound. One can note that the amplitudes can not be compared: indeed, we consider the spectrum of the sound radiated by the real instrument whereas we consider, for the model, the internal acoustic field (which would correspond to a measure under the labium of the instrument). Moreover, to make easier the qualitative comparison of the frequencies, the two spectra have been normalized with respect to their own maximum.

As for the real sound, one can wonder about the quasiperiodic or chaotic nature of the synthesis sound. The study of the Poincaré section allows us to distinguish these differ-

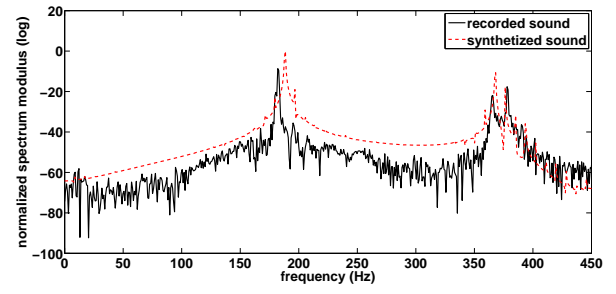


Figure 6. Normalized spectrums of both the sound produced by a "puntera" flauta de chinos (represented in figure 2) and the synthesis sound represented in figure 5.

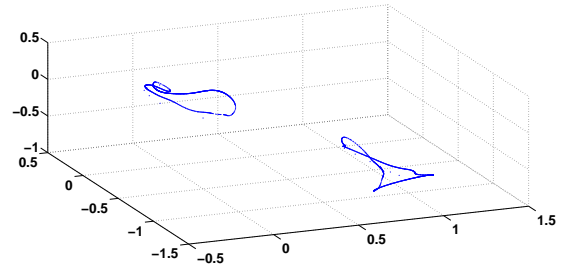


Figure 7. Three-dimensional Poincaré section of the synthesis sound represented in figure 5. The closed curve is characteristic of a two-period quasiperiodic regime.

ent kinds of regime from each other. As shown on figure 7, the three-dimensional Poincaré section of the synthesis signal reveals a densely packed closed curve, which is characteristics of a two-period quasiperiodic solution (see for example [5]), and thus allows to conclude about the quasiperiodic nature of the simulated signal.

As a first conclusion, one can note that this study highlights the ability of the state-of-the-art physical model of flute-like instruments to produce regimes where some of their features such as quasiperiodic nature, fundamental and modulation frequencies, are similar to those of the sound produced by a flauta de chinos.

5. QUASIPERIODIC REGIME: GENERATION MECHANISM

As it provides a more global knowledge of the system dynamics, the study of the bifurcation diagram of the model studied, which ideally represents all its stable and unstable periodic solutions as a function of one of its parameters, has recently shown its interest in the understanding of the functioning of musical instruments [25].

The bifurcation diagram of the model displayed in figure 8 represents the oscillation frequency of the different periodic solutions as a function of the jet velocity U_j . Solutions are computed for a range of the jet velocity U_j for which quasiperiodic regimes may occur in time-domain simulations (see section 4). It shows the existence, in this range of U_j , of two periodic solutions branches: the first corresponds to the first register of the instrument (that is

to say, to oscillations emerging from an instability at a frequency close to the first resonance frequency of the resonator), and the second is related to the second register. Solid and dotted lines respectively represent stable and unstable parts of periodic solution branches.

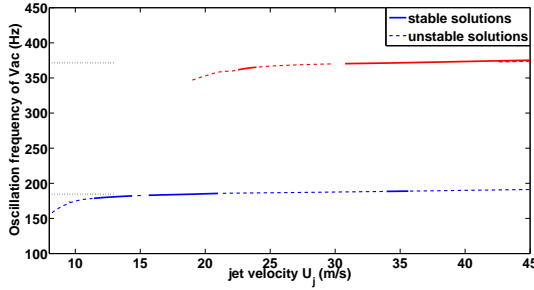


Figure 8. Bifurcation diagram of the studied system, with two periodic solution branches corresponding to the first and the second registers. Abscissa: jet velocity U_j (m/s). Ordinate: oscillation frequency (Hz). The horizontal dotted lines correspond to the two first resonance frequencies of the resonator.

In the previous section, we highlighted that time-domain simulation of the model reveals the existence of a quasiperiodic regime for $U_j = 39$ m/s. For this value, the bifurcation diagram predicts that the first register is unstable, whereas the second register is stable.

In order to understand the origin of this quasiperiodic solution, we computed a second time-domain simulation, during which the jet velocity U_j follows a decreasing ramp, from $U_j = 39$ m/s, to $U_j = 32$ m/s. The result, displayed in figure 9, shows a gentle transition from the quasiperiodic regime studied in the previous section, to a periodic regime. A Fourier analysis of this periodic regime provides an oscillation frequency of 188.4 Hz, which is very close to the first resonance frequency $\frac{\omega_1}{2\pi} = 184.4$ Hz, and thus allows to conclude that this periodic solution corresponds to the first register.

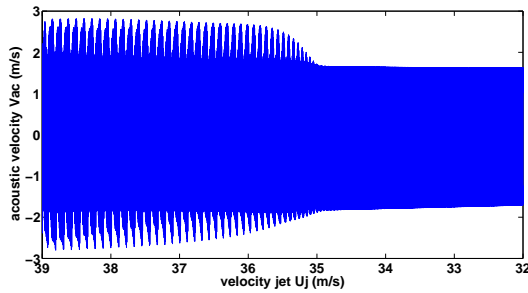


Figure 9. Acoustic velocity computed through a time-domain simulation, for a linear decreasing ramp of the jet velocity U_j . The sampling frequency used for simulation is 22x44.1 kHz.

Thereby, the confrontation of figures 8 and 9 suggests that the quasiperiodic solution is the result of the loss of stability of the first register (predicted by the bifurcation diagram at $U_j = 35.6$ m/s - see figure 8) through a direct

Neimark-Sacker bifurcation [5], .

If the numerical tools used here do not allow computation and continuation of quasiperiodic solution branches, a further analysis of the stability properties of the periodic solutions permits to analyse the birth of quasiperiodic solutions (see section 3.4).

The Floquet multipliers associated to the first register, around the bifurcation point ($U_j = 35.6$ m/s) are shown in figure 10 for both the point of the branch corresponding to $U_j = 35.3$ m/s (just before the loss of stability) and the point corresponding to $U_j = 35.6$ m/s (just after the loss of stability). One can clearly observe that the loss of stability is the result of the crossing of the unit circle by two complex conjugate Floquet multipliers.

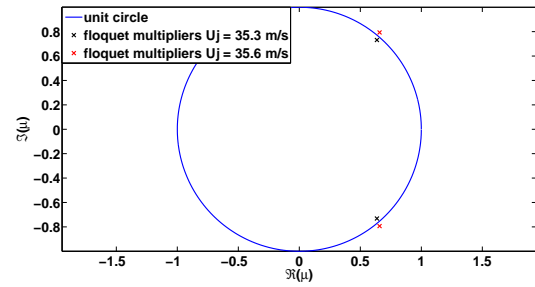


Figure 10. Floquet multipliers represented in the complex plane, for two points of the first register periodic solution branch. The first point is located just before the loss of stability of the first register ($U_j = 35.3$ m/s) and the second one just after ($U_j = 35.6$ m/s).

This analysis not only confirms the generation of a quasiperiodic regime, but it also provides the value of the modulation frequency at the quasiperiodic regime threshold (*i.e* at $U_j = 35.6$ m/s), through the computation of the Floquet exponents γ_f , which are related to the Floquet multipliers μ_f through [5]:

$$\gamma_f = \frac{1}{T} \ln(\mu_f) \quad (12)$$

where T is the period of the periodic solution at the bifurcation point. According to Floquet theory, the imaginary part of the Floquet exponent corresponds to the modulation pulsation ω_{mod} . Here, such an analysis of the Floquet multipliers predicts a modulation frequency: $f_{mod} = \omega_{mod}/2\pi = 26.4$ Hz. A further study of the synthesis signal envelop represented in figure 9 shows a threshold of the quasiperiodic regime at about $U_j = 34.2$ m/s (see a zoom of figure 9 in figure 11). As in section 4, a Fourier analysis of this envelop signal provides the modulation frequency related to the quasiperiodic regime: $f_{modsimu} = 26.3$ Hz. Although the quasiperiodic regime threshold seems to be underestimated in the time-domain simulation (which may be a consequence of both the sampling and the dynamics of the control parameter U_j [26]), the modulation frequency shows good agreement with those predicted by the Floquet analysis.

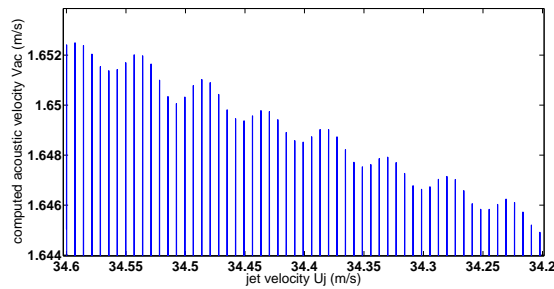


Figure 11. Zoom of figure 9: acoustic velocity computed through a time-domain simulation, for a linear decreasing ramp of the jet velocity U_j . The sampling frequency used for simulation is 22×44.1 kHz.

6. CONCLUSION

This study, rather than proposing a specific physical model of *flautas de chinos*, focuses on the ability of the jet-drive model of flute-like instruments to generate sounds of the same kind as *sonido rajado* produced by these instruments.

As a first conclusion, a study of the sound produced by a *flauta de chinos* has highlighted the quasiperiodic nature of *sonido rajado*.

Secondly, time domain simulations show that not only the jet-drive model can produce, as transverse flutes (for example), periodic solutions associated to the first or second register of the instrument, but also quasiperiodic regimes, which present similarities with the *sonido rajado* of pre-hispanic chilean flutes.

Moreover, an analysis of the system dynamics through its bifurcation diagram highlights the production mechanism of those quasiperiodic regimes. Comparison between the sound produced by *flautas de chinos* and the model dynamics thus suggests that the desired behaviour of such instruments would correspond to a quasiperiodic regime emerging from a direct Neimark-Sacker bifurcation of the periodic solution branch related to the first register of the instrument.

However, a quantitative comparison between experimental and synthesis sounds is impossible here as we study on the one hand the sound radiated by the real instrument whereas we consider, in the model, the internal acoustic field. But more importantly, both the fact that the precise values of different parameters of the model are unknown and the fact that different phenomena are not taken into account (for example, the jet is certainly turbulent [2]) limit such a quantitative comparison. These elements probably explain why significant differences remain between the real sound and the synthesis one, such as, for quasiperiodic regimes, the variation of the modulation frequency with the jet velocity.

7. REFERENCES

[1] H. Wright and D. Campbell, "Analysis of the sound of chilean pifilca flutes," *The Galpin Society Journal*, vol. 51, pp. 51–63, 1998.

[2] F. Blanc, P. de la Cuadra, B. Fabre, G. Castillo, and C. Vergez, "Acoustics of the 'flautas de chinos'," in *Proc. of 20th Int. Symposium on Music Acoustics*, Sydney, 2010.

[3] J. Coltman, "Jet offset, harmonic content, and warble in the flute," *Journal of the Acoustical Society of America*, vol. 120, no. 4, pp. 2312–2319, 2006.

[4] N. Fletcher, "Sound production by organ flue pipes," *Journal of the Acoustical Society of America*, vol. 60, no. 4, pp. 926–936, 1976.

[5] A. Nayfeh and B. Balachandran, *Applied Nonlinear Dynamics*. Wiley, 1995.

[6] V. Gibiat, "Phase space representations of acoustical musical signals," *Journal of Sound and Vibration*, vol. 123, no. 3, pp. 529–536, 1988.

[7] Y. Zou, "Exploring recurrences in quasiperiodic dynamical systems," Ph.D. dissertation, Potsdam University, 2007.

[8] A. Chaigne and J. Kergomard, *Acoustique des instruments de musique (Acoustics of musical instruments)*. Belin (Echelles), 2008, ch. 10.

[9] P. de la Cuadra, C. Vergez, and B. Fabre, "Visualization and analysis of jet oscillation under transverse acoustic perturbation," *Journal of Flow Visualization and Image Processing*, vol. 14, no. 4, pp. 355–374, 2007.

[10] J. Rayleigh, *The theory of sound second edition*. New York, Dover, 1894.

[11] P. de la Cuadra, "The sound of oscillating air jets: Physics, modeling and simulation in flute-like instruments," Ph.D. dissertation, Stanford University, 2005.

[12] A. Nolle, "Sinuous instability of a planar jet: propagation parameters and acoustic excitation," *Journal of the Acoustical Society of America*, vol. 103, no. 6, pp. 3690–3705, 1998.

[13] J. Coltman, "Jet drive mechanisms in edge tones and organ pipes," *Journal of the Acoustical Society of America*, vol. 60, no. 3, pp. 725–733, 1976.

[14] M. Verge, A. Hirschberg, and R. Caussé, "Sound production in recorder-like instruments. ii. a simulation model," *Journal of the Acoustical Society of America*, vol. 101, no. 5, pp. 2925–2939, 1997.

[15] M. Verge, R. Caussé, B. Fabre, A. Hirschberg, A. Wijnands, and A. van Steenberghe, "Jet oscillations and jet drive in recorder-like instruments," *Acta Acustica united with Acustica*, vol. 2, pp. 403–419, 1994.

[16] C. Ségoufin, B. Fabre, M. Verge, A. Hirschberg, and A. Wijnands, "Experimental study of the influence of the mouth geometry on sound production in a recorder-like instrument: windway length and chamfers," *Acta Acustica united with Acustica*, vol. 86, pp. 649–661, 2000.

- [17] B. Fabre, A. Hirschberg, and A. Wijnands, "Vortex shedding in steady oscillation of a flue organ pipe," *Acta Acustica united with Acustica*, vol. 82, no. 6, pp. 863–877, 1996.
- [18] C. Macaluso and J. Dalmont, "Trumpet with near-perfect harmonicity: Design and acoustic results," *Journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 404–414, 2011.
- [19] S. Terrien, R. Auvray, B. Fabre, P. Lagrée, and C. Vergez, "Numerical resolution of a physical model of flute-like instruments: comparison between different approaches," in *Proceedings of Acoustics 2012*, Nantes, France, 2012.
- [20] B. Krauskopf, H. Osinga, and J. Galan-Vioque, *Numerical continuation methods for dynamical systems*. Springer, 2007.
- [21] E. J. Doedel, "AUTO: A program for automatic bifurcation analysis of autonomous systems," *Congressus Numerantium*, vol. 30, pp. 265–284, 1981.
- [22] B. Cochelin and C. Vergez, "A high order purely frequency-based harmonic balance formulation for continuation of periodic solutions," *Journal of Sound and Vibration*, vol. 324, no. 1-2, pp. 243–262, 2009.
- [23] K. Engelborghs, "DDE Biftool: a Matlab package for bifurcation analysis of delay differential equations," Katholieke Universiteit Leuven, Tech. Rep., 2000.
- [24] D. Barton, B. Krauskopf, and R. Wilson, "Collocation schemes for periodic solutions of neutral delay differential equations," *Journal of Difference Equations and Applications*, vol. 12, no. 11, pp. 1087–1101, 2006.
- [25] S. Karkar, C. Vergez, and B. Cochelin, "Oscillation threshold of a clarinet model: a numerical continuation approach," *Journal of the Acoustical Society of America*, vol. 131, no. 1, pp. 698–707, 2012.
- [26] B. Bergeot, A. Almeida, C. Vergez, and B. Gazengel, "Prediction of the dynamic oscillation threshold in a clarinet model with a linearly increasing blowing pressure," *Arxiv preprint arXiv:1207.4636*, 2012.

The design of a chromatic quena: how can linear acoustic help ?

Camille Vauthrin, Benoît Fabre

LAM/d'Alembert

UPMC Univ Paris 06, CNRS UMR 7190

vauthrin@lam.jussieu.fr

Patricio de la Cuadra

CITA

Pontificia Universidad Catolica de Chile, Santiago, Chile

delac@ccrma.stanford.edu

ABSTRACT

While traditional instrument making relies trials and errors process, acoustics may help in designing a new instrument or evolution of existing instruments; resulting in a faster/more efficient designing process and/or better quality instrument. Because of complex and intricate relations between the different building parameters, it also may focus work on the specific parameters. This work can also help to grasp global influence of each geometric parameters on the note's pitch, harmonicity between the registers, the timbre or the field of freedom in play. The definition of the requirements provides the objectives and the limitations of the design process. The requirements are established from linear acoustics, discussions with the flute maker. The prototype of the chromatic quena should correspond to these; especially of the crossover between Andean sounding aesthetics and modern flute playing techniques. Linear acoustic allows to give relatively correct details on resonance frequencies, however the timbre is more difficult to studied. In the aim to approach the diatonic quena's timbre, their distinguished elements are preserved on the chromatic quena as the notch, the knot. Moreover, large holes allows to obtain a regular timbre. This work shows that this design process can help the flute maker in realization of chromatic quena, especially in reducing the number of prototypes.

1. INTRODUCTION

The aim of this study is to help the instrument maker in the process of designing the geometry of new chromatic quena. The traditional quena is a diatonic instrument, producing a diatonic scale of seven notes in one octave by the opening of seven side holes. Semitones that are not found in this diatonic scale can still be produced, using cross fingerings, partial opening of the holes and adjusting the pitch with the lips, resulting in uneven timbre as well as awkward fingerings. On the opposite, the new chromatic quena should include eleven holes to allow for even timbre, clean intonation and ease of playing. Thus, each hole is associated to a semitone of the scale. Indeed, a key mechanism needs to be developed to control the open-

ing of 11 holes using 9 fingers, like that found in modern so-called Boehm flutes, and other woodwind instrument. Wolfe, [1], shows how Boehm's system with large holes allows a certain homogeneous tone. The study presented in this paper is part of a project including flute and quena players from Latin America and the french flute maker, J.Y Roosen. The new chromatic quena is built by the flute maker to allow crossover between Andean sounding aesthetics and modern flute playing techniques. The sounding aesthetics of the new chromatic quena should therefore be in close relation to that of the traditional quena. Therefore certain distinguishing elements belonging to the traditional quena should be preserved. While the method commonly used by instrument makers in such a case would be to go through several prototypes, our study shows how acoustic theory can help the design process, mostly by reducing the number of prototypes needed.

The second section of this paper reformulates the problem in an acoustic design problem, and discusses the acoustic, hypothesis on which the study is based. Then, the acoustic tools and methods are described. The third section describes the process of prototype's design and finally conclusion and expectations are established.

2. HYPOTHESIS, ACOUSTIC THEORIES

2.1 From global description to acoustical parameters

The traditional quena is a diatonic instrument. It is made out of a roughly cylindrical tube on which seven holes are located. The size of the holes is rather large with a typical value of 0.6 times the bore diameter, however the shape is not perfectly circular but carved for tuning purpose. The bore is also not exactly cylindrical because of the natural material used (most often bamboo cane). Two specific features are: the notch at the active extremity and the knot at the passive extremity which corresponds to a knot of bamboo. During the traditional design process, the notch is first carved, then the size of the knot and cylinder length (37-40 cm) adjusted for the tuning of the lowest note (G4). Only after, the holes are drilled and adjusted to the desired scale.

The chromatic quena should approach the timbre and the sounding aesthetics of the traditional diatonic quena while allowing a chromatic equal temperament scale in terms of pitch. To quantify these elements, acoustical parameters can be used. For example, resonance frequencies allow to approach the pitch of the notes played. Resonance frequencies are to be distinguished from the sounding frequencies, the latter corresponding of the instrument in active behavior. Indeed, sounding frequencies evolve according to the

speed of the jet or the register as shown by Auvray [2] and Meissner [3] or the position of the musician's lips, Cossette [4] and Coltman [5]. A deviation between resonance and sounding frequencies can be observed. The timbre or the sounding aesthetics are more difficult to quantify with parameters of linear acoustic, however some elements may be provided. To approach the traditional sound of the quena, the size of the bore, the shape of the notch and the knot should be preserved, moreover to obtain an even timbre across the scale, the holes should have similar sizes. Even if linear acoustic can't predict the timbre, these conditions should allow to stay close to the original sound.

2.2 Acoustic tools and method

Acoustical model of the cylinder: The theory of plane waves described in the Appendix allows to express the transfer matrix of a cylinder. Propagation is described using a transmission line formalism as explained in the books [6], [7]:

$$\begin{pmatrix} P_1 \\ U_1 \end{pmatrix} = \begin{pmatrix} \cosh \Gamma L & Z_c j \sinh \Gamma L \\ Z_c^{-1} j \sinh \Gamma L & \cosh \Gamma L \end{pmatrix} \begin{pmatrix} P_2 \\ U_2 \end{pmatrix} \quad (1)$$

where Γ is the propagation constant ($\Gamma = jk_c$, and k_c the complex wave number which takes into account viscosity and thermal conduction effects), L is the length of the cylinder and Z_c is the characteristic impedance expressed later. This transfer matrix allows to connect pressure P and acoustic flow U at each extremity of the cylinder.

The Stokes number is defined as the ratio of the cylinder radius to the boundary layer thickness: $r_v = |k_v a|$. k_v is expressed by $k_v = \sqrt{-\frac{j\omega}{c_0 l_v}}$, with $l_v = \frac{\mu}{\rho_0 c_0}$. a is the radius of the cylinder, c_0 is the velocity of the sound, ρ_0 is the air density and μ the dynamic viscosity. For large pipes with values of the Stokes number greater than 10, the constant of propagation Γ becomes:

$$\Gamma = \alpha + j \frac{\omega}{v_\phi} \quad (2)$$

where $\alpha = \frac{\omega}{c_0} \left[\frac{\alpha_1}{r_v} + \frac{\alpha_2}{r_v^2} \right]$, $\alpha_1 = 1.044$ and $\alpha_2 = 1.08$, $v_\phi = c_0 \cdot \frac{1}{1 + \alpha_1/r_v}$. Taking into account the visco-thermal effects, the characteristic impedance is written as:

$$Z_c = \frac{\rho_0 c_0}{S} \left[1 + \frac{\bar{\alpha}_1(1-j)}{r_v} - \frac{\bar{\alpha}_2 j}{r_v^2} \right] \quad (3)$$

where $\bar{\alpha}_1 = 0.37$ and $\bar{\alpha}_2 = 1.147$, for air. S is the cross-section of the cylinder.

The temperature-dependent parameters are: the sound velocity c_0 , the density ρ_0 , the dynamic viscosity μ and the thermal conductivity κ .

The holes: The side-holes perturb the acoustic-field inside the bore, the opening or the closing of these modifies the resonances of the instrument and thus allows the musician to play different notes. Therefore, resonances frequencies of the instrument change according to the geometry of the side holes and a precise description of each hole is essential to the prediction of the resonance frequencies.

While the geometry of the holes may be adjusted by the flute maker during the fine tuning of the instrument, we will restrict our study to idealized cylindrical geometry of the holes. The holes are described by their radius, chimney and position on the cylinder. It is assumed that the distance between tone holes is sufficiently large to make higher propagating mode interactions negligible.

A hole inserted between two segments cylindrical duct, relates the input, 1, and output quantities, 2:

$$\begin{pmatrix} P_1 \\ U_1 \end{pmatrix} = T_{cyl} T_{hole} T_{cyl} \begin{pmatrix} P_2 \\ U_2 \end{pmatrix} \quad (4)$$

where T_{hole} and T_{cyl} are respectively the transfer matrix of the hole and the cylinder. The transfer matrix of a side-hole may be approximated by a T-junction, as shown in figure 1.

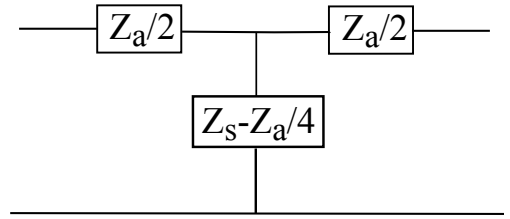


Figure 1: Block diagram showing the T-junction for a tonehole.

This T-junction depends on two parameters: the shunt impedance Z_s and the series impedance Z_a . These are defined differently for an open or closed hole. In the literature, there are several descriptions of tonehole on a pipe. Two geometrical cases are distinguished: a tonehole on a thick pipe and a tonehole with a chimney on a pipe without thickness. Different mathematical expressions for these impedances in low frequencies are available in the literature by Pierce, Nederveen, Dalmont, Dubos, Keefe [7–11] and Lefebvre [12] with finite elements. All of them are summarized and compared in Eveno's thesis [13]. After a comparison of these different mathematical expressions, the model used in this study is Lefebvre's model. In your case, this model is the most relevant, it allows to take into account of the pipe's thickness; which is responsible for the height of the tonehole. The radiation of the tonehole is approximated by that of a piston in an infinite plane baffle. Furthermore, this model is also coherent for small height of holes. This model takes into account the matching volume between the bore cylinder and the hole cylinder.

The transfer matrix corresponding of the T-junction's schema is expressed by:

$$T_{hole} = \begin{pmatrix} 1 + \frac{\bar{Z}_a}{2\bar{Z}_s} & \bar{Z}_a(1 + \frac{\bar{Z}_a}{4\bar{Z}_s}) \\ 1/\bar{Z}_s & 1 + \frac{\bar{Z}_a}{2\bar{Z}_s} \end{pmatrix} \quad (5)$$

where \bar{Z}_s and \bar{Z}_a are the dimensionless impedances defined respectively by $\bar{Z}_s = Z_s/Z_0$ and $\bar{Z}_a = Z_a/Z_0$ with

$Z_0 = \rho_0 c_0 / S$. Expressions impedance are then set according to the configuration of the hole: open (*o*) or closed (*c*) and detailed in [12].

The end knot: The end knot can be considered as a cylinder with a smaller section s_b . Then follows Γ_b and Z_{cb} . The transfer matrix T_b of a cylinder with a length l_b can be associated to this end knot.

Radiation of the instrument: The radiation of the passive extremity is approximated by that of a thin tube baffle, the radiation impedance can be written as follows in low frequencies:

$$Z_r = \frac{\rho_0 c_0}{s_b} \left(jk0.6r_b + \frac{1}{4}(kr_b)^2 \right) \quad (6)$$

with $s_b = \pi r_b^2$, 0.6 is probably underestimated because of the small baffle associated with the knot.

The whole instrument Finally, the whole instrument corresponds to the product of transfer matrices:

$$Z = T_h \cdot \prod_i \{T_{cyl_i} T_{hole_i} T_{cyl_i}\} \cdot T_b \quad (7)$$

Z is thus expressed as a matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, the input impedance of the quena can be expressed as:

$$Z_e = \begin{pmatrix} AZ_r & B \\ CZ_r & D \end{pmatrix} \quad (8)$$

$$Z_e = \frac{AZ_r + B}{CZ_r + D} \quad (9)$$

The quena being an open-open instrument, the resonance frequencies are those of input admittance peaks. The frequencies are estimated from the zero crossing of the imaginary part of the admittance by dichotomy-method.

3. DESIGNING A PROTOTYPE

The quena, figure 2, can be studied in two elements: the active part with the notch and the passive part: the resonator.

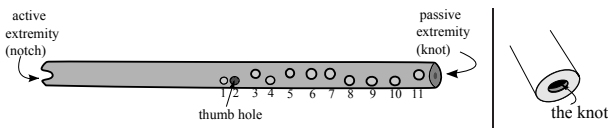


Figure 2: Left: schematic figure of a quena on which are illustrated toneholes, passive and active extremities; right: zoom on the knot of the quena.

This project is developed in collaboration with a flutemaker. To easily interact with him, the design of the chromatic quena is based on the traditional design process. Therefore, no global optimization algorithm, as Le Vey's method [14], is used. The step by step design follows the traditional process. Moreover, this allows to understand the importance of the steps order in the realization, but also to observe the influence of each geometric parameter on the resonance frequencies.

The wood used to the final chromatic quena will be ebony, the pipe is then considered as a cylindrical pipe. The first prototypes presented in this study are built in plastic tube which are also considered as cylindrical.

In a first step, the length correction of the notch is deduced by comparing a realization of a pipe without and with the notch. The model allows to optimize the geometries of the knot and of the cylinder total length to obtain the desired pitch of G4. In a second step, side holes are included and their positions are manually optimized for proper intonation of the first two octaves, the most pertinent geometry of the resonator is searched. Lastly, the fingerings of the third register are established.

3.1 Notches, knot and length cylinder

The notch constitutes the active entrance of the quena, three shapes of notches can be listed: in "U", in "V" and the "little notch". A study measuring the playing range for each shape of notch by L. Garcia [15], shows that a notch with a little window allows a wider range of playing frequencies. On the other hand, the orientation of the labium is to be considered. In fact, when the labium is oriented towards the outside, Garcia notices an expansion of the playing range. The notch selected is a notch between a shape in "V" and a "little notch". The dimensions of the notch are: 9 mm long and 10 mm wide. The labium is carved in the thickness of the tube to 1/3 from the outside and to 2/3 on the inside. This geometry of the notch was selected after observation of traditional diatonic Chilean quenais, to offer an easy adaptation of the blowing technique of quena players as well as similar sounding aesthetics.

It is important to start the study with the notch because it determines the open area at the blowing end, and therefore the corresponding length correction. Its shape strongly influences the resonance frequencies of the cylinder. As a first approximation, the length correction is considered as constant for all the frequency band. The evolution of the length correction according to the register or flutist's position will be discussed section 3.3. In the same way quena-makers do, the notch is carved on a relatively long cylinder which is played, the playing frequency of the pipe with a notch is thus obtained. By comparing this playing frequency of the pipe with the notch and a model of the same pipe without the notch, a length correction l_h can be associated to the notch. In this case it's equivalent to $l_h = 3.4$ cm. Then the knot's geometry can be determined.

The knot is included in the model as discussed in section 2 with two geometric parameters, its length and cross-section. The G4 – 5's harmonicity can be studied as a function of the latter two using our model. G4 – 5's harmonicity changes according to the length or the section of the knot as shows the figure ???. A small variation of 1 mm is added successively on the radius and on the length of the knot. The length of the cylinder is thus adapted to obtain the same pitch of G4 and so making more legible reading the inharmonicity.

The inharmonicity is defined by Debut [16] as the relative difference between the resonance frequency f_n and n times

knot's geometry	f_1 (Hz)	f_2 (Hz)	IH (cents)
$r_b=5$ mm, $l_b=4$ mm	391.7	786.2	6.2
$r_b=5$ mm, $l_b=6$ mm	391.7	787.0	7.9
$r_b=4$ mm, $l_b=4$ mm	391.7	787.4	8.8

Table 1: Table showing the calculating of the inharmonicity of G4 and 5 according to the knot's geometry in a pipe

the first resonance f_1 :

$$IH = \frac{f_n - n f_1}{n f_1} \quad (10)$$

$$IH(\text{cents}) = \frac{IH}{5.78 \times 10^{-4}} \quad (11)$$

f_1 (G4), f_2 (G5) and IH values are referenced in table 1 according to the knot's geometry in a pipe.

The table 1 indicates that G4 – 5 inharmonicity increases when the length of the knot increases. The same for a decreasing cross section. To obtain an equal timbre on each notes of the scale, the geometry of the last open hole should be identical for each fingering. Thus for the G4, the last open hole is the knot, its dimensions should approach the geometry of the side holes. A radius $r_b = 5$ mm and a length $l_b = 4$ mm is a right compromise between the tuning and the geometry of the knot. Indeed the knot's radius is identical to that of the last holes and the knot's thickness approaches the one of the pipe. This knot induces an inharmonicity of about 6 cents between G4 and G5. Lastly, to obtain in the model the resonance frequency of the G4, the length of the pipe should be $L = 37.9$ cm.

3.2 Determination of a simple profile tuning to the equal temperament chromatic scale: the first two registers

Criteria/condition of the tuning In the first step of the design, a simple profile tuning to the equal temperament chromatic scale is sought for the first two registers only. This first tuning is based on resonance frequencies, thus on the geometric parameters (toneholes and cylinder). The input admittance of the quena is calculated following the theory presented in section 2. The algorithm developed during this study is validated by comparison with other software packages, as well as for simple geometrical test cases allowing an analytic resolution. Geometric data of existent traditional diatonic quena are also studied with the program. The first geometric model of a chromatic quena is thus based on resonance frequencies. The two first requirements of the instrument maker are tonehole's diameters smaller than 10 mm by steps of 1/2 mm to comply with existing drill-bits.

When all the toneholes are closed, the instrument plays the lowest note: G4=392 Hz, and when the upper tonehole is open the pitch rises a semitone. This to the F#6=1480 Hz. The chromatic quena includes 11 toneholes, the fingerings are the same for the first two registers.

During the test of a new instrument, the player evaluates its quality firstly by the precision of the intonation. Then

the timbre, the sound and the ease to play are compared. Because of the different qualities searched by the players, these criteria are more difficult to quantify. Since the most important parameter is the intonation of the instrument, this study focuses on the pitch of the quena.

For an instrument like the quena, the fine tuning of the instrument maybe carried by small adjustments of the tone hole geometry in a final stage of tuning. Furthermore, the individual features of each player's lips also influence the tuning. Therefore, the accuracy in tuning that we expect is restricted in the following way, the frequency criteria are established by:

α – Tuning based on the equal temperament chromatic scale A4=440 Hz

β – Maximum frequency deviation of the tuning across two octaves: $\delta f_{max} - \delta f_{min} = 25$ cents

γ – Frequency interval between two consecutive fingerings <4 cents: $\delta f_{n+1} - \delta f_n < 4$ cents where δf corresponds to the interval in cents between resonance frequencies and frequencies of the scale. The last criterion corresponds to the fact that the player may easily adjust his lip position for the pitch correction, and develop a global pitch correction through the whole compass of the instrument. A local compensation between two notes in a small interval is much less easy to achieve.

Tuning profile: After determining the length correction of the notch, the geometry of the knot and the length of the pipe, the 11 toneholes are arbitrarily placed on the pipe with diameters between 7 to 10 mm. This allows to take into account the small volumes in the toneholes when they are closed. Determining the geometry with the algorithm begins with the tonehole 11, the closest to the knot at the passive end. Because we consider large sides holes, adjustments of diameters or positions of this hole have a very little influence on the tuning of the notes lower. The tuning is carried from the tonehole 11 to the first. The geometric determination of tonehole's position and diameters follows three steps:

- the frequency to the first impedance peak is approximately determined by the position of the hole
- the diameter is adjusted for best harmonicity between the first two registers
- the hole's position can be modified to obtain the exact desired frequencies

The fixed parameters are the temperature to $t = 22^\circ C$, the cylinder's radius to $a = 9.5$ mm and the thickness of the pipe $h = 2.95$ mm. The simplest intonation profile obtained is shown in figure 3. This profile corresponds to the instrument's geometry in figure 4. The first intonation profile, figure 3, respects the constraints explained previously and shows a certain regularity between the first two registers. A little gap appears on the G5 due to the knot. The first register is tuned around zero and the second register is a simple decay.

3.3 Link between resonance and sounding frequencies

A deduction of a new intonation profile: As discussed in section 2, there is a shift between resonance and sound-

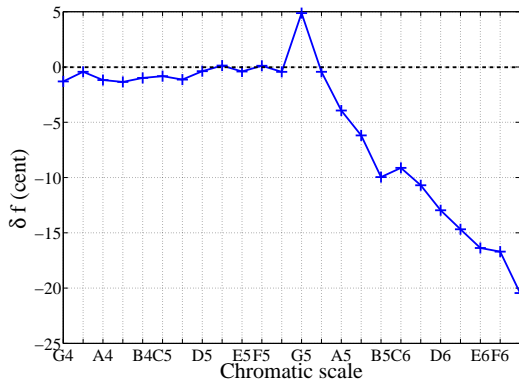


Figure 3: Calculation of the frequency interval (δf) between the intonation profile and the equal temperament chromatic scale with A4 tuned at 440Hz for the geometry of the quena defined in figure 4. Length correction at the notch is taken as 3.4 cm and the temperature $t = 22^\circ\text{C}$ in the calculation. The intonation profile presented corresponds to the resonance frequencies.

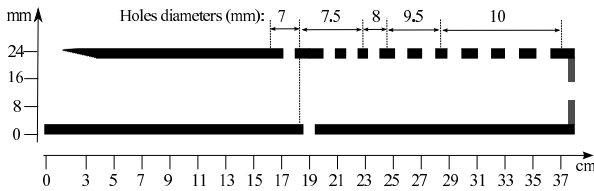


Figure 4: Figure showing the global geometry of the quena in cross-section view.

ing frequencies. Auvray [2] and Meissner [3] show with experiments and models the active behavior of a flute depending on the speed of the jet. Coltman [5] and Cossette [4] explain how the expert musician can control the sounding frequency by approaching the labium resulting in a modification of the open area. To highlight the difference between sounding and resonance frequencies, the instrument described in figure 4 is built in a plastic tubes and played. The interval between sounding frequencies and the tempered scale is represented in figure 5. A target profile at fixed length correction at the lips is then deduced, that corresponds to the difference of resonance and sounding frequencies: $\delta f_{th} = \delta f_{res} - \delta f_{sound}$.

Figure 5 shows that the sounding frequencies follow the resonance frequencies in the lowest part of the first register up to $D\#5$. Next, for the last notes of the first register they are somehow flat. For the second register, they are in average 5 cents sharp. The target profile shows the intonation profile to achieve in order to obtain sounding frequencies closer to the desired scale. A new geometry is then developed.

Determination of the new geometry: The aim is now to determine a new geometry of the quena to approach the target profile. For that, the previous geometry is slightly modified. The constraints of the tuning are the same as above. In a simplistic way, an octave can be reduced by increasing the diameter of the hole and vice versa. The cylinder's di-

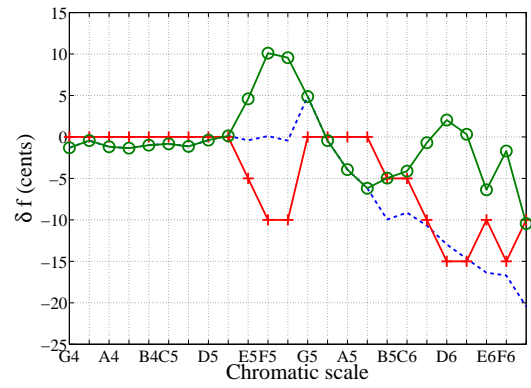


Figure 5: Calculation of the frequency intervals (δf) between the equal temperament scale and the resonance frequencies (dashed), the sounding frequencies (crosses) of the instrument described in figure 4. The target profile is deduced from the difference between the latter two (circles).

mensions are the same: radius, thickness, length, knot and notch. Temperature is also the same. The new geometry of the instrument is plotted on figure 6, highlighting the modifications from the previous geometry.

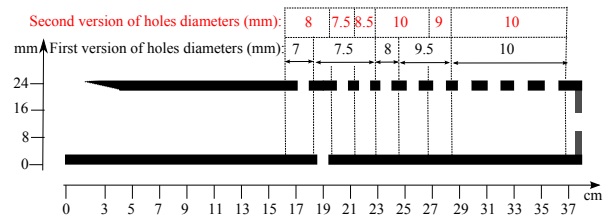


Figure 6: Illustration showing the modifications on tone-hole's geometry (positions and diameters) between the first and the second intonation profiles on the first two registers.

Comparing figures 4 and 6, it is possible to note that the geometry of the lowest holes was not changed. In fact, the tuning of the lowest notes in the scale was correct. The diameters of some upper toneholes were enlarged somewhat to match the desired tuning between the two registers. The intonation profile obtained for this second geometry is presented in figure 7.

The second intonation profile is close to the target tuning up to $A\#5$. For the upper notes, the second intonation profile is a bit lower, maximum of 10 cents. The interval between the resonance frequencies and the target profile for the upper notes (from the $B5$) is reduced of approximately 9 cents. The same between $E5$ and $G5$, the interval is reduced of 6 cents. To further reduce this gap, it would be necessary to increase the diameter of the first holes, but this would not be compatible with the key mechanism. However, these differences are lie inside the desired values presented in section 3.2 and regular, thus easier to control in play.

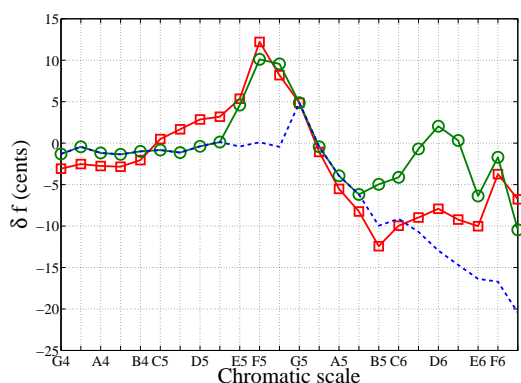


Figure 7: Calculating of the frequency intervals (δf) between the tempered scale and the resonance frequencies from the first geometry described in figure 4 (dashed): already presented in figure 5 is plotted for reference, the target profile deduced from interval between the sounding and resonance frequencies of the first geometry (circles): also presented in figure 5 is plotted for reference and the resonance frequencies from the second geometry described in figure 6 (squares). For the two geometries, the length correction of the notch is 3.4 cm and the temperature $t = 22^\circ C$.

3.4 The third register

The scale from the $G6$ to the $F\#7$ constitutes the third register of the quena. This register isn't played with the same fingerings than the first two octaves. The Boehm's key-mechanism is also adapted to perform this last register. The quena's fingerings are determined from modern flute's fingerings.

The register hole: On the modern flute, small diameter holes are placed between the mouthpiece and the tone holes. These register-holes help the oscillation of upper pipe resonances by reducing the first admittance peaks. In the case of the chromatic quena the hole the closest to the notch acts both as a tonehole for $F\#5 - 6$ and as a register hole for $G6$ and for notes in the third octave of the instrument.

Its position is important because it determines the tuning of $F\#5$ and 6, but also of $G6$. The geometry of this hole is deduced by a compromise between these three notes. The hole's position determines the tuning of $G6$, the diameter the tuning of $F\#5$ and 6. From the previous section, it can be estimated that $G6$ must be tuned between -5 and -10 cents relative to the equal temperament scale. A tuning of -7 cents corresponds with the calculation to a hole's position of 17.7cm. Then, the tuning and harmonicity of $F\#5$ and 6 can be studied according to the hole's radius. The target intonation profile estimates a deviation of $+10$ cents and -10 cents, respectively for the $F\#5$ and 6. The best compromise is to choose the diameter of 8 mm. In fact the deviation is then respectively to 8 cents and -7 cents. This compromise allows to approach best the equal temperament scale in play.

Fingering of the third register: The fingerings for the third register are determined by playing, based on modern flute fingerings. The second instrument described in figure 6 is built in a plastic tube, with the geometry of the register hole described in section 3.4. First, each modern flute fingering is played on the quena. If the pitch is sharper than the desired pitch, an other tonehole is closed from the passive extremity. The effective length of the instrument is then extended and the pitch falls a few cents. These fingerings must obviously comply with the Boehm's key-mechanism.

The resonance frequencies can then be calculated for each fingering. The intervals between the resonances frequencies and the equal temperament scale are represented in figure 8 for the three registers.

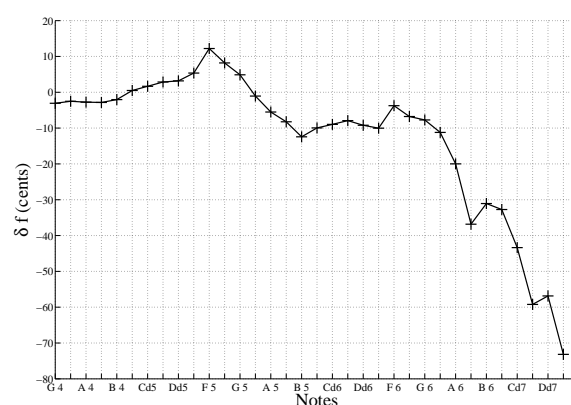


Figure 8: Calculating of the frequency intervals (δf) between the equal temperament scale and the resonance frequencies of the three octaves. These fingerings of the third register are determined by playing the instrument described by the geometry 6 and with the register-hole's geometry described in section 3.4.

Figure 8 shows that the higher the note is, the lower resonance frequency compared to the equal temperament scale is. Playing in the third register requires high jet velocities. This has been observed on modern flute with an increasing of the value of the sounding frequency. However, on modern flute the musician compensates a high air velocity with the distance to the labium, [17].

4. CONCLUSION AND EXPECTATIONS

The first prototype of a new chromatic quena has been designed with a method similar to the one of the flute-maker. The first step consists in observing different diatonic quenans and duplicating them. Then, 1D linear acoustic is used to determine the geometry to be closer to the desired scale. 1D linear acoustic allows to obtain interesting and relevant information on resonance frequencies. However, there is a difference between resonance and sounding frequencies (obtained while playing). To identify the interval between the last two, the first model of quena must be played. It is then observed that this interval increases with the pitch of the note played. The length correction of the notch changes with the frequency, probably due to the forward motion of

the lips towards the labium. It would be interesting to develop the work in this direction, in order to determine how the notch, the position of the lips, the velocity of the jet are involved in the interval between resonance and sounding frequencies.

1D linear acoustic can give efficient information on resonance frequencies, it offers a global view of the influence of geometric modifications, since it presents an intonation profile showing the predicted pitch variations for all the notes. However, no single parameter allows to quantify the sound, the timbre, the intensity or the dynamic. The increase of the side-holes number reduces the distance between each hole. This affects the cutoff frequency of the instrument, which may change the timbre. These parameters are also interesting, effectively the timbre, the sound of the instrument and the ease to play are criteria in the appreciation of an instrument. 1D linear acoustic does not allow to quantify these features and this is the most important limitation of 1D linear acoustic. Therefore, the preservation of the notch and the knot, regular hole's diameters, the timbre of the chromatic quena should remain in close reference to the diatonic quena timbre.

Finally, the 1D linear acoustic allows to help the flute-maker, mostly by speeding up the observation phase and the first phases of geometry determination. Indeed, this design process based on a traditional process, but includes 1D linear acoustic to reduce the number of prototypes to be actually built. However it doesn't replace the later phases of adjustments by the flute maker. Moreover, the quena chromatic will be realized in wood with a key-mechanism. The influence of the keys or of the fingers is not included in this model. It may be possible to include them with a correction term.

acknowledgments The authors would like to thank Laurent Quartier, Arnaud Gérard and J.Y Roosen, the flute-maker, for his suggestions, discussions.

5. APPENDIX

From the plane wave to the matrix transfer of the cylinder: Linear acoustic is based on the equation of loss-less acoustic waves in air:

$$\Delta p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} = 0 \quad (12)$$

where p is the pressure and c_0 is the velocity of sound.

In a cylindrical tube of radius a , at frequencies under the cutoff frequency $f_c = 1.84 \frac{c_0}{2\pi a}$, the general solution is a plane wave, written following d'Alembert as the sum of two progressive waves, $f^+(x - ct)$ and $f^-(x + ct)$ moving respectively, in the $+x$ and $-x$ directions with a velocity c_0 :

$$p(x, t) = f^+(x - c_0 t) + f^-(x + c_0 t) \quad (13)$$

This equation becomes in the frequency domain:

$$P(x, \omega) = P^+(\omega)e^{-jkx} + P^-(\omega)e^{jkx} \quad (14)$$

where $k = \omega/c_0$ is the wave number. Euler's equation allows to obtain the acoustic flow:

$$U(x, \omega) = Z_c^{-1}[P^+(\omega)e^{-jkx} - P^-(\omega)e^{jkx}] \quad (15)$$

with $Z_c = \rho_0 c_0 / S$ the characteristic impedance of the cylinder if ρ_0 is the air density and S the cross-section of the cylinder. Acoustic pressure and flow at any given position x in the cylinder can be written as functions of the pressure and flow at $x = 0$, $P_0 = P(0, \omega)$ and $U_0 = U(0, \omega)$ as:

$$\begin{cases} P(x, \omega) = P_0 \cos kx - U_0 Z_c j \sin kx \\ U(x, \omega) = -P_0 Z_c^{-1} j \sin kx + U_0 \cos kx \end{cases} \quad (16)$$

Equation (16) can be generalized for the abscissa x_1 and x_2 with $x_2 - x_1 = L$, the length of the cylinder:

$$\begin{pmatrix} P_1 \\ U_1 \end{pmatrix} = \begin{pmatrix} \cos kL & Z_c j \sin kL \\ Z_c^{-1} j \sin kL & \cos kL \end{pmatrix} \begin{pmatrix} P_2 \\ U_2 \end{pmatrix} \quad (17)$$

Introducing the "transfer matrix" from the position x_2 to the position x_1 . This "transmission line" formalism will be used throughout the paper.

Two major mechanisms of attenuation are present in a wave-guide: viscosity and thermal conduction effects which are mainly effective close to the inner surface. The theory from Zwikker and Kosten, [18] allows to keep using the plane wave description with uni-dimensional equations considering averaged acoustic quantities. The averaging is carried over the cross section of the cylinder. This theory dissociates the viscous and thermal effects and leads to a formulation by transmission lines, as explained in the books [6], [7].

$$\frac{dP}{dx} = -Z_\nu U \quad \text{and} \quad \frac{dU}{dx} = -Y_t P \quad (18)$$

where

$$Z_\nu = \frac{j\omega\rho_0}{S} \left[1 - \frac{2}{k_\nu a} \frac{J_1(k_\nu a)}{J_0(k_\nu a)} \right]^{-1}, \quad (19)$$

$$Y_t = j\omega\chi_S S \left[1 + (\gamma - 1) \frac{2}{k_t a} \frac{J_1(k_t a)}{J_0(k_t a)} \right] \quad (20)$$

Z_ν and Y_t are respectively the linear impedance in series and the linear admittance in parallel, corresponding to the viscous and thermal conduction effects. J_0 and J_1 are the Bessel functions and χ_S the adiabatic compressibility.

Then the solutions of transmission lines become:

$$P(x, \omega) = P^+(\omega)e^{-\Gamma x} + P^-(\omega)e^{\Gamma x} \quad (21)$$

$$U(x, \omega) = Z_c^{-1}[P^+(\omega)e^{-\Gamma x} - P^-(\omega)e^{\Gamma x}] \quad (22)$$

where Γ is the propagation constant ($\Gamma = jk_c$, and k_c the complex wave number). The transfer matrix of the cylinder can be expressed by:

$$\begin{pmatrix} P_1 \\ U_1 \end{pmatrix} = \begin{pmatrix} \cosh \Gamma L & Z_c j \sinh \Gamma L \\ Z_c^{-1} j \sinh \Gamma L & \cosh \Gamma L \end{pmatrix} \begin{pmatrix} P_2 \\ U_2 \end{pmatrix} \quad (23)$$

6. REFERENCES

- [1] J. Wolfe, J. Smith, J. Tann, and N. Fletcher, "Acoustic impedance spectra of classical and modern flutes," *Journal of Sound and Vibration*, vol. 243, no. 1, pp. 127 – 144, 2001.

- [2] R. Auvray, B. Fabre, and P. Lagree, "Regime change and oscillation thresholds in recorder-like instruments," *The Journal of the Acoustical Society of America*, vol. 131, no. 2, pp. 1574–1585, 2012.
- [3] M. Meissner, "Aerodynamically excited acoustic oscillations in cavity resonator exposed to an air jet," *Acta Acustica united with Acustica*, vol. 88, no. 2, pp. 170–180, 2002.
- [4] I. Cossette, B. Fabre, V. Freour, N. Montgermont, and P. Monaco, "From breath to sound: Linking respiratory mechanics to aeroacoustic sound production in flutes," *Acta Acustica united with Acustica*, vol. 96, no. 4, pp. 654–667, 2010.
- [5] J. Coltman, "Resonance and sounding frequencies of the flute," *The Journal of the Acoustical Society of America*, vol. 40, no. 1, pp. 99–107, 1966.
- [6] A. Chaigne and J. Kergomard, *Acoustique des instruments de musique*, ser. Collection Echelles. Belin, 2008.
- [7] A. Pierce, *Acoustics, an introduction to its physical principles and applications*. Acoustic Society of America, New York, 1989.
- [8] C. Nederveen, J. Jansen, and R. Hassel, "Corrections for woodwind tone-hole calculations," *Acta Acustica united with Acustica*, vol. 84, pp. 957–966, 1998.
- [9] J. Dalmont, C. Nederveen, V. Dubos, S. Ollivier, V. Meserette, and E. te Sligte, "Experimental determination of the equivalent circuit of an open side hole: Linear and non linear behaviour," *Acta Acustica united with Acustica*, vol. 88, no. 4, pp. 567–575, 2002.
- [10] V. Dubos, J. Kergomard, A. Khettabi, J. Dalmont, D. Keefe, and C. Nederveen, "Theory of sound propagation in a duct with a branched tube using modal decomposition," *Acta Acustica united with Acustica*, vol. 85, no. 2, pp. 153–169, 1999.
- [11] D. Keefe, "Theory of the single woodwind tone hole," *The Journal of the Acoustical Society of America*, vol. 72, no. 3, pp. 676–687, 1982. [Online]. Available: <http://link.aip.org/link/?JAS/72/676/1>
- [12] A. Lefebvre, "Computational acoustic methods for the design of woodwind instruments," Thesis, Mc Gill University, 2010.
- [13] P. Eveno, "L'impédance d'entrée pour l'aide à la facture des instruments de musique à vent: mesures, modèles et lien avec les fréquences de jeu," Thesis, UPMC, 2012.
- [14] G. L. Vey, "A non optimization-based method for reconstructing wind instruments bore shape," *Acoustics 2012*, 2012.
- [15] L. Garcia, "Rôle de l'embouchure dans la sonorité des flûtes à encoche," Master's thesis, CNSM, Juin 1996.
- [16] V. Debut, J. Kergomard, and F. Laloë, "Analysis and optimisation of the tuning of the twelfths for a clarinet resonator," *Applied Acoustics*, vol. 66, pp. 365–409, 2005, debut_AA - OR 20 debut_AA - OR 20. [Online]. Available: <http://hal.archives-ouvertes.fr/hal-00000589>
- [17] P. de la Cuadra, B. Fabre, N. Montgermont, and C. Chafe, "Analysis of flute control parameters: A comparison between a novice and an experienced flautist," *Acta Acustica united with Acustica*, vol. 94, no. 5, pp. 740–749, 2008.
- [18] C. Zwikker and C. W. Kosten, *Sound absorbing materials*. Elsevier, New York :, 1949.

Brass instruments

LIP MOTION, THE PLAYING FREQUENCY OF THE TROMBONE AND THE UPSTREAM AND DOWNSTREAM IMPEDANCES

Henri Boutin
University of
New South Wales,
Sydney

Neville Fletcher
Australian National
University,
Canberra

John Smith
University of
New South Wales,
Sydney

Joe Wolfe
University of
New South Wales,
Sydney

henri.boutin@unsw.edu.au, neville.fletcher@anu.edu.au
john.smith@unsw.edu.au, j.wolfe@unsw.edu.au

ABSTRACT

We report the motion of trombone players' lips, its phase with respect to the mouthpiece pressure, the impedances of the bore and the player's vocal tract, and the frequency difference between the bore resonance and played note. The bore resonance frequency shifts very little with playing and often decreases somewhat: the effect of CO₂ can exceed that of temperature and humidity. The bore impedance is usually compliant for the note B \flat 2 and inertive for B \flat 3. The vocal tract impedance measured at the player's mouth is inertive for both notes. In terms of Fletcher's simple model for regeneration (JASA, 93, 2172), the results are consistent with a (+1, -1) valve for B \flat 2 and (-1, +1) for B \flat 3. The pressure in the mouthpiece in both cases rises before the lips separate. For B \flat 3, where the lip motion is mainly transverse, this is consistent with the inertive load. For B \flat 2, the substantial motion of the lips in the direction of flow provides a sweeping motion which produces the current into the bore that precedes lip opening.

1. INTRODUCTION

How and why do the lips of a brass player vibrate? This is an interesting question in music acoustics, with potential applications in music pedagogy and performance. Images of the motion as functions of time are an important contribution to answering the question. Various studies, including the present one, have used stroboscopy or high speed video for this purpose: Copley and Strong [1], Yoshikawa and Muto [2], Tarnopolsky et al [3], Campbell and colleagues [4] and others.

In mathematical models of the vibrating lip, the acoustic impedance spectra of the upstream and downstream side of the lip are important [5, 6]. Simultaneous measurements of the acoustic pressure in the mouth and that in the mouthpiece during trombone playing [7] have shown that the ratio of the former to the latter increases strongly as the players play higher, but varies among players. This increase is in part due to the way the impedance maxima of the bore decrease with increasing frequency. However, the variation

among players strongly suggests that they use their vocal tracts in different ways.

Measuring the upstream impedance Z_{mouth} — the impedance of the player's vocal tract measured near the lips — is non-trivial: the signal produced by the vibrating lip is much more powerful than the probe signal used to measure Z_{mouth} , because the energy of the probe signal has to be divided among hundreds of different frequencies. Nevertheless, this has been reported for the trumpet [8] and the didjeridu [3]. Measuring the downstream impedance Z_{bore} also has challenges because: (i) the instrument operates close to an impedance peak, (ii) the impedance peaks are quite narrow and (iii) the speed of sound and thus the resonant frequency depends subtly on temperature, humidity and the concentration of CO₂ inside the bore. So it is necessary to measure Z_{bore} under conditions close to playing so as to understand the combined effect of temperature, humidity and CO₂. The relative timing of the lip motion and the pressure in the mouthpiece is also important. The timing of lip contact can be measured by the high frequency electrical admittance y_{lip} between upper and lower lips [9, 10].

In order to explain how the lip motion is driven at different playing frequencies, the objectives of the present study are: to investigate the effects of temperature, humidity and concentration of CO₂ on the bore resonances, to measure the mouth and bore impedances during playing, and to compare the lip contact with the pressure in the mouthpiece.

2. MATERIAL AND EXPERIMENTAL SETUP

We report the measurements of the impedance of the trombone bore, the impedance of the player's vocal tract, the pressure in the mouthpiece and a variable related to the opening area between the player's lips. These signals were measured using three different experimental setups described in the next paragraphs. Six amateur players were experimental subjects.

2.1 The trombone and the mouthpiece

All the players played same modified trombone (Yamaha model YBL 321). This instrument has a valve that converts it from the tenor range to the bass range, but the valve was left in the tenor position. In all experiments, the slide was

kept in the same position (first position: all the way in) and the tuning slide was extended 18 mm. The instrument was clamped to a lab bench but mechanically isolated from it with foam.

The players involved in these experiments were asked to play up to five of the lowest notes above the pedal note: B \flat 2 (nominally 117 Hz), F3 (175 Hz), F \flat 3 (233 Hz), D4 (294 Hz) and F4 (349 Hz). However, in this brief report, only the data for two notes are analysed: B \flat 2 (nominally 117 Hz) and B \flat 3 (233 Hz).

The mouthpiece was replaced with an experimental mouthpiece (Figure 1) having the same cup volume (11.6 cm³), diameter (2.53 cm) and throat geometry as the original mouthpiece. The mouthpiece is transparent with two flat glass windows to allow high-speed video and stroboscopic observation of the motion of the lips. In a hole of diameter 2.3 mm in one side is sealed an *Endevco* piezoresistive pressure transducer (model 8507C-2). Two brass electrodes are set into the rim of the mouthpiece, so that they contact the face above the upper lip and below the lower. They are connected to the input of an electroglottograph (*EGG*, model EG2-PCX2) from *Glottal Enterprises*.

2.2 The bore impedance

For the temperature studies, the capillary method is used to measure the bore impedance Z_{bore} so that it could be sealed to the mouthpiece very quickly. The impedance head comprises an acoustic current source with high impedance and a microphone. The current source is similar to that used by Wolfe et al [11] and comprises a truncated cone concentric with a conical hole, but separated from it by three 100 μ m wires placed at 120° angles. The source and a microphone (*B&K* model 4944A) are located side by side in the plane against which the trombone mouthpiece is clamped. The impedance head is calibrated with a semi-infinite pipe of length 142 m and of diameter 7.8 mm.

For calibration and measurements, a loudspeaker generates a broadband acoustic current at the reference plane. This contains frequency components ranging between 50 Hz and 600 Hz with a resolution of 0.67 Hz ($= 44.1 \text{ kHz}/2^{16}$). During a calibration iteration, the amplitudes of each of the spectral components of this current are equalized [11].

2.3 The vocal tract impedance

The method used to measure the vocal tract impedance of the trombone player is the three microphone technique using non-resonant loads, described by Dickens et al [12]. Its advantage over the capillary method for this application is that a more powerful probe signal is possible, and this is helpful in the presence of the broadband noise measured close to the player's lips.

We use our smallest impedance head (of outer diameter 4.8 mm) to minimise the perturbation of the player, who is asked to hold this pipe between the lips at the corner of the mouth during playing, while orienting it so that the measurement plane is between the upper and lower teeth and behind the central portion of the lips. A flange on the head

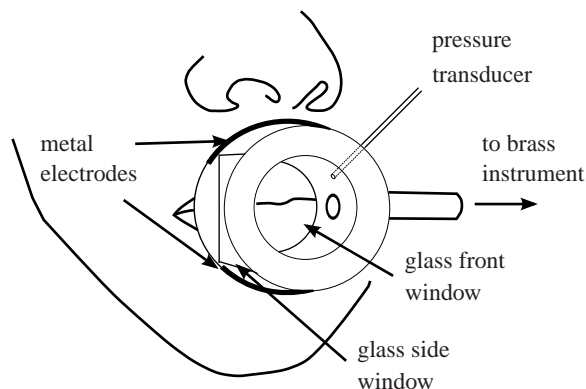


Figure 1. The modified mouthpiece equipped with two brass electrodes connected to an electroglottograph (*EGG*) to measure lip contact after Wolfe and Smith [9], and Freour [10]), a piezoresistive pressure transducer connected to the mouthpiece cup. Its design follows Ayers [13]: the two glass windows provide undistorted front view and side view of the lip motion.

at 25 mm from the measurement plane limits the maximum insertion into the mouth. The probe signal contains 34 cycles of length 2^{16} points (about 1.49 s), so that the measurements can last up to 50 s. 820 frequencies spaced at 0.67 Hz cover the range 50 Hz to 600 Hz.

2.4 The pressure in the mouthpiece, the lip motion and their relative phase

The pressure transducer in the mouthpiece, flush with the wall of the cup, is connected to a bridge amplifier and then to a sound card (*MOTU*, model *Firewire* 828). The sound card gain was calibrated using a known voltage source (*Topward* model 8120) of variable frequency. Then we replaced the transparent back plate of the mouthpiece by a plug in epoxy resin, equipped with a reference microphone (*B&K* 4944A). The mouthpiece was then driven with a broadband signal over the range [50 Hz, 1 kHz] and the ratio between the spectrum of signals from the pressure transducer and the reference microphone recorded. In the following experiments, with the transparent window replaced, this ratio determines the pressure in the mouthpiece cup from the transducer signal.

The *EGG* connected to the brass electrodes provides a signal negatively correlated with the electrical admittance between the two lips. It has no DC component, is maximum when the lips are open and minimum when they are closed. Thus it is correlated with and approximately in phase with the opening area between the lips, so we call it $OV(t)$ (for *open variable*). In order to evaluate the delay introduced by the *EGG* we connected a voltage-controlled resistance at its input. The delay (0.1 ms between 100 Hz and 1 kHz) is not negligible compared to the period of the notes played (≈ 10 ms), but this difference can be added when comparing the phase difference between the pressure in the mouthpiece and the lip opening area as given by the outputs of the piezotransducer and the *EGG*.

3. TEMPERATURE AND GAS EFFECTS ON THE RESONANCES OF THE BORE

As the trombonist starts to play, the temperature, the humidity and the concentration of CO_2 all increase in the bore. How do the impedance peaks vary with playing? We measured the input impedance of the trombone before the instrumentalist plays, and after he has played for 3 s, 10 s, 30 s and 4 min. Between measurements, the air inside the bore was flushed, and the bore impedance was measured again. During the 3 s and 10 s experiments, the player was asked to take a breath and to play a single sustained note. These experiments reproduce the playing conditions met by the player when his instrument is initially dry and at ambient temperature (26.8°C and relative humidity 57% in this instance). In the longer experiments (30 s and 4 min) the performer was asked to play several long notes and to inhale at will. The impedance head, coated with a thin layer of petroleum jelly, was sealed to the mouthpiece rim, within 3 s of the end of the playing and the impedance measured for successive cycles over the following 50 s. This protocol was followed once or twice with every player. The impedance spectra were sampled at 0.67 Hz. Around the peaks, these data were interpolated with a fitted cubic function, giving an estimated frequency resolution of 0.1 Hz. The mean curves corresponding to each playing duration are shown in Figure 2.

For all playing conditions, the second, third and fourth impedance peaks fall within the ranges 113.0 – 113.9 Hz, 170.9 – 172.1 Hz, and 227.7 – 229.4 Hz. These are the peaks used to play B \flat 2, F \sharp 3 and B \flat 3 respectively. For each, we note that the peak frequencies and amplitudes decrease after playing for 3 s. After 10 s and 30 s of playing, the frequencies are slightly greater than after 3 s and the amplitudes almost unchanged. After 4 min, the peak frequencies return approximately to their initial values, while the amplitudes decrease further. The same trend was observed on the 9 peaks between 50 Hz and 600 Hz. The changes in their amplitudes and frequencies are plotted on Figure 3.

The repeated measurements made before playing on different days are always located at less than 0.26% from their average frequency, with a standard deviation of 0.09%. Figure 3 shows that, after playing for 3 s and 10 s, the peak frequencies decrease in average by 0.48% and 0.62% respectively. Because humidity and temperature both increase the speed of sound, we conclude that, in these experiments, increased concentration in CO_2 in the bore more than compensates for the increases in temperature and humidity. This is not surprising: 3 s is easily enough time for the player to replace the air in the bore, but probably not enough time for the air to warm the instrument. Rises in water concentration are limited by the temperature.

In the 30 s-experiment, the players inhaled at least once between notes. So the CO_2 ceased increasing. This would explain why the peak frequencies stopped decreasing. In the 4 minute experiment, the players breathed several times as well, and there was time for temperature and thus water concentration to rise. Here, resonant frequencies differed little from the dry, ambient conditions: an average increase of 0.02%.

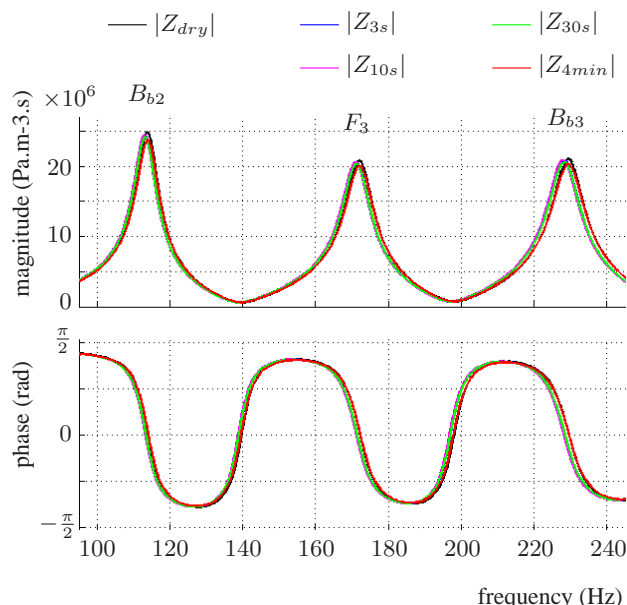


Figure 2. This graph shows several measurements of the impedance of the bore of the trombone equipped with the experimental mouthpiece when the slide is in 1st position and the trigger valve is not depressed. The three peaks in magnitude are the 2nd, 3rd and 4th resonances of the bore. The corresponding notes are labelled on the graph. Z_{dry} is the average curve obtained before playing; Z_{3s} , Z_{10s} , Z_{30s} and Z_{4min} are the average curves obtained after the musicians played the notes they wished for these durations. All measurements were carried out the same day. The temperature and the relative humidity at the time were respectively 26.8°C and 57%.

The decrease in the amplitude of the impedance peaks can possibly be attributed to humidity. We expect that gas composition would have very little effect on the characteristic impedance $\rho c/A$. However, as suggested by Coltman [14], air near 100% humidity could, in a standing wave, evaporate and condense during each cycle, which would increase losses. This could explain why after playing 4 min, the minima in Z_{bore} increase in magnitude (by 3.3% in average) while the maxima decrease (by 3.8% in average).

4. THE VOCAL TRACT IMPEDANCE

Among the previous studies focusing on the brass players' lips, two different lip motions in different ranges are reported [1, 2]. In the lower register, components of motion parallel and perpendicular to the flow have similar amplitudes, while in the high register the parallel component is much smaller. A simple model by Fletcher [6] associates qualitatively different kinds of auto-oscillatory lip valve motion with different values of reactive components of the upstream and downstream impedances.

Measurements of the upstream impedance spectrum used a small impedance head inserted in the corner of the mouth with its end lying between upper and lower teeth and close behind the lips. None of the players reported any difficulty to play the notes in this study (B \flat 2 and B \flat 3) with the

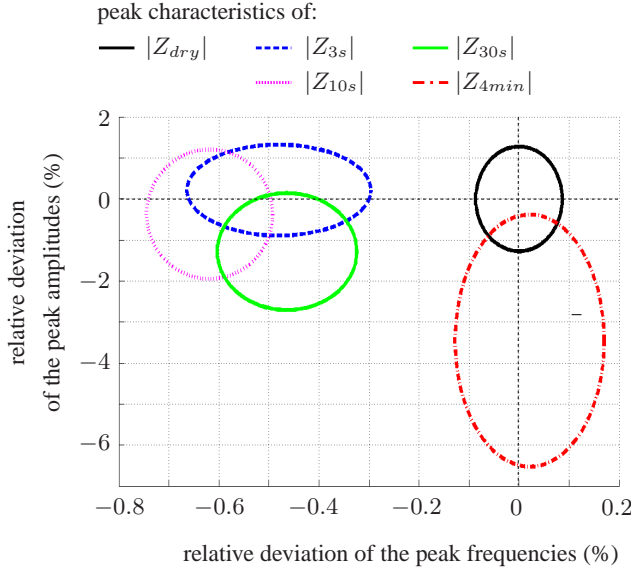


Figure 3. This plot displays the relative deviations (in %) of the impedance peak amplitudes and frequencies, after the musician played for 3s, 10s, 30s and 4min. The center of each ellipse is the mean of all measurements in each condition and the semi-axes show the standard deviations.

impedance head between the teeth and some could comfortably play several notes above these. We asked the players to play long notes, then to take another breath when they were out of air, and to start again. Each measurement lasted 30 s and allowed the subjects to play usually between 2 and 3 notes of 10 s. The experiment was carried out several times with each performer playing the lower pitch note Bb2 and then the higher pitch note Bb3. For each vocal tract configuration, the impedance curves obtained were fairly similar; two of them are displayed on Figure 4.

The players reported that they raised their tongue while playing a high pitch note. This would be expected to raise the characteristic impedance, could explain why Z_{mouth} is typically higher when the player is playing Bb3. (An analogous result was observed by Wolfe et al [15], using a mechanical trombone-playing machine with an outwards opening valve.)

Finally, we removed the impedance probe from the player's mouth and asked each player to play notes (Bb2 and Bb3) for about 10 s. We recorded the sound pressure in the mouthpiece, p_{mp} , and calculated its spectrum over a 1 s interval, in the sustained period, after the attack. We compare it with the bore resonances measured after 10 s, since it is approximately the duration of the notes played. Henceforth, Z_{bore} refers to Z_{10s} . f_p varies somewhat for different performers and samples. However, for all six players, f_p is always located above the bore resonance when the note played is Bb2 and below when it is Bb3. (These inequalities are true for all of the Z_{bore} spectra measured in this study under the various conditions described above.)

From Fletcher's simple model [6], a required condition

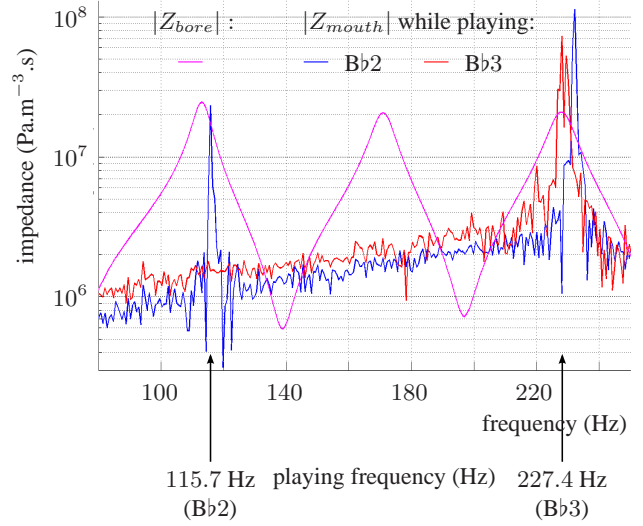


Figure 4. This plot shows the magnitude of the trombone impedance (after playing 10 s) and that of the vocal tract impedance of the musician while playing a Bb2, tongue in low position, and while playing a Bb3, tongue in high position. Both measurements used an impedance head (4.8mm-diameter) introduced in the mouth of the player. The narrow peaks on the curves of vocal tract impedance are artefacts due to the sound of the fundamental playing frequency and, for Bb2, its second harmonic.

for a valve to auto-oscillate is:

$$Im \left(\frac{W \bar{p}_1 \mu (\sigma_2 Z_2 - \sigma_1 Z_1)}{\sqrt{2\rho \bar{p}_1} + W \bar{x} (Z_1 + Z_2)} \right) > 2\pi f_p \times k, \quad (1)$$

where W is the valve width perpendicular to the flow, \bar{x} the static component of the valve opening and \bar{p}_1 the static pressure upstream of the valve. ρ is the air density, k the attenuation coefficient of the valve and f_p its playing frequency. μ is proportional to the valve flap area divided by the effective mass involved in the oscillation. The pair $(\sigma_1, \sigma_2) = (\pm 1, \pm 1)$ describes how the valve operates: $\sigma_1 = +1$ if a positive pressure upstream of the valve tends to open it further, and -1 if it tends to close it. σ_2 is similarly defined with respect to the pressure downstream of the valve. Finally Z_1 is the impedance 'looking into' the upstream side of the valve in the direction of the load, and Z_2 'looking into' the downstream side of the valve. Here they correspond to Z_{mouth} and Z_{bore} respectively. This simplified model assumes that the valve channel is negligible length, the static pressure downstream is zero and that the valve flaps never collide with each other and undergo simple harmonic motion. In addition it neglects the inertance of air inside the channel and the flow generated by the flap motion. Even though these simplifying assumptions are not satisfied by the lip valve, the model is still qualitatively helpful in discussing whether a vibration mode can auto-oscillate or whether it is driven by an other mode.

Our measurements show that the imaginary parts of the impedances (X_{mouth} and X_{bore}) have magnitudes comparable with the real parts (R_{mouth} and R_{bore}) at the playing frequencies, cf. Figure 5. This means that the losses in the

vocal tract and the bore and its radiation impedance cannot be neglected. Thus (1) leads to the following conditions:

$$\sigma_2 X_{bore} - \sigma_1 X_{mouth} > \alpha (X_{mouth} + X_{bore}) \times \dots$$

$$..(\sigma_2 R_{bore} - \sigma_1 R_{mouth}), \quad (2)$$

$$\text{with } \alpha = \frac{W\bar{x}}{\sqrt{2\rho\bar{p}_{mouth}} + W\bar{x}(R_{mouth} + R_{bore})} \quad (3)$$

Typical values of X and R at both playing frequencies studied are obtained from the measurements of Figure 5. This involves removing the very narrow peaks at the harmonics of the note being played and interpolating and smoothing the real and imaginary parts of the vocal tract impedance with a Savitsky-Golay filter:

f_p	Z_{bore} (MPa.m ⁻³ s)		Z_{mouth} (MPa.m ⁻³ s)	
	R_{bore}	X_{bore}	R_{mouth}	X_{mouth}
115.7 Hz (Bb2)	13.8	-11.1	0.95	0.76
227.4 Hz (Bb3)	20.4	3.5	2.6	0.98

Table 1. Typical values of the real and imaginary parts of Z_{bore} and Z_{mouth} at both playing frequencies 115.7 Hz (Bb2) and 227.4 Hz (Bb3)

Since R_{mouth} , R_{bore} and \bar{x} are positive, $\alpha > 0$ in (2). In the case of Bb2, $-X_{bore} > X_{mouth} > 0$. Then the only condition for (2) to be satisfied with any positive values of R_{mouth} is $(\sigma_1, \sigma_2) = (+1, -1)$. This is the regeneration mode for an *outward swinging valve*, or *swinging door valve*.

For Bb3, $X_{bore} > X_{mouth} > 0$. Then, a $(-1, +1)$ valve always satisfies the auto-oscillation condition with any positive value of the real part R_{mouth} . By giving plausible order of magnitude to the model parameters, $\bar{p}_1 = 2\text{kPa}$, $W = 1\text{cm}$ and $\bar{x}_1 = 1\text{mm}$, the simple model predicts that a $(-1, -1)$ valve could also auto-oscillate. But this also depends on the real part of Z_{mouth} . As a result, in the higher range, our results and this model together suggest that the brass player's lips probably do not operate as a $(+1, +1)$ valve but more likely as a $(-1, +1)$.

5. THE TIME SIGNALS

In the next experiment, we investigate the phase relationship between the pressure in the mouthpiece p_{mp} and the motion of the lips. As detailed in Section 2, the contact electrodes on the experimental mouthpiece allow us to measure the AC component of the electrical admittance between the lips y_{lip} . During the previous experiment, as we asked the performers to play Bb2 and Bb3, we measured y_{lip} and the pressure in the mouthpiece p_{mp} simultaneously. The signal shown on Figure 6 is inversely related to this admittance: it is a minimum when the contact area between the lips is maximum and *vice versa*. For this reason, we call this the open variable $OV(t)$. Copley and Strong [1] show that the lips of a trombonist close completely during each cycle, as do those of a hornist [2]. This was also observed by high speed video in the present study.

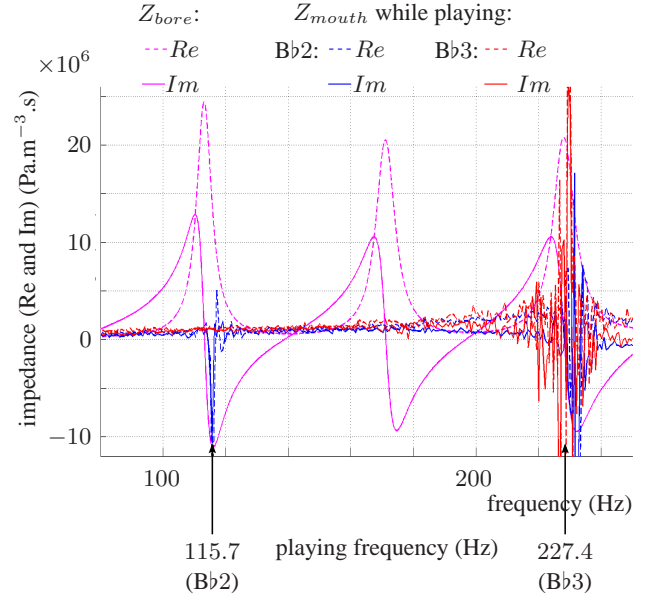


Figure 5. This plot shows the real and imaginary parts of the bore impedance (after playing 10s) and of the vocal tract impedance of the musician while adopting two different vocal tract positions: tongue in low position to play a Bb2 and in higher position to play a Bb3.

During this closed phase, air flow between the lips is zero, y_{lip} is maximum and OV is a minimum.

Figure 6 shows an example of $OV(t)$ and $p_{mp}(t)$. For each of the played notes, the pressure in the mouthpiece starts to increase before the lips begin to open and reaches its minimum value just before the lips close completely. This feature was observed for each note played and each performer. The air flow between the lips is zero when the lips are closed. When the lips are open, however, it is not simple to relate the flow to the lip opening or to OV .

For the higher played note Bb3, the impedance measurements show that Z_{bore} is inductive at the playing frequency (Figure 4), so one would naïvely expect the flow into the mouthpiece to lag the pressure p_{mp} . (In practice, the flow between the lips is broadly related to lip opening, but the motion is complicated by higher harmonics, whose phase relations may in principle be different from that of the fundamental.)

In contrast, for the lower note Bb2, Z_{bore} is compliant at the playing frequency and so the flow in the mouthpiece U_{bore} is expected to lead the mouthpiece pressure p_{mp} . This raises an obvious question: how can the flow into the mouthpiece begin while the lips are still closed?

Yoshikawa and Muto [2] provided images from side viewing of the lip motion in a transparent horn mouthpiece. They show that the upper lip moves outward significantly in the lower range and that the lips separate after they have moved forward into the mouthpiece.

Video images taken in the present project give similar results: for Bb2, the lip motion has a large component in the direction of the flow, and this motion leads the transverse motion in phase. In contrast, for Bb3 the motion is almost entirely transverse. For Bb2, this sweeping motion

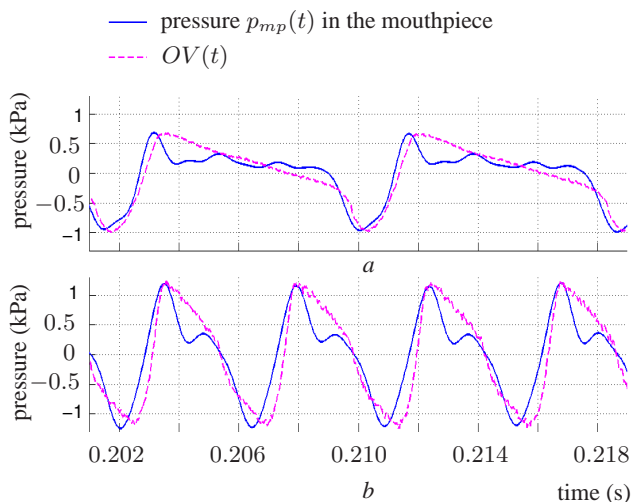


Figure 6. The pressure in the mouthpiece $p_{bore}(t)$ and the open variable $OV(t)$ (OV is inversely related to the electrical admittance between the upper and lower lip and is a maximum when the lips are separated and a minimum when they are in contact). (a) the musician played a lower pitch (Bb2) note, and (b), a higher note (Bb3)

generates a positive flow into the mouthpiece before the lips separate and before pressure begins increasing in the mouthpiece. In this preliminary account, we do not report the magnitudes of this ‘sweeping’ current, but simple calculations show that its magnitude is of the same order as the total acoustic flow into the mouthpiece. In contrast, for the higher note Bb3, the lip motion is almost entirely perpendicular to the flow, so U_{bore} essentially equals the air flow between the lips, and this current can start to increase only when the lips open.

6. CONCLUSIONS

These observations suggest that, for Bb2, the lips operate as an outwards opening valve and for Bb3, as a sideways sliding valve. Fletcher’s simple model [6] predicts the regeneration conditions for such valves in terms of the acoustic impedance upstream and downstream of the lip valve.

The upstream impedance, Z_{mouth} depends on the configuration of the player’s vocal tract and, at the playing frequencies reported here, it increases when the player raises his tongue.

According to the curves of Figure 2, the frequency and amplitude of the peaks in Z_{bore} depend only weakly on the duration of playing, probably because of compensating effects. So the changes in the frequency of these peaks are rather smaller than the displacement of the playing frequency from the peak in Z_{bore} .

Further measurements carried out with six performers showed that Z_{bore} is usually compliant when the note played is Bb2 and inertive when the note played is Bb3. According to a simple theoretical model of valve [6], this change of sign in the imaginary part of Z_{bore} explains why the lip valve cannot have the same operating mode for both played notes. The model predicts that the lip valve is likely to oscillate as a (+1,−1) valve in the low range and as a

(−1,+1) valve in the high range.

Plots of $p_{mp}(t)$ and $OV(t)$ (Figure 6) showed that, during playing, the pressure in the mouthpiece leads the lip opening. This result is explained for the higher note, when Z_{bore} is inertive, by flow between the lips. For the lower note, however, this phase relation requires that the flow into the bore has a large component that begins before the lips separate. This is explained by the sweeping action of the lip: for the lower note, the lips move forwards into the mouthpiece before they separate, generating a flow whose magnitude is of the same order as the acoustic flow into the bore.

This preliminary report presents only some of the data in the early part of this study. The conference presentation will have further data on more players, a greater range of notes, and a larger set of techniques.

Acknowledgments

We thank the Australian Research Council for support and our volunteer subjects.

7. REFERENCES

- [1] D. C. Copley and W. J. Strong, “A stroboscopic study of lip vibrations in a trombone,” *The Journal of the Acoustical Society of America*, vol. 99, no. 2, pp. 1219–1226, 1996.
- [2] S. Yoshikawa and Y. Muto, “Lip-wave generation in horn players and the estimation of lip-tissue elasticity,” *Acta Acustica united with Acustica*, vol. 89, no. 1, pp. 145–162, 2003.
- [3] A. Z. Tarnopolsky, N. H. Fletcher, L. C. Hollenberg, B. D. Lange, J. Smith, and J. Wolfe, “Vocal tract resonances and the sound of the Australian didgeridu (yidaki) i. experiment,” *The Journal of the Acoustical Society of America*, vol. 119, pp. 1194–1204, 2006.
- [4] M. J. Newton, M. Campbell, and J. Gilbert, “Mechanical response measurements of real and artificial brass players lips,” *The Journal of the Acoustical Society of America*, vol. 123, no. 1, pp. EL14–EL20, 2007.
- [5] S. Elliott and J. Bowsher, “Regeneration in brass wind instruments,” *Journal of Sound and Vibration*, vol. 83, no. 2, pp. 181–217, 1982.
- [6] N. Fletcher, “Autonomous vibration of simple pressure-controlled valves in gas flows,” *The Journal of the Acoustical Society of America*, vol. 93, pp. 2172–2180, April 1993.
- [7] V. Fréour and G. P. Scavone, “Vocal-tract influence in trombone performance,” *Proc. Int. Symp. Music Acoustics, Sydney & Katoomba*, 2010.
- [8] J.-M. Chen, J. Smith, and J. Wolfe, “Do trumpet players tune resonances of the vocal tract?” *The Journal of the Acoustical Society of America*, vol. 131, no. 1, pp. 722–727, 2012.

- [9] J. Wolfe and J. Smith, "Acoustical coupling between lip valves and vocal folds," in *Acoustics Australia*, vol. 36, 2008, pp. 23–27.
- [10] V. Fréour and G. P. Scavone, "Development of an electrolabograph embedded in a trombone mouthpiece for the study of lip oscillation mechanisms in brass instrument performance," *Canadian Acoustics*, vol. 39, no. 3, pp. 130–131, Oct. 2011.
- [11] J. Wolfe, J. Smith, J. Tann, and N. H. Fletcher, "Acoustic impedance spectra of classical and modern flutes," *Journal of Sound and Vibration*, vol. 243, no. 1, pp. 127–144, 2001.
- [12] P. Dickens, J. Smith, and J. Wolfe, "Improved precision in acoustic impedance measurements by using calibration loads without resonances," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 3015–3015, 2007.
- [13] R. D. Ayers and M. S. Lodin, "Alternative mouthpiece design for viewing the lip reed in motion," *The Journal of the Acoustical Society of America*, vol. 107, p. 2843, 2000.
- [14] J. W. Coltman, "Acoustical losses in wet instrument bores (I)," *The Journal of the Acoustical Society of America*, vol. 114, p. 1221, 2003.
- [15] J. Wolfe, J. M. Chen, and J. Smith, "The acoustics of wind instruments-and of the musicians who play them," in *Proceedings of the 20th International Congress on Acoustics, ICA-2010*, August 2010.

TIMPANI-HORN INTERACTIONS AT THE PLAYER'S LIPS

Jer-Ming Chen

School of Physics, The University of New South Wales, Sydney
jerming@unsw.edu.au

John Smith

john.smith@unsw.edu.au

Joe Wolfe

j.wolfe@unsw.edu.au

ABSTRACT

This study investigates the observation by some horn players that a timpani sounding nearby can interfere with their playing. By determining the horn's transfer function and measuring the pressure response in the bell and mouthpiece during moderate to loud timpani strokes, the horn is found to behave as an acoustic impedance-matching device capable of transmitting an overall impulse gain response of at least ~ 16 dB from the bell to the mouthpiece, while some non-linear propagation in the bore is also observed. Further resonance interactions between the bore of the horn and the timpani stroke show gain responses of up to ~ 26 dB, which depend on the timpani's tuning. Lastly, pressure measurements in the mouthpiece made during horn playing show that timpani strokes played near the bell can affect the amplitude, periodicity and frequency of the pressure signal generated at the horn player's lips, and may be large enough to perturb the player's musical performance.

1. INTRODUCTION

The function of the bell of a horn is well known: it is an impedance matcher. When the horn is played, the bell efficiently radiates high frequencies outwards, and so contributes to the instrument's characteristic timbre. In the inwards direction, however, the bell is expected to increase the pressure amplitude of waves travelling into the horn from outside the bell. This property may explain the observation by orchestral horn players and teachers [1, 2, 3] that, when the horn and the timpani play in close proximity, and especially when the bell of the horn faces the timpani, there is a tendency for timpani strokes to interfere disruptively with horn playing.

The celebrated horn player, composer, conductor and jazz musician Gunther Schuller (b. 1925) writes: "The timpani's spreading wave-lengths back up through the horn, violently jarring the player's lips. Under these conditions split notes abound and what notes can be played develop a strong rasp. A half minute of this and the horn player will retain no sensitivity in his lips." [1]

The scope and explanation of this phenomenon remains an active source of discussion amongst horn players and teachers (e.g. online horn forums [4]) but, to the authors'

knowledge, there have been no acoustical studies on this matter so far.

Accordingly, this paper reports preliminary measurements of the pressure measured in the mouthpiece of the horn in response to external sounds (either timpani strokes or sustained broadband excitation) applied outside the bell of the horn.

2. MATERIALS AND METHODS

Three experimental setups were used to make the measurements in this study.

2.1 Measurement of Horn Transfer Function

In a room treated to reduce external noise and reverberation, a Yamaha YHR-664 double horn is suspended over a loudspeaker, such that its bell faces the loudspeaker coaxially with a separation of one bell radius.

The transfer function of the horn is usually measured from mouthpiece to bell. Here, it is measured from bell to mouthpiece using a source at the bell. Two $\frac{1}{4}$ -inch pressure-field microphones (Brüel & Kjær 4944A) are used: one is positioned at the plane of the bell, near the centre, while the second is fitted into a specially modified horn mouthpiece which enables the microphone to measure the pressure at the mouthpiece via a 1 mm vent drilled into the cup. The mouthpiece, with the microphone attached, is sealed and isolated from the external radiation field using a specially fitted nylon cap.

A broadband probe signal (25-1000 Hz, at 2.7 Hz intervals) is produced by the loudspeaker. The pressure spectrum of this broadband probe signal was 'flattened' with respect to the microphone situated at the plane of the bell, using the software ACUZ [5]. The resulting FFT of the two microphone signals are then time-averaged and divided to yield the horn transfer function.

Measurements were made for both F and B \flat horns for the fingerings 000, 0X0, X00 and XX0, where X means depressed for index, middle and ring fingers respectively.

2.2 Impulse Measurement Using Timpani Strokes

In the same room, the same horn (with the same microphones located at the bell and mouthpiece) is now suspended over a single timpani (26" drum, Evans USA, nominal sounding range F2-E3), such that its bell faces the drum skin coaxially with a separation of one bell radius. The horn transfer function from bell to mouthpiece is not simply related to that from mouthpiece to bell.

Nevertheless, the peaks in the transfer function measured here correspond approximately to the sounding frequencies of the horn.

For the 000 fingering on the F horn:

- 2nd resonance, 86.1 Hz (F2-23 cents)
- 3rd resonance, 131.9 Hz (C3+14 cents)
- minimum between the 2nd and 3rd resonance, 107.7 Hz (A2-37 cents)

For the 000 fingering on the B \flat horn:

- 2nd resonance, 118.4 Hz (A#2+27 cents)
- minimum between the 2nd and 3rd resonance, 145.3 Hz (D3-18 cents)
- the 3rd resonance occurred above the range of the timpani and so was not measured

To investigate the effect of a timpani tuned close to horn resonances, the timpani was tuned over a range of pitches deviating up to ± 100 cents from the above frequencies, and struck at dynamic levels ranging from *mp* to *mf*. The resulting pressure signals at the bell and in the mouthpiece were then recorded by the two microphones, and analysed.

For reproducibility, these measurements were made without a hand in the bell.

2.3 Timpani Strokes During Horn Playing

The horn remains positioned as before with the cap removed from the mouthpiece so the instrument can be played, but without the hand in the bell. Greater sound pressure levels are now expected in the mouthpiece, and consequently the mouthpiece microphone is replaced by a piezoresistive pressure transducer (Endevco 8507c-2). While the player plays sustained, steady notes at the horn resonances identified earlier (sounding F2, C3 and B \flat 2, but written C3, G3 and F3 respectively for the horn in F) at *p* and *mf* dynamic levels, the timpani (tuned to these notes, and also tuned to ± 70 cents) is struck at dynamic levels ranging from *mf* to *ff*, while the pressure in the mouthpiece is recorded and analysed.

Lastly, an ‘ecological’ measurement is made while the horn player plays sitting in the normal concert position, hand in the bell in the usual position, with the bell pointing at the timpani, struck 1 meter away.

3. RESULTS AND DISCUSSION

3.1 Transfer Function Measurements

The acoustic transfer functions measured for the 000 fingering on the F and B \flat horn are shown in Figure 1.

In both plots, the overall transmission gain from bell to mouthpiece increases steadily with frequency, and local maxima indicate resonances of the horn. For the 000 fingering on the F horn, the second resonance has a gain of 20 dB, rising steadily to 27 dB by the sixth resonance. For the B \flat horn with 000 fingering, the second resonance has a gain of 23 dB, and rises to 28 dB at the sixth resonance. (The first resonance in the horn is below the pedal note and is not played.) A comparable gain profile is observed for the transfer functions of other fingerings measured.

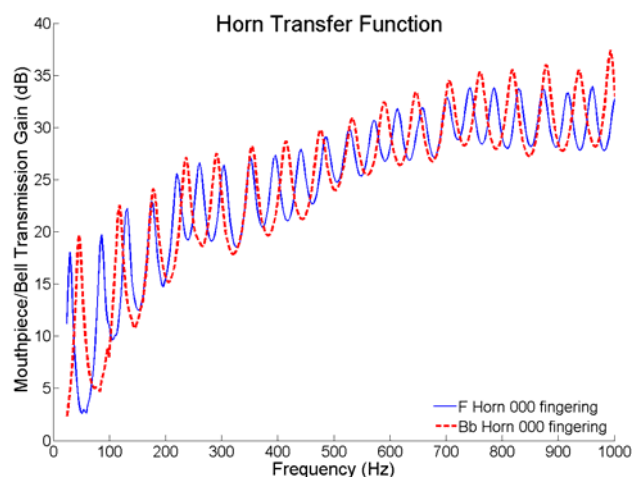


Figure 1. Bell-to-mouthpiece horn transfer function measured for the 000 fingering on the F and B \flat horn.

3.2 Impulse Measurements Using Timpani Strokes

Figure 2 shows a typical pressure pulse waveform of a timpani stroke, measured in the bell (top) and in the mouthpiece (bottom), both shown on the same scale. In this example, the timpani is tuned nominally to A2+25 cents and the 000 fingering on the B \flat horn is used.

Here, the pressure pulse arriving at the mouthpiece arrives 8 ms after the pulse enters the bell, consistent with the ~ 2.75 m length of the B \flat horn. The largest trough of the pressure signal in the mouthpiece exceeds that in the bell by 17 dB. Similarly, the first pressure peak arriving at the mouthpiece exceeds that in the bell by 16 dB. The initial trough and peak at the bell and mouthpiece are comparable in shape but the subsequent pressure signal received in the mouthpiece is noticeably different from that measured at the bell, because of the standing waves produced in the bore.

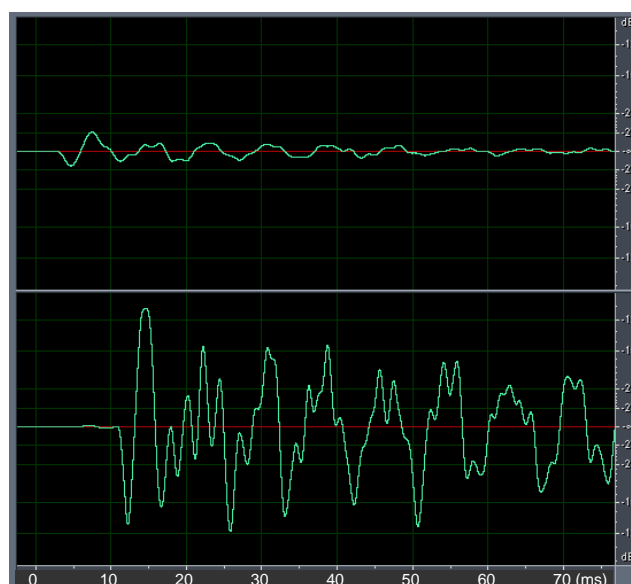


Figure 2. Typical waveforms of the pressure pulse of a timpani stroke, nominally tuned A2+25 cents, measured in the bell (top) and in the mouthpiece (bottom) for the B \flat horn 000 fingering, shown on the same scale and measured using equal microphone amplifier gains.

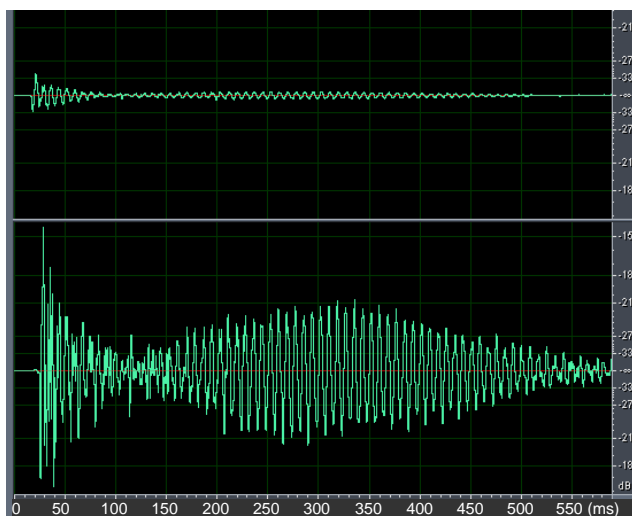
Impulse gain values measured for F Horn 000 fingering (dB)			
2 nd resonance (R2)		3 rd resonance (R3)	Trough
R2-32 cents	14.8±2.7	R3-134 cents	15.6±1.8
R2-22 cents	15.6±2.0	R3-104 cents	16.9±1.1
R2+02 cents	15.4±2.5	R3-74 cents	16.5±1.3
R2+13 cents	16.0±2.7	R3-34 cents	16.8±1.3
R2+38 cents	16.3±1.9	R3-04 cents	15.9±1.2
R2+58 cents	16.1±1.7	R3+21 cents	16.0±0.7
R2+88 cents	16.2±3.1	R3+46 cents	16.2±1.0
(Average	15.8±0.6)	R3+76 cents	15.5±1.5
		R3+106 cents	16.7±1.3
		(Average	16.2±0.5)

Impulse gain values measured for Bb Horn 000 fingering (dB)			
2 nd resonance (R2)		Trough	
R2-102 cents	17.2±1.0	-98 cents	17.7±0.8
R2-52 cents	17.3±1.0	+02 cents	17.2±0.8
R2-02 cents	17.4±0.7	+102 cents	17.7±0.6
R2+03 cents	17.3±1.1	(Average	17.5±0.3)
R2+48 cents	18.0±0.8		
R2+98 cents	17.7±0.7		
(Average	17.5±0.3)		

Table 1. The gain values (average \pm standard deviation) of the pressure impulse (initial trough & peak) extracted for each timpani stroke, measured at a range of timpani pitches tuned near a corresponding peak (resonance) and trough of the measured horn transfer function.

The averaged impulse gain values of the initial trough and peak in the pressure signal, extracted for each timpani stroke and played at a range of timpani tunings, are collated for the 000 fingering on the F and B \flat horn and tabulated in Table 1.

For each fingering, the impulse gain values obtained are fairly consistent in magnitude and are only weakly dependent on the frequency difference between the timpani note and the horn resonance: a consistent gain of ~ 16 dB for the F horn 000 fingering, while the B \flat horn 000 fingering (which has a shorter pipe and hence less attenuation) shows a higher gain of ~ 18 dB. On a time scale too short for reflections, resonances are irrelevant and the horn acts as an acoustic impedance-matching transformer for the external pressure impulse from the struck timpani. For pressure amplitudes up to the linear limit, which at this separation corresponds to timpani notes up to *mf*, this gain is approximately independent of the magnitude of the pressure pulse.



Pressure pulses exceeding ~ 1 kPa (~ 150 dB) are sometimes measured in the mouthpiece if the external impulse signal at the bell is of the order of 100 Pa (~ 130 dB) or greater. At these larger amplitudes (strokes $> mf$), the pressure pulse is observed to arrive at the mouthpiece with somewhat larger gain and an altered shape, *e.g.* for the F horn 000 fingering, peaks arrive 5% (0.6 ms) sooner than the trough, on average. This is consistent with the effects of nonlinear propagation in the bore, which are expected to be noticeable at this sound level because of the relatively long distance travelled in the narrow bore [6].

For the initial impulse, there is no dependence on the relative tuning of timpani and horn. In the later response, once the energy transmitted from the timpani sets up standing waves in the bore of the horn, we should expect to see the effects of such tuning. Figure 3 shows two contrasting timpani strokes both measured using the B \flat horn 000 fingering: one is tuned to the second horn resonance, and the other is tuned to the transfer function minimum between the second and third resonances.

In both strokes, the large-amplitude aperiodic transient of the timpani signal at the bell lasts about 0.1 s and is followed by a quasi-periodic slow decay. During this quasi-periodic mode, we can observe the effects of resonance in the case where the timpani is tuned to the horn resonance. The envelope of the signal measured at the bell, which is largely due to the signal produced by the timpani, decays almost monotonically. In the mouthpiece signal, however, the amplitude of the quasi-periodic signal rises smoothly from about 0.13 to 0.3 s, as more energy from the timpani is gradually stored in the standing wave in the bore. This peaks at a gain of about 26 dB, and remains near that level until about 0.5 s. In contrast, the stroke tuned away from the horn resonance receives no ‘help’ from the horn resonance and thus has no boost observed in the mouthpiece signal; its decay envelope is comparable with that measured at the bell.

In a large majority of orchestral scores, the timpani play the tonic (or less commonly the fifth) of the chord, which is also played by the horns. In very many cases, therefore, one or more of the horns would be using a fingering in

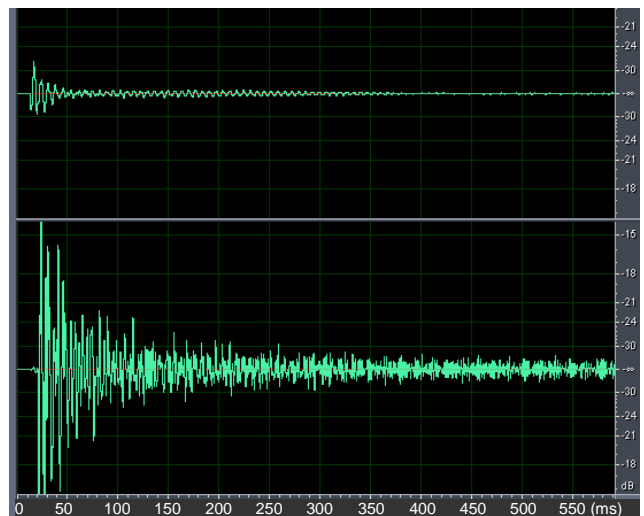


Figure 3. Two contrasting timpani strokes, measured using the B \flat horn 000 fingering, showing the microphone signal at the bell (top) and in the mouthpiece (bottom), on the same scale. *Left:* timpani tuned to A#2+25 cents to coincide with the 2nd horn resonance. *Right:* timpani tuned to D3-20 cents to coincide with the transfer function minimum between the 2nd and 3rd resonance.

which one of the resonances of the horn would be tuned close to the frequency of the timpani note.

This poses two difficulties for the horn player. First, s/he will receive a large transient, produced by the timpani, and arriving at the lips during or at best soon after the transient of the horn note. Even in the absence of the timpani, this starting transient can already be hard for the player to play cleanly, because the first several or more vibrations of the lips must be produced before the reflection from the bell has produced standing waves that stabilise the vibrations of the lips. Second, the timpani sets up a slowly increasing periodic wave that adds to the standing wave produced by the player and potentially interferes with the motion of the lips.

3.3 Timpani Strokes During Horn Playing

In many cases, the pressure impulse signal from the timpani stroke can be easily observed in the mouthpiece during horn playing at both the *p* and *mf* dynamic levels measured (~152 dB and ~158 dB respectively, measured in the mouthpiece): the arrival of the timpani stroke is indicated by a region of large-amplitude aperiodic transients (up to ~6 dB larger than the quiescent lip pressure signal) lasting several lip pressure cycles (~50 ms); alternatively, the pressure pulse might arrive at exactly the right moment to destructively interfere with and to reduce the lip pressure signal for several cycles (up to ~6 dB less than the quiescent lip pressure signal). Further, if the timpani plays at a frequency close to a horn resonance, as would often be the case, resonance-driven interactions in the horn's bore (similar to that reported earlier in §3.2) sometimes persist up to 0.5 s.

Figure 4 shows an example of a measurement made of a horn player playing at the *p* dynamic level (149 dB in the mouthpiece here), sitting in the normal concert position with his hand in the bell and the bell pointing at the timpani, which is struck 1 meter away.

Here, the pressure signal (playing at A#2+20 cents) generated by the player's lips (bottom) is quasi-periodic (quiescent) up until the arrival of the timpani stroke (seen 8 ms beforehand in the top signal, measured at the bell) where it becomes disrupted: strong irregular transients are observed in the first 50 ms (reaching 4.5 dB above the quiescent lip pressure signal here), while irregularities of amplitude and structure persist in the lip signal up to 0.5 s before the lips resume quiescent vibration.

Other perturbations are also observed in this close proximity of instruments: the player's sounding pitch can sometimes become unstable immediately following a stroke, deviating by several cents, if the timpani is tuned close to the playing pitch. However, in some instances where the timpani was tuned ~80 cents flatter than the horn sounding pitch, the lip signal was pulled similarly ~80 cents flatter for up to 0.5 s following the stroke.

A related disruptive effect is sometimes also reported when horn players are seated closely together and playing a high passage at a loud dynamic level in unison (a fairly common occurrence in an orchestral climax). Under these conditions, it is sometimes reported to be difficult for the players to sustain the notes [1]. The relative phase between the waves produced by a player and his/her neigh-

bours has no predictable relationship, therefore potentially disruptive interference might also be possible here.

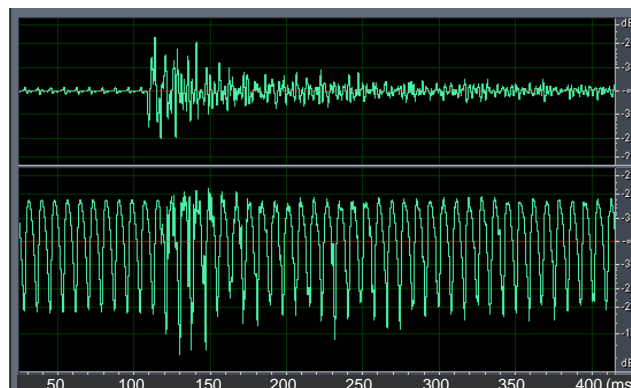


Figure 4. A typical waveform of the pressure pulse of a timpani stroke during horn playing, both sounding A#2+20 cents, measured in the bell (top) and in the mouthpiece (bottom) on the B \flat horn 000 fingering. The horn is played softly in the normal concert position (hand in the bell), with the bell pointing at the timpani, struck 1 m away. The microphones have different gains.

4. CONCLUSIONS

Transfer function measurements of the horn at various fingerings show gains of at least ~20 dB between the periodic pressure signal input at the bell and the signal which is transmitted to the mouthpiece.

Measurements of timpani strokes made near the horn reveal an overall impulse gain response of at least ~16 dB, because the horn is behaving as an acoustic impedance-matching receiver in this case. However, when the timpani is also tuned near a resonance in the horn, as would normally be the case in orchestral performance, a dramatic gain of ~26 dB can be observed once the timpani signal excites standing waves in the bore. Further, evidence of nonlinear wave propagation in the horn has been observed, allowing even greater transmission to the mouthpiece if the external impulse signal at the bell is of the order of 100 Pa or greater.

Pressure measurements in the mouthpiece made during horn playing show that both the large-amplitude aperiodic transient and the quasi-periodic decay of the timpani stroke can interact with the pressure signal from the horn player's lips to affect its amplitude (both constructively and destructively), periodicity and frequency. This interaction may be large enough to interfere with the player's control of his/her lips during musical performance.

Acknowledgments

We thank the Australian Research Council for support, our volunteer horn player and Emery Schubert for loan of his instruments.

5. REFERENCES

- [1] G. Schuller, *Horn Technique*, Oxford University Press, 1962, pp. 83-84.
- [2] Douglas Hill, *Collected Thoughts: On Teaching and Learning, Creativity, and Horn Performance*, Warner Bros Publications, 2001, p. 71.

- [3] Martin Buckle, “The Horn and the Alphorn”, Website <http://www.martinatnewton.com/page4.htm> (last altered/updated on 23 March 2013).
- [4] Yahoo Groups Horn Forum: <http://launch.groups.yahoo.com/group/horn/message/58568> (last altered/updated 12 Feb 2012).
- [5] J.R. Smith, N. Henrich, and J. Wolfe, “The acoustic impedance of the Bøhm flute: standard and some non-standard fingerings,” in *Proc. Inst. Acoustics*, 1997, vol. 19, pp. 315–320.
- [6] A. Hirschberg, J. Gilbert, R. Msallam, and A.P.J. Wijnands, “Shock waves in trombones,” *J. Acoust. Soc. Am.*, 1997, vol. 99(3), pp. 1754–1758.

The use of the input impedance for characterising historical serpents

Pauline Eveno and Sandie Le Conte

Musée de la Musique
221 avenue Jean-Jaurès
75019 Paris, France
pauline.eveno@gmail.com

ABSTRACT

This article describes how the input impedance is used as a descriptor to classify and characterise the corpus of serpents from the Musée de la Musique. The study focuses on the products of a family of craftsmen (Baudouin father and son) with the objective to find a trace of the knowhow transmission. Input impedance measurements highlight small differences between serpents made by Baudouin and serpents from other craftsmen but do not allow for discrimination between the instruments made by the father or the son.

1. INTRODUCTION

The Musée de la musique keeps in its collection more than five thousand musical instruments. Among them, only five percent are in a functionable state. Indeed, a musical instrument is of historic, aesthetic, cultural, technic and musical import and curators have to evaluate their cultural values in order to favour some of them in the context of an exhibition. For wind instruments, the general conservation rules are not to play them, to avoid a quick and huge hygrometric variation. When the musical value is interesting, and to allow the public to have a chance of listening a proposition of historical sound, the Musée orders a facsimile. The facsimile is a copy of an original musical instrument, using historical techniques and the same materials. The collection of serpents in the Musée de la musique includes thirty five instruments and one facsimile.

The serpent is described in the Dictionnaire Technologique written by Francoeur in 1831 [1] as a wind instrument of low register. It is made of several pieces of wood joined together thanks to proteinic glue and maintained with a thin layer of leather. It is played with a mouthpiece, and can reach 2.4 meters long. Only six finger holes are made and, according to the available methods, its fundamental is D1 and many fingerings exist to produce the same note. The pitch of the notes had still not been normalized at that time, but a spread value was the A 392 Hz, which will be used in this article.

Very little information about this corpus is available. For example, without any historical documentation, it is im-

possible to date the instruments. Nevertheless, some of them are signed by the maker Baudouin. Paris archives [2] show that George Antoine Baudouin claimed to be a serpent maker early in the 19th century. He died in 1816 and his son, Antoine Gabriel Baudouin, took over the family business. One of the curator's goals is to distinguish the instruments made by the father from those made by the son.

The input impedance helps characterising the acoustical response of a musical instrument at different frequencies. It depends on the geometry of the bore and small bore variations lead to different impedances [3, p. 65-67]. Impedance measurement can be used by craftsmen to evaluate the reproducibility of wind instruments making process [4-6], or to characterise the influence of a process step, as the bending for example [7]. It is more difficult to use the input impedance with ancient instruments, mostly owing to the presence of leaks. This article proposes to describe the advantages and the limitations of the input impedance measurement of a historical musical instruments' corpus.

2. PROTOCOL

2.1 Impedance measurement

Many techniques are used to measure the input impedance [8, 9]. The difference between these methods consists in the measurement of the acoustical flow. Using one or two couples of microphones is the most popular technique [10, 11]. In this study, the experimental set up developed jointly by the LAUM¹ and the CTTM² has been used [12]. In this system, one microphone is positioned in a cavity behind the source which allows an estimation of the flow. The source is a piezo electric buzzer. The instrument is connected in front of the buzzer via a small cavity in which a second microphone is placed to measure the acoustical pressure. The transfer function between the two microphones, taking into account their sensitivities, allows the estimation of the input impedance.

The absolute uncertainty of the resonance frequencies measurement with this sensor is estimated to be 0.3%, which is 5 cents [13]. However, when comparing different resonances of the same instrument, the relative error is estimated to be ± 3 cents, as systematic errors, such as those due to temperature or sensor geometry, are partially compensated.

Copyright: ©2013 Pauline Eveno et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ Laboratoire d'Acoustique de l'Université du Maine, Avenue Olivier Messiaen, 72085 LE MANS Cedex 9

² Centre de Transfert de Technologie du Mans, 20 rue Thalès de Milet, 72000 Le Mans



Figure 1. Experimental set up used to measure the input impedance of historical Serpents.

The input impedance measurement is sensitive to the air temperature and to the airtightness between the device and the instrument. Silicone adaptors are made for each instrument (Figure 1). These adaptors allow for perfect contact between the device and the serpent and respect the conservation rules (silicone is non corrosive to both metal and wood). Moreover, the measurements are carried out in a climate-controlled room.

2.2 Measured serpents

Five of the nine instruments studied were made by Baudouin from 1812 to 1827. It is actually the most represented maker in the collection. Some of those serpents were most likely made by the father, and other ones by the son. The curator's hypothesis is to attribute to the father the serpents to which there is only one mark, and to the son, the instruments which have two marks (see Figure 2): first, because serpents with one mark are rarer than the one with two marks and second, because serpents with two marks seem to have been provided with keys from the outset [2]. Another serpent from the same period but made by another craftsman, Degalle, is also presented in this article. Finally, the three last studied serpents are contemporary. One is a facsimile by Berger (leather) and Wetter (wood) of a serpent made by Coeffet around 1830. The other two are made by Ribo and one of them is a copy of a serpent kept in the MIM (Musée des Instruments de Musiques) in Bruxelles. All these serpents are showed in Table 1.

Serpents kept in the Musée de la musique are identified with their inventory number. The input impedance is measured on the entire instrument, i.e with the crook and the mouthpiece (which are kept with the instrument, but it is not always possible to be sure that they come from the same instrument), positioned according to traces of use. Another measurement is also taken on the resonators only.

3. RESULTS

3.1 Precision of the measurement

To have reliable and repeatable measurements, it is necessary to firmly affix the instrument flat on the impedance



(a) 2 marks attributed to Baudouin son









(b) 1 mark attributed to Baudouin father

Figure 2. Baudouin's signature on his serpents.

sensor in order to avoid any leak and to prevent it from moving. The silicone adaptors guarantee a good airtightness but do not stand for a rigid mounting. Moreover, the presence of leaks in the instrument, the difficulty to open or close some keys, etc. make the measurements difficult to repeat with a same result. So, in the case of serpent measurements, it is not possible to reach the level of precision indicated in the previous section. Measured amplitudes of a same instrument can differ by 10% whereas the frequency varies by 2%. Fortunately, frequency gives more information on the acoustics of the instruments than the amplitude, especially regarding the tuning. It may be questionable to compare several instruments with such a low reproducibility rate. Nevertheless, even with instruments that are supposed to be identical, differences can be highlighted by the input impedance. For example, Mamou-Mani et al. [5] compared five Howarth S10 student model oboes and found that the peak amplitudes can vary by as much as 30%, while the peak frequencies differ by no more than 2%. Therefore, it is consistent trying to distinguish different instrument craftsmanship with the input impedance measurement, especially when it is possible to have an estimation of resonance frequencies with a 2% accuracy.

3.2 Serpent impedance and the EFP

A typical serpent input impedance, as in Figure 3, shows several peaks which correspond to the eigenmodes of the

Inventory number	Type	Photo	Inventory number	Type	Photo
E575	Baudoin 1 mark 6 holes		E980.2.675	Baudoin 2 marks 8 holes 2 keys	
DAD32046	Baudoin 1 mark 7 holes 1 key (missing)		E1292	Baudoin 2 marks 6 holes	
DECL19193	Baudoin 2 marks 6 holes		Facsimile	Berger and Wetter 6 holes	

continued on next page




<i>continued from previous page</i>					
Inventory number	Type	Photo	Inventory number	Type	Photo
E2011.14.1	Degalle 6 holes		Ribo-6holes	Ribo 6 holes	
Ribo-10holes	Ribo (copy MIM) 10 holes 4 keys	 (photo of the original)			

Table 1. Photos and features of the nine studied serpents.

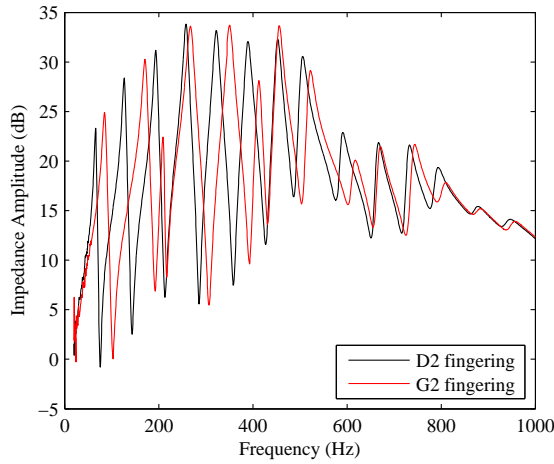


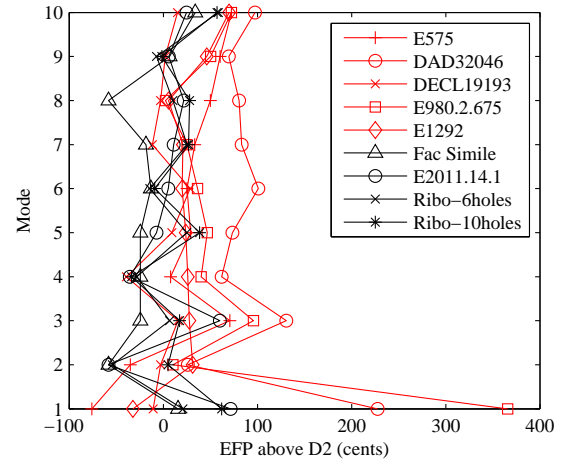
Figure 3. Input impedance of the E2204 facsimile for two fingerings: D2 fingering with all the holes closed (in black) and G2 fingering with the last three holes open (in grey).

instrument. The notes playable by the musician are supposed to have a frequency close to the resonance frequencies of the impedance. Figure 3 shows that, when all the holes are closed, the resonances are in a quite harmonic relationship, but when the holes are open this is not the case anymore. It is possible to observe the instrument harmonicity using the input impedance by plotting the equivalent fundamental pitch (EFP) [14]. It is defined as follows:

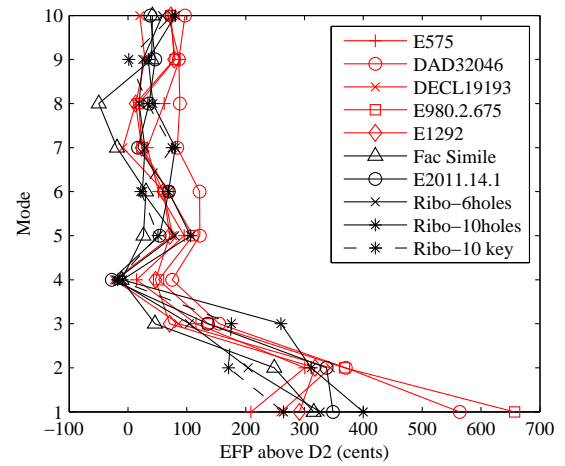
$$\text{EFP}(n) = \frac{1200}{\log 2} \log \left(\frac{f(n)}{nf_0} \right), \quad (1)$$

where $f(n)$ is the frequency of the n^{th} peak and f_0 is the frequency of a reference tone. The EFP represents the interval (in cents) between the frequency of each mode and the frequency of the corresponding mode for an ideal conical instrument of fundamental frequency f_0 . EFP(n) is the pitch of the harmonic series whose n^{th} member has frequency $f(n)$. Most of the methods give D2 as the fundamental frequency of the serpent, which corresponds to a frequency of 65.4 Hz, since the diapason was A 392 Hz at the time.

Figure 4 shows the EFP of the nine serpents for two fingerings: D2 fingering, where all the holes are closed, and E2 fingering where the last hole is open and the other are closed (for the Ribo serpent with 10 holes there are two options: opening the last two holes to be in the same configuration as the other serpents, or opening just the last hole with the key in order to have a longer air column). Regarding the EFP of the D2 fingering, four serpents have, except for the first mode, an identical behaviour: serpents E575, E980.2.675, DAD32046 and E2011.14.1. Indeed, the four have an EFP that increases by around 100 cents from the second mode to the third, then it decreases by approximately 50 cents from the third mode to the fourth and finally, for the superior modes it keeps more or less constant. Since there are a lot of leaks in these old serpents, it is often difficult to measure the first impedance peak, this is why there are many differences. DAD32046 and E980.2.675



(a) All holes closed (D2 fingering)



(b) The last hole open (E2 fingering)

Figure 4. EFP of the six serpents for two fingerings: D2 fingering where all the holes are closed and E2 fingering with the last hole open. Baudoin's serpents are given in red and the others in black.

have a too high first resonance. That should not be taken into account, this is in all likelihood due to leaks. It can also be observed that the facsimile has lower resonances than the others and serpents E1292 and DECL19193 are more harmonic than the others. Moreover, the two serpents made by Ribo have close resonances. This is because Ribo used the same shape for his resonators, which was copied with a machine shown in Figure 5. This first figure therefore shows that it is possible to distinguish the work of one factor: Ribo. With modern tools it is indeed easier to reproduce identical instruments. Moreover, it is possible to classify Baudoin's serpents into two groups: E575, DAD32046 and E980.2.675 in one group and DECL19193 and E1292 in another. That might prove the work of two different persons, the father and the son, but in that case, it is not the number of marks that can help classify them.

The second plot in Figure 4 can show if the last hole is set up differently in some serpents. It is possible to measure the position directly on the instrument, which can seem



Figure 5. Machine used by Ribo to copy the shape of his resonators.

easier than doing an impedance measurement, but the advantage of the impedance is that it directly shows the impact of the holes in the air column of the instrument. All the instruments in this figure, without considering the first mode, and with exception to the two made by Ribo, have a quite similar behaviour. This to say, a huge decrease of the EFP of about 200 cents between the second and the third mode and then an EFP more or less constant. It is therefore only possible to distinguish the work of Ribo from the one of other artisans. Regarding the relation between the two first resonances it is however possible to highlight the work made by Baudoin. Indeed, for E575, DECL19193 and E1292, the second mode is higher than the first one, whereas it is the contrary for all the other serpents. When plotting the EFP for other fingerings, the results have similar behaviours.

In Figure 4, the EFP is plotted for the entire serpents (with mouthpiece and crook), which is interesting to know which notes can be played with the instrument. Nevertheless, it is not possible to be sure if the mouthpiece is original, and the crook can have different positions, which significantly modifies the input impedance. All these elements add a lot of uncertainty in the comparison of the instruments. This is why it is interesting to plot the EFP³ of the resonator alone, in order to only characterise the craftsmanship of the wood.

Figure 6 shows the EFP of the serpents resonators. Unfortunately, for the serpent DECL19193 it was impossible to remove the crook which was stuck inside the resonator. Moreover, measurements of Ribo's serpents were taken rapidly at an exhibition and only the entire serpents, which were more interesting for the craftsman, were measured. This is why only the EFP of six serpents are plotted in this figure. In Figure 6, higher modes can be separated into three groups: 1) DAD32046, E1292 and E.980.2.675, 2) Facsimile and E2011.14.1 and 3) E575. These results show the importance of the embouchure and the crook on the acoustics of the instrument. Indeed, the serpent DAD32046 whose resonator acoustics is very close to the E1292 at high frequency behaves in a completely different manner when it is provided with a crook and a mouthpiece. More-

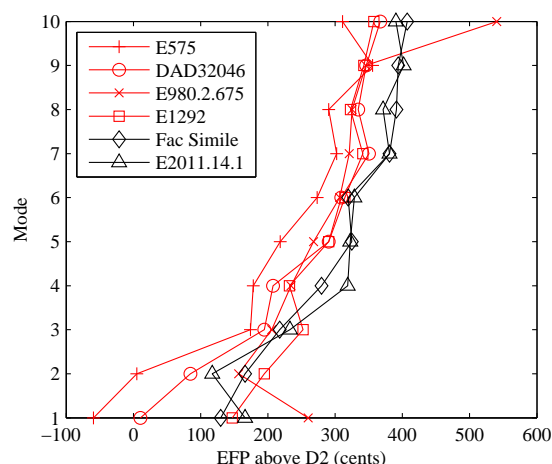


Figure 6. EFP of the resonators of six serpents with all their holes closed. Baudoin's serpents are given in red and the others in black.

over, in Figure 6 the serpent DAD32046 does not have its first impedance peak much higher than the others, as in Figure 4. This seems to indicate that the mouthpiece found with the instrument is not suitable or that there is a problem with its crook. At low frequency, behaviour differs and serpents can be split into two groups: 1) E575, DAD32046, E1292 and Facsimile which have a second mode higher than the first, and 2) E980.2.675 and E2011.14.1 for which it is the contrary.

4. CONCLUSIONS

The input impedance is useful to characterise musical instruments that cannot be played for conservation reasons. It gives information about the notes that can be played by the instrument and it can help reconstruct the shape of the bore. This is why it was used to try to understand the work of the craftsmen Baudoin. Measurements are sometimes difficult to conduct since instruments are old, have a lot of leakages and are difficult to firmly mount on the impedance probe. Nevertheless, it is possible to obtain repeatable measurements, especially regarding the frequencies. By looking at the entire serpents, it is possible to distinguish the work of a contemporary craftsman, as he uses a machine that can reproduce similar shapes. Some distinctive features of Baudoin's work have been highlighted but they do not differ much from the work of other craftsmen. Moreover, those features are not the same for entire serpents or for only resonators.

This shows the limits of the input impedance for answering this kind of question. No explanation can be given on the number of marks since some 1 mark serpents are acoustically closer to serpents with 2 marks than others with 1 mark. The input impedance gives some clues but a part of mystery remains... Radiographies of the serpent resonators will be done, but it should only confirm the results obtained with the input impedance.

³ We keep the EFP diagram to simplify, even if there is no musical interpretation that can be done on it.

Acknowledgments

The authors would like to acknowledge T. Maniguet who allowed these measurements as the curator of the corpus, S. Vaiedelich who helped for the protocol setup, P. Ribo who provided his serpents and the French Culture ministry which supports this study. Thank you to R. Jilwah and M. Ghamsari for proofreading.

5. REFERENCES

- [1] L.-B. Francoeur, *Dictionnaire technologique ou nouveau dictionnaire universel des arts et métiers, et de l'économie industrielle et commerciale; par une société de savants et d'artiste*. Thomine et Fortic, Paris, 1831, ch. Tome 19, pp. 300–303.
- [2] T. Maniguet, “L’usage du serpent et de ses formes dérivées dans la première moitié du xixe siècle,” *Musique, Images, Instruments*, vol. 14, in press.
- [3] P. Eveno, “L’impédance d’entrée pour l’aide à la facture des instruments de musique à vent : mesures, modèles et lien avec les fréquences de jeu (the input impedance for the support of the musical instruments making: measurements, models and link with the playing frequencies),” Ph.D. dissertation, Univ. Pierre et Marie Curie (Paris VI), 2012. [Online]. Available: <http://goo.gl/RvujY>
- [4] V. Gibiat and J. Selmer, “Impedance comparison of high-quality alto and soprano saxophones,” *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2292–2292, 2002.
- [5] A. Mamou-Mani, D. B. Sharp, T. Meurisse, and W. Ring, “Investigating the consistency of woodwind instrument manufacturing by comparing five nominally identical oboes,” *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp. 728–736, 2012.
- [6] A. Mamou-Mani and D. B. Sharp, “Evaluating the suitability of acoustical measurement techniques and psychophysical testing for studying the consistency of musical wind instrument manufacturing,” *Applied Acoustics*, vol. 71, no. 7, pp. 668–674, 2010.
- [7] J.-P. Dalmont, M. Curtit, and F. Yahaya, “Bore reconstruction from input impedance,” *Proceedings of Forum Acusticum, Aalborg, Denmark*, 2011.
- [8] J.-P. Dalmont, “Acoustic impedance measurements, part i : A review,” *Journal of Sound and Vibration*, vol. 243, no. 3, pp. 427–439, 2001.
- [9] A. H. Benade and M. I. Ibisi, “Survey of impedance methods and a new piezo-disk-driven impedance head for air columns,” *J. Acoust. Soc. Am.*, vol. 81, no. 4, pp. 1152–1167, 1987.
- [10] V. Gibiat and F. Laloë, “Acoustical impedance measurements by the two microphone-three-calibration (tmtc) method,” *J. Acoust. Soc. Am.*, vol. 88, no. 6, pp. 2533–2545, 1990.
- [11] M. van Walstijn, D. M. Campbell, J. Kemp, and D. Sharp, “Wideband measurement of the acoustic impedance of tubular objects,” *Acta Acustica United with Acustica*, vol. 93, pp. 435–446, 2005.
- [12] J.-C. L. Roux, J.-P. Dalmont, and B. Gazengel, “A new impedance tube for large frequency band measurement of absorbing materials,” *Proceedings of Acoustics’08*, 2008.
- [13] C. A. Macaluso and J.-P. Dalmont, “Trumpet with near-perfect harmonicity: design and acoustic results,” *J. Acoust. Soc. Am.*, vol. 129, no. 1, pp. 404–414, 2011.
- [14] D. M. Campbell, “Cornett acoustics : some experimental studies,” *Galpin Society Journal*, vol. 49, pp. 180–196, 1996.

TROMBONE SOUND SIMULATION UNDER VARYING UPSTREAM COUPLING CONDITIONS

Vincent Fréour and Gary P. Scavone

Computational Acoustic Modeling Laboratory (CAML)

Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT)

McGill University

vincent.freour@mail.mcgill.ca

ABSTRACT

The acoustical influence of the upstream airways is an important issue in brass performance. Analyzing the modalities of upstream resonance adjustments around the playing frequency will improve our understanding of vocal-tract tuning and lip-valve mechanics in brass instrument playing. In this study, different conditions of upstream coupling are simulated at the fundamental frequency using a simple one-mass model of the lips coupled to a trombone resonator. Maintaining a constant amplitude of the upstream acoustic pressure, variations of the phase of the upstream relative to the downstream input impedance at the fundamental frequency f_0 result in changes in playing frequency, as well as in the downstream acoustic energy produced. Further analysis shows that this upstream acoustic control can displace the playing frequency near the lip natural frequency, allowing optimal efficiency of the mechanical lip-valve system. These results highlight the possible importance of upstream phase tuning as part of a vocal-tract tuning strategy in brass performance. It further suggests a new experimental method for the estimation of lip natural frequencies on artificial player systems.

1. INTRODUCTION

Recent measurements of vocal-tract influence in trombone and trumpet performance have highlighted the ability of proficient players to create an upstream resonance around the playing frequency that supports or even overrides the regenerative effect of the downstream air-column in the higher register [1–3]. This important contribution of upstream airways also appears to be dependent on the phase of the vocal-tract input impedance at the playing frequency. Experimental investigations conducted on an artificial player system allowed the phase of an upstream resonance at f_0 to be varied independently from the amplitude, leading to variations of the playing frequency around a downstream impedance peak, as well as changes in the downstream acoustic energy produced [4]. These results suggest that an optimal upstream phase tuning, matching with mechanical constraints from the vibrating lips, induces maxi-

mal lip displacement. In this paper, we present the results from numerical simulations using a simple vibrating lip model coupled to a downstream trombone air-column, and involving controlled upstream coupling conditions. Fritz proposed a frequency-domain simulation, as well as an analytical formulation of the frequency shift induced by a vocal-tract resonance in a clarinet model [5]. Contrary to the latter technique, the method proposed in this paper is based on the time-domain resolution of the coupled equations governing the dynamics of the system. This approach allows for the theoretical investigation of the effect of upstream phase tuning on the behaviour of this simple lip model.

2. LIP MODEL

2.1 Mechanical equation

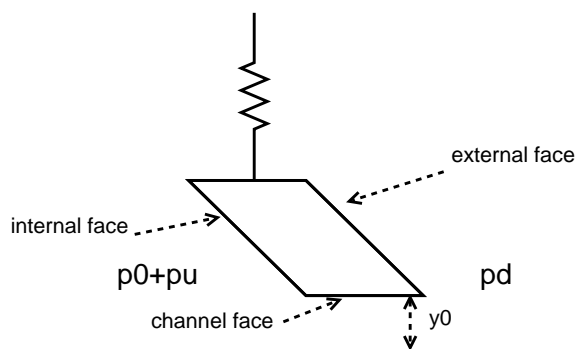


Figure 1. One-mass model of the lips.

The mechanical model chosen for the lips is inspired from [6] and presented in Fig. 1. The two lips are assumed to be identical and placed symmetrically on the mouthpiece. Each lip is allowed to move along a vertical axis and its vertical position relative to the equilibrium position y_0 is noted y . Neglecting the Bernoulli pressure applied on the channel face, an overpressure on the upstream side tends to increase the lip opening area, while an overpressure on the downstream side tends to close it. According to Helmholtz [7], this model corresponds to an outward striking model of the lips. Although a brass player's lips do not seem to support the outward striking model over the full range [8–11], we will assume, in order to keep the

problem simple, that this model adequately simulates the mechanical behaviour of a trombone player's lips over the traditional playing range. The upstream acoustic pressure (mouth side) arising from the upstream acoustical feedback is noted p_u and the quasi-static mouth pressure p_0 . The downstream acoustic pressure (mouthpiece side) arising from the downstream acoustical feedback is noted p_d . In this model, the dynamics of each lip can be represented by the simple second order oscillator equation:

$$\frac{d^2y}{dt^2} + \frac{\omega_{lip}}{Q_{lip}} \frac{dy}{dt} + \omega_{lip}^2(y - y_0) = \frac{F}{m_{lip}}, \quad (1)$$

where ω_{lip} is the lip angular frequency ($\omega_{lip} = 2\pi f_{lip}$), Q_{lip} the quality factor, m_{lip} the mass of one lip and y_0 the lip vertical equilibrium position. F is the vertical component of the force acting on the lip. F is proportional to the pressure difference between mouth and mouthpiece i.e. $F/m_{lip} = (p_0 + p_u - p_d)/\mu$, where μ is the effective mass of the lips, m_{lip} , divided by the effective area of the internal face of the lips; this force therefore supports an outward striking reed behaviour. By substituting F/m_{lip} , Eq. 1 becomes:

$$\frac{d^2y}{dt^2} + \frac{\omega_{lip}}{Q_{lip}} \frac{dy}{dt} + \omega_{lip}^2(y - y_0) = \frac{p_0 + p_u - p_d}{\mu}. \quad (2)$$

2.2 Flow equation

The lip channel is assumed to have constant width b so that the time-varying lip opening area s_{lip} is estimated by the expression $s_{lip} = 2by$. The volume-flow u through this channel is assumed to be quasi-stationary, frictionless and incompressible. Furthermore, we consider $s_{lip} \ll s_{cup}$, with s_{cup} the area of the mouthpiece entryway, and thus no pressure recovery at the mouthpiece cup. Under these assumptions, u can be expressed as a function of the pressure difference $p_0 + p_u - p_d$ across the lips using the Bernoulli equation:

$$u(t) = \sqrt{\frac{2(p_0(t) + p_u(t) - p_d(t))}{\rho}} \cdot 2by(t), \quad (3)$$

where ρ is the average air density.

3. DOWNSTREAM AND UPSTREAM FEEDBACK

3.1 Downstream coupling

According to previous studies [12–14], the downstream acoustical feedback equation can be expressed as follows:

$$p_d(t) = Z_c u(t) + \int_0^\infty r(s) \{Z_c u(t-s) + p_d(t-s)\} ds, \quad (4)$$

where $Z_c = \rho c / s_{cup}$ is the characteristic wave impedance of a cylindrical tube of cross section area s_{cup} and c is the speed of sound. $r(t)$ is the reflection function of the trombone derived from the input impedance Z_d measured experimentally using the experimental set-up described in [15].

As the reliability of the measurements becomes critical below 80 Hz, the real and imaginary parts of Z_d are linearly interpolated to zero below that value. The Fourier transform $R(\omega)$ of $r(t)$ can then be expressed as follows:

$$R(\omega) = \frac{Z_d(\omega) - Z_c}{Z_d(\omega) + Z_c}. \quad (5)$$

A quarter-period sinusoidal window is applied to $R(\omega)$ in order to remove the high frequency noise above 12 kHz. After symmetrization of $R(\omega)$ around the Nyquist frequency, $r(t)$ is calculated as the real part of the inverse Fourier transform of $R(\omega)$. The first 100 ms of $r(t)$ are represented in Fig. 2.

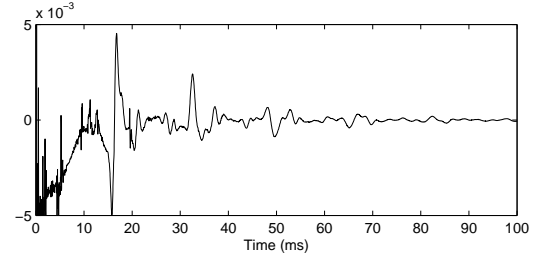


Figure 2. Reflection function $r(t)$ of a 2B King trombone (slide in the closed position) calculated from an input impedance measurement performed using the multi-microphone technique described in [15].

3.2 Upstream coupling

Contrary to previous numerical simulations of upstream coupling [16, 17], our approach does not rely on the modeling of vocal-tract resonances but only simulates an upstream feedback at f_0 . Once oscillations are established and sustained with a constant amplitude, the system is disturbed through the injection of a sinusoidal upstream acoustic pressure of the same instantaneous fundamental frequency as p_d . The amplitude of p_u is maintained constant so that the upstream acoustical feedback energy provided to the lips is constant along the tone duration. However, the instantaneous phase of p_u relative to the phase of p_d at f_0 is varied along the tone duration so that the phase difference $\angle Z_u(f_0) - \angle Z_d(f_0)$ varies linearly in time. Consequently, this protocol enables a specific investigation of the effect of the phase of the upstream impedance at f_0 , independently from the amplitude of the acoustic energy regenerated on the upstream side of the lips. This procedure is implemented according to the following looped sequence:

1. After sample $p_d(n)$ has been calculated from Eq. 4, a 2nd order resonant filter (center frequency $f_c = f_{lip}$) is applied to vector $-p_d$. The output vector p_s is therefore a sinusoidal waveform of amplitude given by the resonant filter coefficients and out of phase with $p_d(f_0)$ by 180° .
2. p_s is normalized to its maximum value over its two last periods and scaled by a constant factor C . This

step enables the amplitude of acoustic upstream energy to be maintained constant, independently from the amplitude of the downstream pressure produced.

3. $p_u(n)$ is set from the normalized and scaled value of p_s , with a time-shift given by a phase-shift parameter $\Phi(n)$; $p_u(n) = p_s[n - \Phi(n)]$.
4. $p_u(n)$ is applied as in Eq. 2 and 3.
5. The procedure is iterated by one time step.

Consequently, when f_0 is constant, this procedure allows the phase difference $\angle P_u(f_0) - \angle P_d(f_0)$ to be linearly varied through the array parameter Φ . Assuming continuity of the volume flow at the reed junction [18], the following equation can be written:

$$\frac{Z_u(\omega)}{Z_d(\omega)} = -\frac{P_u(\omega)}{P_d(\omega)}, \quad (6)$$

which leads to:

$$\angle P_u(f_0) - \angle P_d(f_0) = \angle Z_u(f_0) - \angle Z_d(f_0) + \pi. \quad (7)$$

From Eq. 7, we observe that controlling the phase difference $\angle P_u(f_0) - \angle P_d(f_0)$ results in varying the phase difference $\angle Z_u(f_0) - \angle Z_d(f_0)$.

4. TIME-DOMAIN SIMULATION

Simulations are performed using the following parameter values: $f_{lip} = 310$ Hz, $Q_{lip} = 6$, $\mu = 25$ kg/m², $b = 10$ mm, $y_0 = 0$ mm, $s_{cup} = 4.8$ cm², $\rho = 1.1769$ kg/m³, $c = 347.23$ m/s. It is worth mentioning that in the downstream configuration chosen, the fifth and the sixth resonances of the downstream input impedance are respectively located around 300 Hz and 355 Hz. It is then clear that the chosen lip resonance frequency f_{lip} is between two acoustic resonances, rather than very close to one.

Time-domain simulations are performed by discretization of the differential equations using the forward Euler method and applying the trapezoidal approximation for the integration as performed in [14]. Sampling frequency is set to 48 kHz.

The quasi-static pressure p_0 is specified using a quarter-period sinusoidal onset of 10 ms, a steady phase at 15 kPa, and a quarter-period sinusoidal decay of 10 ms. The phase shift vector Φ contains values varying linearly from 110 to 240; the value contained in Φ at index n indicates the number of shifted samples applied to p_s in order to derive the p_u sample at index n . It therefore refers to the instantaneous phase shift between p_u and p_d waveforms at f_0 . The boundary values in Φ were derived empirically so that the phase shift between p_u and p_d covers the maximal range allowing the system to maintain oscillations. Scaling parameter C is set so that the p_u amplitude is of the same order of magnitude as p_d , hence producing significant effect on the system. Such order of magnitude of $P_u(f_0)/P_d(f_0)$ amplitude ratio has been observed on trombone players in the higher register [1,2]. A quarter-period sinusoidal onset

envelope is applied to the scaling factor so that the p_u amplitude grows smoothly from zero to C when injected into the system.

Simulations start with no upstream feedback ($p_u = 0$). Once the permanent regime of oscillation is reached (time t_1 determined empirically), upstream coupling is added according to the procedure described in the previous section.

5. RESULTS AND DISCUSSION

5.1 Results from simulation

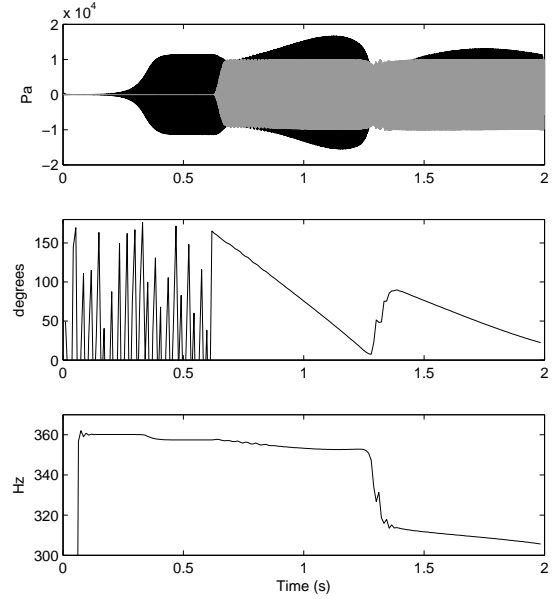


Figure 3. Results from a 96000 sample simulation at 48 kHz. From top to bottom: waveforms of p_d (black) and p_u (gray); phase difference between p_u and p_d at the fundamental frequency; fundamental frequency f_0 .

The waveforms of p_d and p_u calculated from a two-second simulation are presented in Fig. 3. Around 0.6 s, the upstream feedback is introduced and results in significant variations in p_d amplitude as the phase difference $\angle P_u - \angle P_d$ at f_0 varies linearly in time. This variation is also accompanied by changes in f_0 as shown in Fig. 3. At time 1.3 s, a change of register is observed to a lower mode, resulting in a smaller slope of the phase shift due to a smaller value of f_0 . As for the higher mode, f_0 decays slowly while $\angle P_u - \angle P_d$ is decreased until the end of the tone.

5.2 Lip mobility

The frequency response of the lip $G(\omega)$, also referred to as the lip mobility in [18], is defined as follows:

$$G(\omega) = \frac{Y(\omega)}{\Delta P(\omega)}, \quad (8)$$

where $\Delta P(\omega) = P_u(\omega) - P_d(\omega)$.

From Eq. 2, the magnitude and phase of G are represented by solid lines in Fig. 4. As predicted by the mechanical model, the outward striking character of the lip-valve system results in a $+90^\circ$ phase shift between Y and ΔP at the resonance.

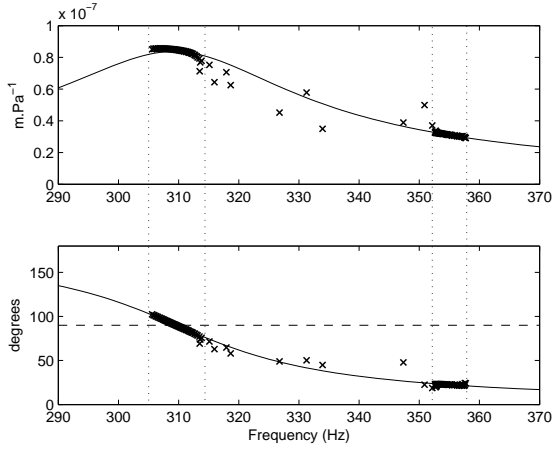


Figure 4. Amplitude (top) and phase (bottom) of lip mobility $G(\omega)$: theoretical value (solid line) and calculated from numerical simulation results (cross). The horizontal dashed line indicates 90° phase angle. The dotted lines indicate the two regions of stable oscillations corresponding to the two produced tones.

In the same figure, the amplitude and phase of the lip mobility at f_0 calculated from simulated waveforms of p_d , p_u and y are represented in crosses. As predicted by the model, and over the frequency range where acoustic energy is observed, the amplitude and phase of G estimated from simulations overlap with its theoretical values. In this particular case, variations of G at f_0 caused by the upstream feedback allow oscillations to occur over a significant frequency range, including the lip natural frequency identified by $\angle G(f_0) = 90^\circ$. This specific tuning point induces a maximum of lip displacement, hence allowing the lip valve to operate with maximal “efficiency”.

5.3 Downstream and upstream input impedances

The influence of downstream and upstream acoustical feedback can be further discussed in terms of input impedances Z_d and Z_u . The total impedance “seen” by the lip-valve $Z = -U/\Delta P$ can be expressed from Eq. 6 as the sum of the downstream and upstream impedances:

$$Z = Z_d + Z_u. \quad (9)$$

Therefore, from experimental measurement of Z_d , the complex quantities Z_u and Z can be derived from Eq. 6 and 9 at frequencies where acoustic energy is observed. The amplitude and phase of Z_d , as well as amplitude and phase of Z_u and Z over the frequency range covered, are represented in Fig. 5. For the highest tone produced with upstream support (353-358 Hz region), a decrease in Z_u amplitude is observed with decrease in frequency. This reveals a decrease in the amplitude ratio P_u/P_d at f_0 , most

probably due to the increase in $|Z_d|$. The upstream coupling also lowers the value of $\angle Z$. As $\angle Z$ becomes closer to zero, oscillations become unstable and a change of regime occurs around 353 Hz to the lower mode. Looking closer at the frequency region between 352 and 358 Hz (Fig. 6), we notice a hysteresis effect at the lower frequency limit; the change in regime occurs around 352.9 Hz when $\angle Z$ is maximized, whereas the lower playing frequency is reached around 352.7 Hz for a larger value of $\angle Z_u - \angle Z_d$ difference. In the region near the lip’s mechanical resonance (305-315 Hz interval), $|Z_u|$ roughly overlaps with $|Z_d|$, and decreasing values of $\angle Z_u$ force $\angle Z$ towards a low value plateau around -100° .

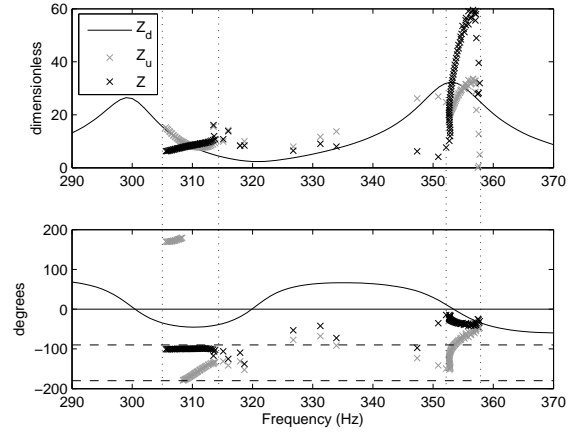


Figure 5. Amplitude (top) and phase (bottom) of $Z_d(\omega)$ measured experimentally (solid line), as well as $Z_u(\omega)$ (gray cross) and $Z(\omega)$ (black cross) at frequencies where acoustic energy is observed during simulations. The two horizontal dashed lines indicate -90° and -180° phase angles. The dotted lines indicate the two regions of stable oscillations corresponding to the two produced tones.

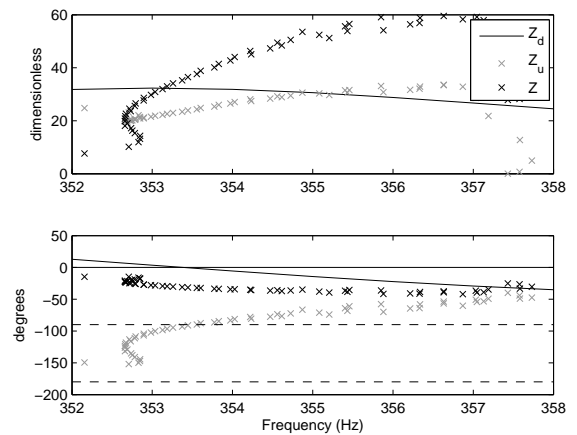


Figure 6. Closer view at Fig. 5 between 352 and 358 Hz.

These observations can be further discussed in light of the linear theory of oscillation. According to previous work [14,

18], the phase condition of regeneration under which sustained oscillations are maintained can be expressed as follows:

$$\angle Z + \angle G = 0. \quad (10)$$

This phase condition relies on the linearisation of the flow equation (Eq. 3) and therefore assumes that s_{lip} and u oscillate in phase.

From Eq. 10, we see that injection of upstream acoustic energy allowing $\angle Z$ to reach values below -90° enables $\angle G$ to extend above 90° . It thus enables oscillations to occur both above and below the lip natural frequency f_{lip} . According to Eq. 10, and given the nature of $\angle Z_d$ ($-90^\circ < \angle Z_d < 90^\circ$), oscillations could not occur at f_{lip} with an outward striking model without appropriate upstream coupling. In the 353-358 Hz region, the lowered value of $\angle Z$ enables oscillations to occur at a lower frequency and hence closer to a maximum in $|Z_d|$. In the 305-315 Hz region, the upstream coupling enables $\angle Z$ to be displaced around -100° and oscillations to occur at the lip natural frequency f_{lip} for which $\angle G = +90^\circ$.

5.4 Extension to the inward striking lip model

The lips of a brass player have been shown to potentially exhibit different mechanisms of oscillation across register [9, 11]. Overall, a transition from a dominant outward striking to dominant inward or upward striking behaviour is observed with increase in playing frequency. The theoretical evaluation presented for an outward model is extended to an inward striking model derived from Eq. 2 by setting the value of μ to -0.04. The negative value of μ results in a negative term F_{lip} on the right side of Eq. 2. This negative force tends to close the valve when the pressure increases on the upstream side of the lips, and to open it when the pressure rises on the downstream side. The other parameters are unchanged except the equilibrium position y_0 set to 0.5 mm; by construction, this valve requires y_0 to be higher than zero in order to initiate oscillations.

The waveforms of p_d and p_u calculated from a two-second simulation are presented in Fig. 7. As in previous simulations, the upstream feedback is introduced around 0.6 s and results in significant variations in p_d amplitude as the phase difference $\angle P_u - \angle P_d$ varies linearly in time. This variation is also accompanied by changes in f_0 but without any change in register (oscillations only occur near the downstream mode located at 300 Hz). Contrary to the outward striking model, f_0 is not linearly correlated to the phase difference $\angle P_u - \angle P_d$. Injection of the disturbing signal results in a small increase in f_0 , preceding a non-linear decreasing phase until the end of the tone. Around 1.5 s, small oscillations are visible in the phase and frequency traces, as well as in the upstream pressure waveform, probably indicating a limit of stability of the system.

The values of G calculated from simulations perfectly overlap with its theoretical value as shown in Fig. 8. Contrary to the outward striking model, oscillations occur on the lower frequency side of the lip resonance. The general decrease in playing frequency induced by the phase

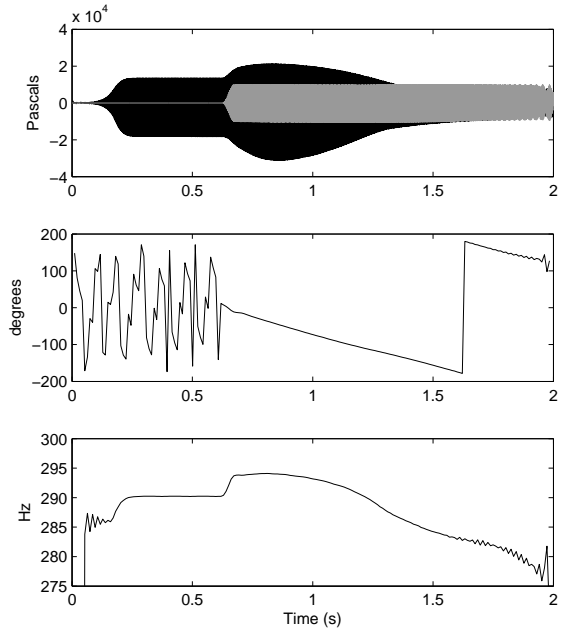


Figure 7. Results from a 96000 sample simulation at 48 kHz. From top to bottom: waveforms of p_d (black) and p_u (gray); phase difference between p_u and p_d at the fundamental frequency; fundamental frequency f_0 .

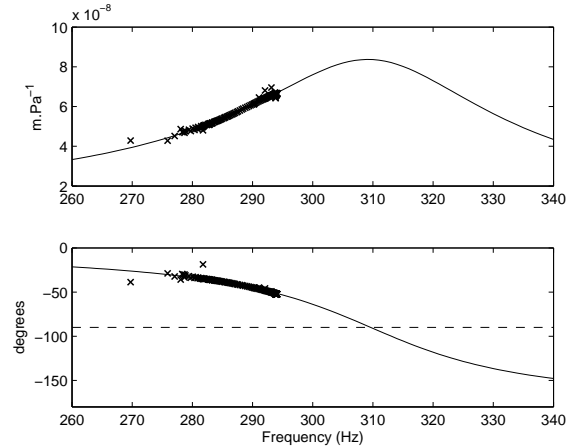


Figure 8. Amplitude (top) and phase (bottom) of lip mobility $G(\omega)$: theoretical value (solid line) and calculated from numerical simulation results (cross). The horizontal dashed line indicates -90° phase angle. The dotted lines indicate the two regions of stable oscillations corresponding to the two produced tones.

shift between P_d and P_u hence results in displacing oscillations away from the lip resonance. However, the increase in f_0 observed at the injection of the upstream perturbation tends to displace oscillations of the system closer to $|G|$ maximum. This contributes to explain the significant increase in the amplitude of p_d waveform observed in Fig. 7 between 0.65 and 0.9 s.

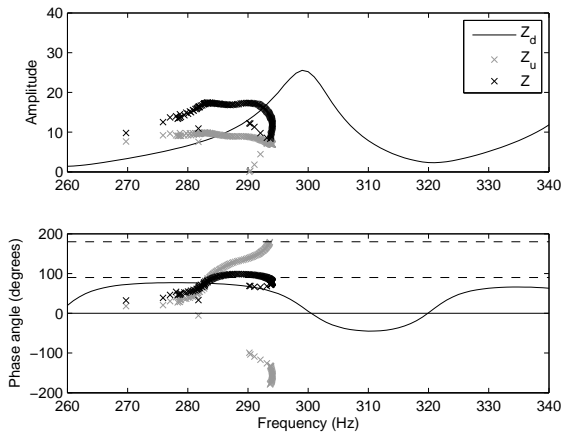


Figure 9. Amplitude (top) and phase (bottom) of $Z_d(\omega)$ measured experimentally (solid line), as well as $Z_u(\omega)$ (gray cross) and $Z(\omega)$ (black cross) at frequencies where acoustic energy is observed during simulations. The two horizontal dashed lines indicate 90° and 180° phase angles.

Looking at the amplitude and phase of Z_d , Z_u and Z at the playing frequency in Fig. 9, we observe that the upstream support allows $\angle Z$ to be raised above 90° in the region where acoustic energy is produced. According to the linear theory of oscillation, the following equality should be verified: $\angle Z = -\angle G$ as stated by Eq. 10. From Fig. 8 and 9, we observe that the validity of this relation is weak: although $\angle Z$ reaches values above 90° , $\angle G$ only decreases down to -50° . This further suggests that the downstream and upstream acoustic feedback introduce significant phase shift between the oscillatory lip opening area and acoustic flow generated at the lips.

Overall, the combined effect of the growing value of $|G|$ and $|Z_d|$ induced by the upstream coupling maximizes lip motion and lip opening area, as well as the downstream acoustic feedback. This results in boosting the amplitude of the acoustic pressure created in the mouthpiece, hence providing optimal efficiency of the sound production process.

6. CONCLUSIONS

These results focus our attention on the relevance of upstream phase tuning on the oscillation of a lip-valve system. In response to previous in-vivo and in-vitro experimental studies on the influence of upstream airways in trombone performance, this study further supports two hypothesis.

Firstly, in a region around the lips' mechanical resonance, proper upstream phase tuning may be part of trombone players' tuning strategy, so that oscillations occur very near to the mechanical resonance of the lips. Vocal-tract tuning may then result from a trade-off between the production of a high amplitude Z at f_0 , and matching f_0 with f_{lip} . However, the reasons underlying a favorable upstream phase tuning remain uncertain for the highest mode excited in

this simulation (far from the lip resonance). This should be the object of further attention.

Secondly, this method based on variations in $\angle Z_u - \angle Z_d$ at f_0 , may be implemented on an artificial player system to allow for estimation of lip natural frequency during sound production. On the contrary to other methods requiring the lips to be at rest during measurements [6, 11, 19, 20], the proposed approach enables the lip natural frequency to be evaluated while the lips are vibrating and under the constraint of the static mouth pressure. Given the sensitivity of the lips to external forces, this approach may produce more reliable estimates of f_{lip} . The results recently obtained in [4] should therefore be further analyzed with regards to identification of lip mechanical resonances.

Finally, further numerical simulations should be performed to evaluate the influence of upstream tuning on more realistic physical representations of the lips. For instance, the approach presented in this paper can be extended to the two-dimensional model proposed by Adachi [14], or two-mass models proposed by Cullen [21]. These investigations should contribute to improve our understanding of the behaviour of lip-reed valves in brass instruments, particularly with respect to the transition between different mechanisms of oscillation.

Acknowledgments

The authors would like to thank the Centre of Interdisciplinary Research in Music Media and Technology, the Schulich School of Music and the Natural Sciences and Engineering Research Council of Canada for helping fund this project.

7. REFERENCES

- [1] V. Freour and G. Scavone, "Vocal-tract influence in trombone performance," in *Proc. 2010 International Symposium on Musical Acoustics*, Sydney, Australia, 2010.
- [2] V. Fréour and G. Scavone, "Investigation of the effect of upstream airways impedance on regeneration of lip oscillations in trombone performance," in *Proc. Acoustics 2012 Nantes Conference*, Nantes, France, 2012, pp. 2225–2230.
- [3] T. Kaburagi, N. Yamada, T. Fukui, and E. Minamiya, "A methodological and preliminary study on the acoustic effect of a trumpet player's vocal tract," *J. Acoust. Soc. Am.*, vol. 130, no. 1, pp. 536–545, 2011.
- [4] V. Fréour, N. Lopes, T. Hélie, R. Caussé, and G. Scavone, "Simulating different upstream coupling conditions on an artificial trombone player system using an active sound control approach," in *Proc. ICA 2013*, Montreal, Canada, 2013.
- [5] C. Fritz, J. Wolfe, J. Kergomard, and R. Caussé, "Playing frequency shift due to the interaction between the vocal tract of the musician and the clarinet," in *Proc. Stockholm Music Acoustics Conference 2003*, 2003, p. 263266.

- [6] J. Cullen, J. Gilbert, and D. Campbell, "Brass instruments: linear stability analysis and experiments with an artificial mouth," *Acta Acustica*, vol. 86, pp. 704–724, 2000.
- [7] H. L. F. Helmholtz, *On the Sensations of Tone*. New York: Dover, 1954.
- [8] N. Fletcher, "Autonomous vibration of simple pressure-controlled valve in gas flows," *J. Acoust. Soc. Am.*, vol. 93, no. 4, pp. 2172–2180, 1993.
- [9] S. Yoshikawa, "Acoustical behavior of brass player's lips," *J. Acoust. Soc. Am.*, vol. 97, no. 3, pp. 1929–1939, 1995.
- [10] F.-C. Chen and G. Weinreich, "Nature of the lip reed," *J. Acoust. Soc. Am.*, vol. 99, no. 2, pp. 1227–1233, 1996.
- [11] J. Gilbert, S. Ponthus, and J.-F. Petiot, "Artificial buzzing lips and brass instruments: Experimental results," *J. Acoust. Soc. Am.*, vol. 104, no. 3, pp. 1627–1632, 1998.
- [12] R. Schumacher, "*Ab initio* calculations of the oscillations of a clarinet," *Acustica*, vol. 48, pp. 71–85, 1981.
- [13] S. Adachi and M. Sato, "Time-domain simulation of sound production in the brass instrument," *J. Acoust. Soc. Am.*, vol. 97, no. 6, pp. 3850–3861, 1995.
- [14] —, "Trumpet sound simulation using a two-dimensional lip vibration model," *J. Acoust. Soc. Am.*, vol. 99, no. 2, pp. 1200–1209, 1996.
- [15] A. Lefebvre and G. Scavone, "A comparison of saxophone impedances and their playing behaviour," in *Proc. 2011 Forum Acusticum*, Aalborg, Denmark, 2011.
- [16] G. Scavone, "Modeling vocal-tract influence in reed wind instruments," in *Proc. Stockholm Music Acoustics Conference 2003*, 2003, pp. 291–294.
- [17] P. Guillemain, "Some roles of the vocal tract in clarinet breath attacks: Natural sounds analysis and model-based synthesis," *J. Acoust. Soc. Am.*, vol. 121, no. 4, pp. 2396–2406, 2007.
- [18] S. Elliot and J. Bowsher, "Regeneration in brass instruments," *Journal of Sound and Vibration*, vol. 83, no. 2, pp. 181–217, 1982.
- [19] V. Fréour, R. Caussé, and K. Buys, "Mechanical behaviour of artificial lips," in *Proc. 2007 International Symposium on Musical Acoustics*, Barcelona, Spain, 2007.
- [20] M.-J. Newton, "Experimental mechanical and fluid mechanical investigations of the brass instrument lip-reed and the human vocal folds," Ph.D. dissertation, University of Edinburgh, Edinburgh, United Kingdom, 2009.
- [21] J. Cullen, "A study of brass instrument acoustics using an artificial reed mechanism, laser doppler anemometry and other techniques," Ph.D. dissertation, University of Edinburgh, Edinburgh, United Kingdom, 2000.

Sensor Based Hand and Lip Weight and Pressure Measurements in Trombone Playing

Tobias Grosshauser, Gerhard Tröster

Institute of Electronics - WearLab

ETH Zurich

grotobia@ethz.ch,

troester@ifee.ee.ethz.ch

Bernhard Hufnagl, Adina Mornell

Hochschule für Musik und Theater

München

bernhard.hufnagl@gmx.de

adina.mornell@hmtm.de

ABSTRACT

Trombone players are exposed to high physical stress while playing. This paper describes a first step towards a physical stress measurement setup. It is based on flexible force sensitive resistors (FSRs) fixed to the trombone. The load of the left hand holding the instrument, the right hand and the overall pressure on the lips is observed. The results show the pressure distribution between the fingers of different players and the overall pressure on the lips. An evaluation of measurements made with six trombone players while playing was conducted and a questionnaire based evaluation about the influence of the sensors while playing was carried out. In addition to audio and video recordings, this approach allows real-world measurements out of lab into the daily playing and exercising environment of the musicians. The setup also provides further possibilities because it can be used in the sense of augmented instruments, allowing new and additional forms of expression, if the sensor data are used for real-time sound synthesis or effects. Further supportive applications in the field of pedagogy, practicing, physical stress avoidance and choosing the right instrument are possible.

1. INTRODUCTION

Many audio and gesture parameters have already been explored and described in practicing, teaching, and performing of musical instruments. The most used technologies are based on the use of video, optical tracking techniques, and sensors like acceleration sensors, magnetometers, and gyroscopes. The method suggested in this paper extends the approved practices. The basic technology used is a high sensitive and flexible FSR based pressure sensor. The main aim was to unobtrusively integrate sensors into a traditional brass instrument, here the trombone, for pressure measurements at the contact points between the musician and the instrument. Borchers et al. in [1] carried out a strain gauge based measurement integrated into the mouthpiece and an intra-oral metal appliance counter part to obtain the overall pressure data of brass players on the lips.

Self-made and easy to build FSR in music context was described by Koehly et al. [2] and on musical instruments for performance analysis by Grosshauser et al. [3] for violin, piano, and bow measurements.

Several sensor based system for violin exist, further vision based approaches are explored, mostly VICON technology, e.g. described by Ng [4]. In piano sensing applications, researchers favored isolated finger work, only including arm movements for the horizontal displacement of the hand. Engelman [5] made measurements of perioral pressure during wind instrument playing and Petiot [6] made measurements of force applied to the mouthpiece while brass instrument playing. On the performance side, several hyperinstruments for performance purpose with similar position-sensing technologies exist (see Machover [7]). The FSR-based setup described here might also be a useful contribution to this field.

2. TECHNICAL DESCRIPTION

The main challenge of the setup presented here is the integration of additional technology into an existing musical instrument without influencing the musician while playing. FSR-based pressure measurement is used in many applications in music making (see Koehly et al. [2]). Pressure measurements of the contact point between musicians and the musical instruments for performance analysis is described by Grosshauser et al. [3] for violin bowing and violinists' chin-/shoulder pressure studies. Flexible FSRs were fixed to the curved surface of the violin bow and chin and shoulder rest. The same technology is now being applied to the trombone. The lightweight, thinness, and flexibility allows unhindered playing and might give a realistic picture of the forces transferred to the instrument by the musician while playing.

In this study, trombone players are observed while playing with the attached setup. The pressure between the fingers on the instrument and the pressure between mouthpiece and the lips is captured while playing. Flexible FSRs (see Fig. 1 and 2) are fixed at the contact points between fingers, lips, and the musical instrument.

2.1 Technical Description of the sensors

The most basic method of interfacing to a FSR is depicted in Fig. 4. In this configuration a FSR is used as a voltage divider. In this case R_{F1} from Fig. 4 (circuit drawing



Figure 1. Trombone with eight pressure sensors fixed to the instrument. The contact points (left hand, right hand, and lips) between the musician and the instrument are covered with flexible FSRs to measure the applied pressure.

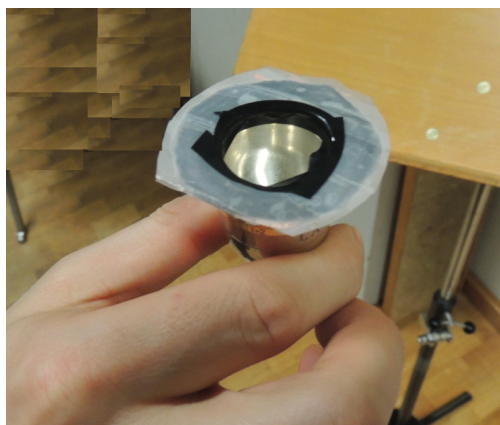


Figure 2. FSR sensors fixed on the mouth piece of the trombone. The sensor is flexible and thin allowing unhindered playing.



Figure 3. Contact points between the trombone and the fingers of the player. The right hand is missing in this picture but also measured. The numbers indicate each contact point at which sensors are attached to the instrument.

partly from Princeton sensor tutorial¹ is the force sensing resistor. The force sensing resistive element consists of a contact element and a resistive surface. The resistance of the contact varies according to the amount of exerted pressure. An increase in force results in a decrease in the value of R_F and an increase in the output voltage. R_{Var} is a potentiometer to calibrate the sensor and adjust the sensitivity.

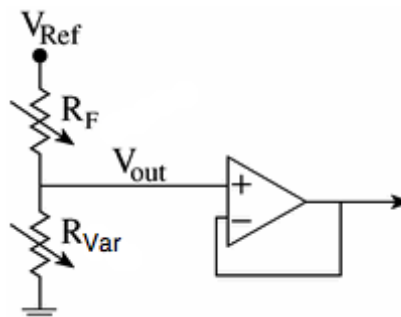


Figure 4. Circuit diagram of a voltage divider based on an adjustable reference resistor (R_{Var}) and a force sensitive resistor (R_F). R_{Var} is used to adjust the sensitivity of the sensor to the pressure range of each contact point. There are eight measurement points, each consisting of this voltage divider/sensor combination.

2.2 Fixation of the Sensors

The eight flexible FSRs are self-made, based on flexible conductive foils. They are self-adhesive and fixed on the mouthpiece and the finger-to-instrument contact areas. These are shown in Fig. 3. If the hand position changed from player to player, the sensor position was adjusted to the right position. The setup further consists of a microcontroller board and a printed circuit board (PCB) with potentiometers to calibrate the eight used measurement channels. The sampling frequency is 100 Hz and could be easily raised e.g. for realtime sound synthesis. For the goal of this study, the lower sampling frequency was sufficient. Further a DV camcorder for audio and video recording is used for aligning the sensor, audio and video data for later examination. Due to latency and synchronization issues, the setup is realized with wire-based data transfer.

2.3 Calibration of the Sensors

Although only the pressure distribution between the contact points between the instrument and the musicians was considered, the method shows that it's possible to measure physical stress of the lips, hand, and fingers. Further, we did a calibration to get an idea of the absolute pressure values in grams. We calibrated the sensors with a standard strain gauge based weighing machine. We placed the weighing machines in front of the mouthpiece of the trombone player to get the values and the relation to the FSR sensor (mouthpiece/weighing machine combination see Fig. 5). The trombone player played a certain scale

¹ <http://soundlab.cs.princeton.edu/learning/tutorials/sensors/node17.html>

with the mouth piece only and simultaneously holding the weighing machine and pressing it against the lips, similar to playing a real trombone.



Figure 5. Setup with a scale for pressure sensor calibration.

3. COURSE AND RESULTS OF THE EVALUATION

3.1 Course of Experiment and Results

Six trombone students were measured while playing a fixed set of notes and scales. Additionally, the study was recorded with an A/V Camera. All measurements were performed on a prepared trombone. Prior to measurement, volunteers were allowed to warm-up in their preferred way for a brief period of time (range 5-10 minutes) and to adapt to the instrument they had to play. An acoustic metronome paced playing speed. All the scores for the segments played were placed on a music stand. All participants performed standing. Each measurement trial took five minutes. All participants played the same notes at the same tempo from the score. After the measurements, the subjects filled out a form with some questions regarding the experiment and were asked for further comments.

3.2 Results of the Questionnaire

The results of the survey are depicted in Table 1. The participants were between 19 and 23 years old, all studying trombone at University of Music and Performing Arts Munich with an overall learning time of the instrument between 4 and 16 years. Two female and six male test subjects participated in the study. The acceptance of the sensor setup was relatively high among musicians. Three of eight trombone players declared the set up had influence on their playing. After questioning the influence, the participants stated, that it was based on the existence of cables on the trombone and did not hinder while playing. Seven participants considered the gathered data as useful, one as not useful. The follow up question, what the data could be useful for, they named possible applications in teaching and practicing support.

3.3 Results of the Measurements

The first measurements show the increase of overall lip pressure in a B major scale from b' to b'' (see Fig. 6). The diagram shows the data in grams, calibrated with a scale, but the error of the FSR values is around 10 %. A tendency of increasing lip pressure of higher notes is clearly distinguishable and congruent with the results of Borchers et al. in [1].

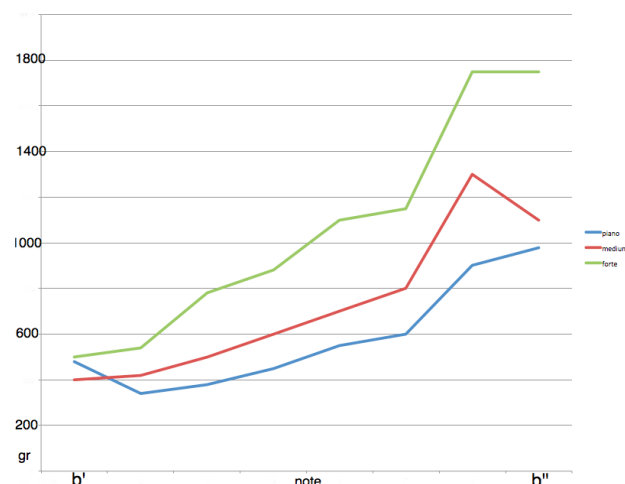


Figure 6. Overall trombone on lip pressure increase within one octave played from b' to b'' .

Fig. 7 shows the increase of the pressure in the hand palm indicating different pattern of each player. Not only the overall pressure level differs while playing at each player, also the distribution while playing is different according to individual playing modalities.

Fig. 8 shows the pressure distribution of the middle finger of different players. Different pressure levels of each musician indicate more or less used force while playing. Also the increase and decrease of pressure while performing the phrase is different for each player. For exact performance analysis, the system has to be further improved, but with this basic setup, individual and diverse pressure changes while playing are already recognizable.

4. DISCUSSION

The described method is a first step into obtaining objective data in music making for daily practicing, teaching, and performing situations, especially for new music and augmented instruments. According to some students' feedback, the system does not hinder playing, but improvements have to be developed. The calibration method will be optimized in the next iterations to improve the measurement accuracy. Furthermore, the work represents a new step towards novel measurement setups to quantify usually hidden parameters pivotal to music making, which are hard to hear and difficult to be objectively shown. Especially in music pedagogy the represented setup could provide meaningful data to the teacher and student. A potential next step will be the implementation of real-time feedback and the objective evaluation of exercises supposed to

Table 1. Statistical data of the experience. The data were obtained with a questionnaire filled by the students after the experiment.

	Sex	Age	Years Instruments Played	Influence of Sensors	Data Useful
P-01	M	20	4	No	Yes
P-02	M	20	12	No	Yes
P-03	F	21	08	Yes	Yes
P-04	M	19	10	Yes	No
P-05	F	21	07	Yes	Yes
P-06	M	23	14	No	Yes
P-07	M	22	15	No	Yes
P-08	M	23	16	No	Yes

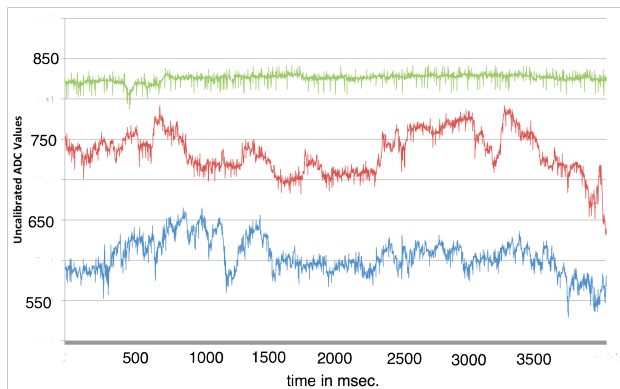


Figure 7. Hand palm pressure (see Fig. 3 measurement area nr. 5) of three trombone players performing the same piece in same tempo and playing mode.

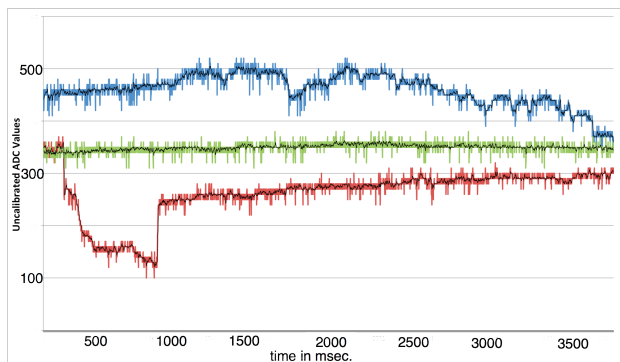


Figure 8. Pressure applied to the middle finger while playing (see Fig. 3 measurement area nr. 3). The variation of the applied pressure shows high pressure differences between three trombone players.

improve e.g. left hand/ right hand and embouchure coordination. Further a long-term study to find possible correlations between pain and measured data would round up the study. Based on this information, the development of alternative exercises might be attempted, which can be, in turn, objectively evaluated by using a comparable setup, regarding their efficiency of improving a specific task. Next steps will also include the simplification of the present setup and its refinement to further enhance the unobtrusiveness and acceptability among musicians. In addition, other relevant parameters not included here should be observed, also in combination with diverse existing technologies. Easy understandable real-time feedback modalities and the use of the sensor setup in other types of musical instruments, especially wind and brass, will be a long-term goal. This may ultimately contribute to the development of augmented instruments and new ways of music making. Further, this additional data allows the combination with any sound synthesis software to enable augmented compositions and improvisations.

5. REFERENCES

- [1] L. Borchers, M. Gebert, and T. Jung, "Measurement of tooth displacements and mouthpiece forces during brass instrument playing," in *Medical Engineering and Physics*, vol. 17, 1995, pp. 567–570.
- [2] R. Koehly, D. Curtil, and M. Wanderley, "Paper fsrs and latex/fabric traction sensors: methods for the development of home-made touch sensors," in *6th International Conference on New Interfaces for Musical Expression, NIME06*. Paris, France, France: IRCAM — Centre Pompidou, 2006, pp. 230–233.
- [3] T. Grosshauser, "Low force pressure measurement: Pressure sensor matrices for gesture analysis, stiffness recognition and augmented instruments," in *8th International Conference on New Interfaces for Musical Expression, NIME08*, S. G. Volpe, A. Camurri, Ed., 2008.
- [4] K. Ng, "3d motion data analysis and visualisation for technology-enhanced learning and heritage preservation," in *AIKED'09: Proceedings of the 8th WSEAS international conference on Artificial intelligence, knowledge engineering and data bases*. Stevens Point,

Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2009, pp. 384–389.

- [5] J. A. Engelman, “Measurement of perioral pressures during playing of musical wind instruments,” in *American Journal of Orthodontics*, vol. 51(11), 1965, pp. 856–864.
- [6] J. F. Petiot, “Measurements of the force applied to the mouthpiece during brass instrument playing,” in *Proceedings of the SMAC03 (Stockholm Music Acoustics Conference 2003)*, vol. 1, pp. 225–228.
- [7] C. J. Machover, “Hyperinstruments: Musically intelligent and interactive performance and creativity systems,” in *International Computer Music Conference*, 1989.

TIME DOMAIN SIMULATION OF STANDING WAVES IN BRASS WIND INSTRUMENTS TAKING NON-LINEAR WAVE STEEPENING INTO ACCOUNT

Wilfried Kausel

IWK - University of Music and
Performing Arts Vienna
kausel@mdw.ac.at

Clemens Bernhard Geyer

IWK - University of Music and
Performing Arts Vienna
clemens.geyer@gmail.com

ABSTRACT

Nonlinear wave steepening up to the degree of shock wave formation is commonly associated with the observed spectral enrichment of brass wind instrument sounds with increasing dynamic level. By modulating fractional delay stages - similar to a method known for producing arbitrary non-linear audio effects - the known physical relationships between local sound pressure, temperature, fluid velocity and wave propagation speed can be enforced in time domain simulations. This way a realistic model of bi-directional non-linear wave propagation can be established. A chain of such non-linear propagation elements combined with traditional Digital-Wave-Guide (DWG) scattering elements and loss filters can be used for closed loop simulations including vibrating lips and realistic radiation conditions. Arbitrary acoustical ducts, like real brass wind instruments defined by their accurate bore profiles, can be simulated and their characteristic sound synthesized. For obtaining the numerical results shown below, ART (Acoustic Research Tool)¹ has been used, which is an Open-Source simulation framework and model library for acoustical simulations in the frequency and time domain. The simulated shock formation distance is in good agreement with analytic results, eg. 4.3-8.6m for sinusoidal stimulus at 170dB between 87 and 175Hz (the low octave of a tenor trombone).

1. INTRODUCTION

In classical linear theory the propagation speed of waves is always assumed to be constant, regardless of whether the wave equation is solved in the frequency domain for a given frequency or whether it is discretized in space and time in order to do a recursive numerical evaluation for a sequence of time steps on an equidistant spatial grid. However, it is known that the speed of sound c is a function of temperature T , actually $\frac{c[m/s]}{c_0[m/s]}$ is proportional to $\sqrt{\frac{T[K]}{T_0[K]}}$. It is further known that for adiabatic processes the temperature depends on the time variant pressure p according to

¹<http://artool.sourceforge.net>

$\frac{T[K]}{T_0[K]} = \left(\frac{p}{p_0}\right)^{\frac{\kappa-1}{\kappa}}$, with κ being the ratio $\frac{C_p}{C_v}$ of specific heat capacitances C at constant pressure and volume.

The temperature- and therefore pressure-dependent sound speed is not the only deviation from the linear theory. A second and typically dominating source of non-linearity is the time variant local fluid velocity v itself. The effective time and space variant wave propagation speed $c_w = c + v$, where $v = \frac{p-p_0}{\rho c_0}$. The fluid velocity v is positive in over-pressure zones where $p > p_0$ and negative where the pressure $p < p_0$ with p_0 being the atmospheric pressure and ρ the density of air.

For unidirectional wave propagation the two different effects are usually combined [1] and an effective wave propagation speed c_w is calculated according to $c_w = c_0 + \frac{(\kappa+1)}{2}v$. This is not possible when standing waves are represented as the sum of a forward and a backward traveling wave. The question whether this decomposition is valid in the case of non-linear propagation has not finally been answered but the work of Menguy and Gilbert [2] is a strong indication that this concept can be applied at least in the weakly nonlinear case.

Doing so, the local speed of sound $c = \sqrt{\frac{\delta p}{\delta \rho}} = \sqrt{\frac{\kappa p}{\rho}}$ according to [3], is related to the local temperature and must be the same for the forward and the backward traveling wave. The cause of its fluctuations is the local physical pressure, the sum of the forward and backward component. The resulting fluid velocity $v = v_f + v_b$, however, which is again the sum of a forward and a backward wave velocity, affects forward and backward component in an opposite way. It acts just like a current on ships cruising in opposite directions. Since $v_f = \frac{p_f - p_0}{\rho c_0}$ and $v_b = -\frac{p_b - p_0}{\rho c_0}$, the effective sum

$$v = \frac{p_f - p_b}{\rho c_0} \quad (1)$$

and is thus proportional to the difference of the partial pressures.

A positive v supports the forward and impedes the backward component, a negative v speeds up the backward and slows down the forward component. At pressure antinodes of standing waves the convective effects will cancel as there is no resulting fluid velocity. At pressure nodes the resulting fluid velocity will be maximal. The component traveling from the positive to the negative pressure maximum will obtain maximum acceleration and the other one maximum deceleration.

The effect of nonlinear wave propagation causes sound waves in cylindrical sections of wind instruments to steepen since wave crests propagate faster than wave troughs and can even overtake the troughs leading to almost infinite gradients in pressure and velocity, the so-called shock fronts. Shock waves exiting a trombone played at some fortissimo level have been observed by Hirschberg [4] in 1996.

Using an analytic expression for the shock formation distance he came to the conclusion: “As we will explain further on, the nonlinearity in the transfer function depends crucially on a primary nonlinearity of the source.” and “As X_s is determined by dp/dx we understand the importance of the initial nonlinearity of the source which we discussed above.”

He refers to the fact, that typical brass wind instruments do not have sufficiently long cylindrical sections to generate shock waves from a sinusoidal stimulus, at least at realistic playing pressures and in their lower register. Since they do generate shock waves rather easily even in the very low register, it is required that the lip generated stimulus pressure does contain at least a few higher harmonics and/or non-linear wave propagation does take place in the regeneration.

The shock formation distance $X_s = \frac{2c_0}{(\kappa+1)(-\delta u/\delta x)_{max}}$ (according to [1]) is the minimum distance a plane wave with a certain initial velocity or pressure gradient has to travel inside a cylindrical tube until nonlinear wave steepening makes it a shock front with infinite gradient. For a sinusoidal stimulus with frequency f and pressure amplitude \hat{p} at atmospheric mean pressure P_{at} we obtain

$$X_s = \frac{\kappa}{\kappa+1} \frac{1}{\pi} \frac{P_{at} c_0}{\hat{p} f}. \quad (2)$$

Results for a trombone have already been given in the abstract.

The length of the cylindrical part of prominent brass wind instruments has been measured by Benade and can be found in [5] on p192. It is about 40cm (29% of 136cm) in trumpets, 143 cm (52% of 275cm) in trombones and 228cm (61% of 374cm) in French horns. These proportions vary as valves or slides are operated. The above numbers are for instruments with no valves pressed and the slide fully in.

Knowing that brassy sound can easily be generated by skilled players even in the lowest register and without extreme blowing pressures, we can conclude with Hirschberg that there must be relatively steep pressure gradients already present in the mouthpiece. Alternatively Hirschberg's assumption that high frequencies are radiated and do not participate in the regeneration process can only be correct for a few high notes which are close to the cut-off frequency of the bell.

As far as the low register is concerned it is obvious that several higher harmonics will be below cut-off and therefore reflected at the bell back into the regeneration loop. A second harmonic already halves the shock formation distance ($\delta p/\delta x = \hat{p}2\pi(f + 2f/2 + 3f/3 + \dots)$) and - if reflected at the bell - triples the available run length for wave steepening.

This should be enough motivation to include non-linear wave steepening in the study of the regeneration process.

Estimating brassiness for various types of brass wind instruments and the influence of certain parameters or the player's embouchure on the brassiness potential has become an important topic in brass wind instrument acoustics in the recent years [6–11].

2. MODELING NON-LINEAR WAVE PROPAGATION

By extending the acoustic wave equation for propagation of plane waves in a viscous heat-conducting fluid and applying the averaging method to non-linear resonant oscillations in closed tubes, Ochmann was able to obtain an inhomogeneous Burgers equation which he solved in 1985 for the case of piston excitation by linearization and transformation into a Hills equation [12].

In 2000 Menguy and Gilbert managed to obtain uncoupled Burgers equations representing forward and backward traveling waves [2]. This indicates that a local linear approximation may be valid when the coupling between both components is properly taken into account [13].

Thompson and Strong were using the Burgers equations in 2001 to introduce nonlinear wave propagation in frequency domain simulations of a tenor trombone [14]. In 2008 Gilbert et al. published a simulation tool for brassiness studies [15] based on their theoretical results.

The Burger equations contain fractional derivatives and other mathematical challenges and there is no analytical solution in the general case. Even for special cases solutions can only be approximated using series expansions and iterative approaches. Especially for real time sound synthesis, a computationally more efficient method would be very desirable.

Since time domain simulators separate wave propagation and wave scattering it is straightforward to keep everything but modify the propagation elements to reflect the modulation of the speed of sound. Msallam proposed to use fractional delay stages to simulate a varying speed of sound by modulating the propagation delay according to the local sound pressure [16, 17]. This method did not provide any distributed coupling between forward and backward branch and was therefore completely nonphysical.

In 2010 Cooper published an improved version of this method [18] as an audio effect for “brassification” of sounds. He mentioned wind instrument sound synthesis as a possible application but he did not go so far as to derive any physical model parameters nor to show any results for standing waves at resonance conditions.

3. PROPOSED MODEL OF WAVE STEEPENING

In order to use Cooper's method for physical modeling several essential details have to be changed. First, the time varying delay must be derived from basic physical laws instead of choosing it arbitrarily with the aim of producing an attractive brassification effect. With the nominal propagation length of a single duct element $L = \frac{c_0}{f_s}$ at sampling rate f_s and sampling interval $D_0 = \frac{1}{f_s}$ we obtain a formulation for the time varying temperature dependent unit

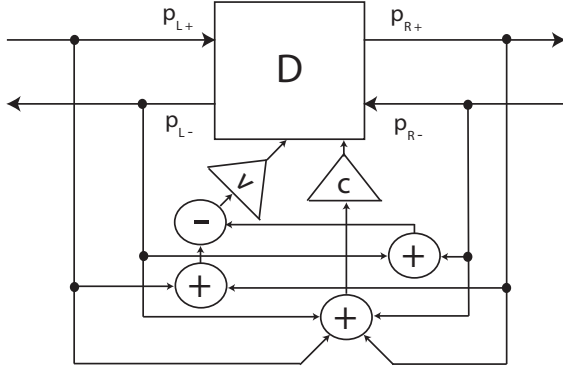


Figure 1. Bidirectional delay element with delay D taking local sound speed c and velocity difference v into account

delay D_T according to

$$\frac{D_T}{D_0} = \frac{c_0}{c} = \left(\frac{p_0}{p}\right)^{\left(\frac{\kappa-1}{2\kappa}\right)} \quad (3)$$

Second, the coupling of forward and backward traveling waves has to be changed. Cooper proposes a single-sided coupling, which means that the physical local pressure, the superposition of the forward and backward traveling component causing the temperature change which is one source of non-linearity, is taken at the right side of each propagation element to be used as a control signal for modulating the propagation delay. But this has as a consequence that forward and backward traveling waves are no longer treated symmetrically. The non-linear effect of the forward traveling wave is shifted by one unit in time and space.

Considering steep gradients in pressure and a spatial resolution chosen to be as coarse as possible for economical reasons, it is doubtless better to have symmetric treatment of forward and backward wave as shown in Figure 1. This can easily be achieved by taking the arithmetic mean between the physical pressure left and right of the propagation element to control this part of the propagation delay.

Third, the local fluid velocity has to be taken into account as it is responsible for about 80% of the total effect. Therefore an additional delay component due to the effective fluid velocity according to Equation 1 has to be added which is controlled by the sum of the wave velocities and therefore - because of the negative sign of the backward velocity - by the difference of the forward and backward wave pressure. It makes sense to average left and right side pressure of both waves and to use the local speed of sound which has already been calculated.

Another most important implementation detail concerns the implementation of the fractional delay itself. In order to avoid higher order interpolation filters a first order Lagrange stage (linear interpolation) is desirable. This requirement is not so much due to computational efficiency considerations but much more due to the fact that a time varying propagation delay must be implemented. First, filter parameters would have to be recalculated every time the propagation delay changes. Second, efficient interpolation filters are typically FIR filters which usually behave

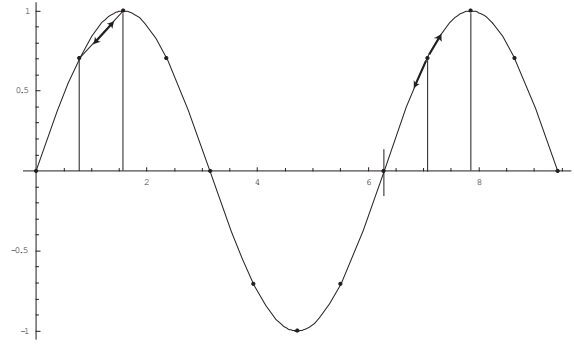


Figure 2. Linear interpolation between two samples (left) and at and around an exact sample position (right)

in a very ugly way when coefficients are modified dynamically.

But linear interpolation has a severe draw back. All interpolated values have a smaller magnitude than the original function value would have, at least if the second derivative of the signal has always the opposite sign as the signal itself. As this is true for all signals with a strong fundamental, extremely strong numerical damping would occur (see left case of Figure 2).

By carefully studying this problem a simple work around was found. By choosing an interpolation interval of two time steps with three breakpoints instead of one time step with two breakpoints, the average interpolation result without non-linearity now coincides with the existing signal sample at the center of the interpolation range instead of falling exactly in between two neighboring samples (see right case of Figure 2). Considering the non-linearity as a second order effect, the interpolation error suddenly becomes third order. Actual simulation results did prove that after this modification numerical damping caused by interpolation no longer was an issue.

4. ART IMPLEMENTATION

The proposed model has been implemented using the Open-Source simulation framework and model library ART [19] (Acoustic Research Tool), which is described in greater detail in another paper of the authors [20]. It is a recent result of the work of work group WG2 of the Technical Committee Musical Acoustics TCMA of the European Acoustics Association EAA [21].

In ART it is straightforward to declare a transmission element with the proposed properties. All elements representing bore segments of an instrument are created in a loop starting at line 8 and ending at line 35 of the heavily commented Python code shown in Figure 3. Each delay element has an input and an output port on its left and on its right side carrying the bidirectional pressure signals. The suffixes p (plus) and m (minus) of the port designators indicate the forward and backward propagating signal branch while port number 1 is the left and port number 2 the right port of all delay and junction modules.

The interpolation requires a conditional assignment depending on whether the time variant propagation delay of

```

1 # *****
2 # create simulator named TimeSimulator and save the object pointer in the variable sim
3 sim = ARTCreateSimulator("TimeSimulator", "TimeDomain", "")
4 # create list objects for the object pointers of all delay modules and scattering junctions created later
5 dMod = list(); jMod = list()
6 # append propagation and scattering elements, define their ports and local parameters, connect ports
7 # do this for tlen+1 elements in a loop starting with element 0 up to element tlen
8 for i in range(0, tlen+1) :
9 # append TimeModules named Delay0, Delay1, Delay2... to simulator sim, store their addresses in list dMod
10 dMod.append(ARTCreateModule(sim, "Delay{0}".format(i), "TimeModule"))
11 # add local parameter c/c0 and how it depends on p to the module which has just been created
12 ARTAddParam(dMod[i], "c = 0.192752*((1.e5+p)^0.143)") # c/c0 = ((p+p0)/p0)^0.143 w/ p0=100k
13 # add local parameter v/c0 and how it depends on dp to the module which has just been created
14 ARTAddParam(dMod[i], "v = 0.00241583475*dp/c") # v/c0 = p/(c rho)/c0 w/ rho=1.2068, c0=343
15 # add local parameter k and how it depends on c and v to the module which has just been created
16 ARTAddParam(dMod[i], "k = 2/(c+v)-2.") # D/D0 = c0/(c0(c+v)) w/ D0=2, D=2+k: 2+k=2/(c+v)
17 # add local parameter p and how it is to be derived from past samples at ports p1+ -> p2+ and p1- <- p2-
18 ARTAddParam(dMod[i], "p = (p1p[t-1]+p1m[t-1]+p2p[t-1]+p2m[t-1])/2")
19 # add local parameter dp and how it is to be derived from past samples at ports p1+ -> p2+ and p1- <- p2-
20 ARTAddParam(dMod[i], "dp = (p1p[t-1]+p2p[t-1]-p1m[t-1]-p2m[t-1])/2")
21 # calculate current sample at output port p2+ (conditional assignment based on sign of k)
22 ARTAddPort(dMod[i], "p2p[t] = (k>=0)?((1-k)*p1p[t-2]+k*p1p[t-3]):(-k*p1p[t-1]+(1+k)*p1p[t-2])")
23 # calculate current sample for output p1- of the backward traveling wave
24 ARTAddPort(dMod[i], "p1m[t] = (k>=0)?((1-k)*p2m[t-2]+k*p2m[t-3]):(-k*p2m[t-1]+(1+k)*p2m[t-2])")
25 # add junction elements named Junction0, Junction1... to simulator sim, store their addresses in list jMod
26 if i < tlen: # do not append a junction after last delay element
27 jMod.append(ARTCreateModule(sim, "Junction{0}".format(i), "DWGCylinderJunctionModule"))
28 # read r1 and r2 from previously created array "borelist" and set the radius parameters of each junction
29 ARTSetParam(jMod[i], "r1 = {0}".format(borelist[i]))
30 ARTSetParam(jMod[i], "r2 = {0}".format(borelist[i+1]))
31 # connect junction input p1+ with delay output p2+ as well as delay input p2- with junction output p1-
32 ARTConnectPorts(sim, "Junction{0}.p1p = Delay{0}.p2p; Delay{0}.p2m = Junction{0}.p1m".format(i))
33 if i != 0: # connect to previous group
34 # connect inputs of previous group (marked {0}) with outputs of current group (marked {1}) and vice versa
35 ARTConnectPorts(sim, "Junction{0}.p2m = Delay{1}.p1m; Delay{1}.p1p = Junction{0}.p2p".format(i-1, i))
36 # *****

```

Figure 3. Python code for setting up instrument.

an element needs to be smaller or greater than two standard unit delays. In lines 22 and 24 of the Python code the fractional delay for the forward and backward traveling wave is implemented.

Depending on parameter k (with $[-1 \leq k \leq 1]$) a linear interpolation between time step $t-1$ and $t-2$ or between time step $t-2$ and $t-3$ is performed. A positive $k \leq 1$ corresponds to the longer delay between 2 and 3 time steps, a negative $k \geq -1$ will create a shorter delay between 1 and 2 time steps. The interpolation itself is a weighted average of the two samples with the weights k and $1-k$.

For $k = 0$ the element is in the linear mode and a standard propagation delay of exactly 2 time steps is created. As non-linearities are small for all physical problems the magnitude of k will be very small, which makes interpolation errors negligible.

The parameter k (Figure 3, line 16) is calculated according to

$$k = \frac{2}{(c/c_0 + v/c_0)} - 2 \quad (4)$$

for a unit delay $D_0 = 2$ and an actual delay $D = D_0 + k$ with $D/D_0 = c_0/(c+v)$ and Equations 1 and 3. The ratios $c/c_0 = ((p_s + p_0)/p_0)^{(\kappa-1)/(2\kappa)}$ (line 12, parameter c) and $v/c_0 = p_s/(c\rho)/c_0$ (line 14, parameter v) only depend on the local sound pressure p_s (parameter p) and some basic physical constants like air density $\rho = 1.2068$, $\kappa = 1.4$, atmospheric pressure $p_0 = 100k$ and average

sound speed $c_0 = 343m/s$.

The total physical sound pressure p is the average between the sum of p_{1+} and p_{1-} on the left side and the sum of p_{2+} and p_{2-} on the right side (line 18). The difference between the forward and backward traveling pressure wave required in Equation 1 is the average of p_{1+} and p_{2+} minus the average of p_{1-} and p_{2-} (line 20).

The junction modules in ART are predefined scattering modules representing the reflection and transmission coefficients at bore diameter discontinuities or of short conical slices as described by Walstijn [22].

Line 3 creates and initializes a simulator for the time domain. Line 5 creates Python lists for keeping track of all delay and junction elements of the instrument. These elements are then being created and initialized in the loop.

Line 10 creates the delay elements, line 27 the junction elements, which are inserted in between the delay elements. In lines 29 and 30 bore diameters from the instrument's bore profile are assigned to the corresponding local parameters. The spatial discrimination of the bore profile has been chosen in a way to make $L = 2c_0/f_s$.

In lines 32 and 35 right and left sided ports belonging to the forward or backward propagating branch are interconnected to create uninterrupted signal chains in plus and minus direction.

Closing the loop by establishing appropriate reflection conditions at both ends of the instrument, as well as injecting a stimulus signal and plotting the results, are not

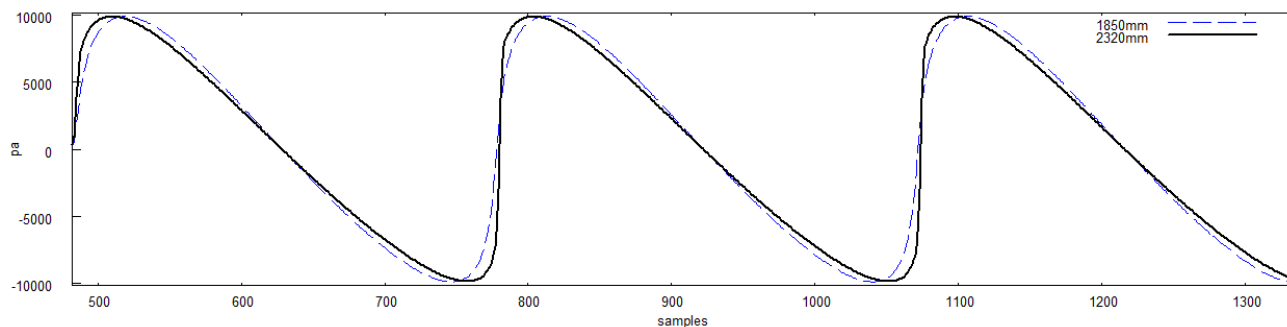


Figure 4. Simulated shock formation distance in a cylindrical tube for sinusoidal stimulus at 10kpa and 300Hz (analytical result is 1857mm, signals plotted at distances of 1850mm and 2320mm).

shown in the code snippet. However, the whole working example is included in the ART package, just like several other examples like "trumpet sound generation with single mass lip model" and "clarinet sound generation with reed", and can be downloaded from the SourceForge server.

5. RESULTS

First, a forward propagating wave stimulated by a sinusoidal signal has been simulated in order to verify the resulting shock formation distance using the analytical expression of Equation 2. The time variant signal at two positions inside a long cylindrical tube has been plotted. One slightly below the theoretical shock formation distance and a second one above.

It can be seen that the simulation behaves as expected by theory, creating a shock front after about 2m. It should be noted here, that a 10kpa amplitude in the cylindrical part of a real brass wind instrument would require an unrealistic mouthpiece pressure of about 40kpa (183dB). This can be concluded from the trombone measurements shown by Hirschberg in [4]. Considering that the simulation has been done at 300Hz and that the low register of a trombone extends down to below 100Hz, it can be concluded that the cylindrical part of trombones will not be long enough to facilitate shock waves unless regeneration extends the available run length and adds higher harmonics to the stimulus signal.

Therefore a closed loop simulation demonstrating non-linear regeneration has been performed in a second step. A cylindrical tube of length 3092mm composed from 401 bidirectional transmission elements each having a length of 7.71mm has been excited at its third resonance frequency of 137.5Hz by means of a sinusoidal pressure signal. This 1000pa stimulus pressure was added to the forward propagation path at the left end of the tube. The sampling frequency was chosen to be 88200Hz which results in a segment delay of 2 sampling intervals at a speed of sound of 340m/s.

The tube was terminated by a reflection coefficient of -0.95 at the open end on the right side and by a perfect reflection coefficient of 1 at the closed end on the left side. A realistic bore profile with a physical radiation impedance and distributed viscous losses will be studied in a future work.

The resulting pressure signals from the stimulated end of a cylindrical tube during a starting transient are shown in Figure 5. The upper Figure 5a shows the total sound pressure and the forward and backward traveling components at the location of the stimulus. Figure 5b shows the derivatives illustrating the steady steepening of the total pressure at the mouthpiece side which creates higher harmonics during regeneration.

Figure 5c shows the corresponding pressure signals at the open end of the tube. As amplitudes are growing higher, characteristic pressure spikes appear, which indicates the formation of shock waves.

6. CONCLUSIONS

Non-linear sound propagation within brass wind instruments played at high dynamic levels can be modeled in a physically correct way by modulating the bidirectional propagation delay of elementary wave guide elements according to a spatially smoothed but time varying pressure profile, which is the source of a corresponding temperature and fluid velocity profile and is linked to a derived speed of sound profile.

The underlying assumptions are adiabatic pressure variations and a spatial discrimination fine enough to have ten or more elements per wave length of the fundamental frequency which is considered. The spatial resolution limits the pressure gradient, which means that good resolution in space and time is indispensable when the buildup of shock-like conditions are to be studied.

The proposed simulation method allows to study the effect of nonlinearities and its role in regeneration within instruments represented by their accurate bore profiles. It allows the use of different viscous loss filters and proven radiation models. While all simulations presented in this paper have been done for a cylindrical tube, it will be part of future work to simulate real instruments and to compare theoretical predictions with measured data.

Simulations indicate that physical non-linearities are too weak to allow shock wave formation at possible playing levels in the low register of brass wind instruments if stimulated by a purely sinusoidal signal and if wave steepening is only taken into account for the forward path. This is consistent with unpublished observations of Gregor Widholm

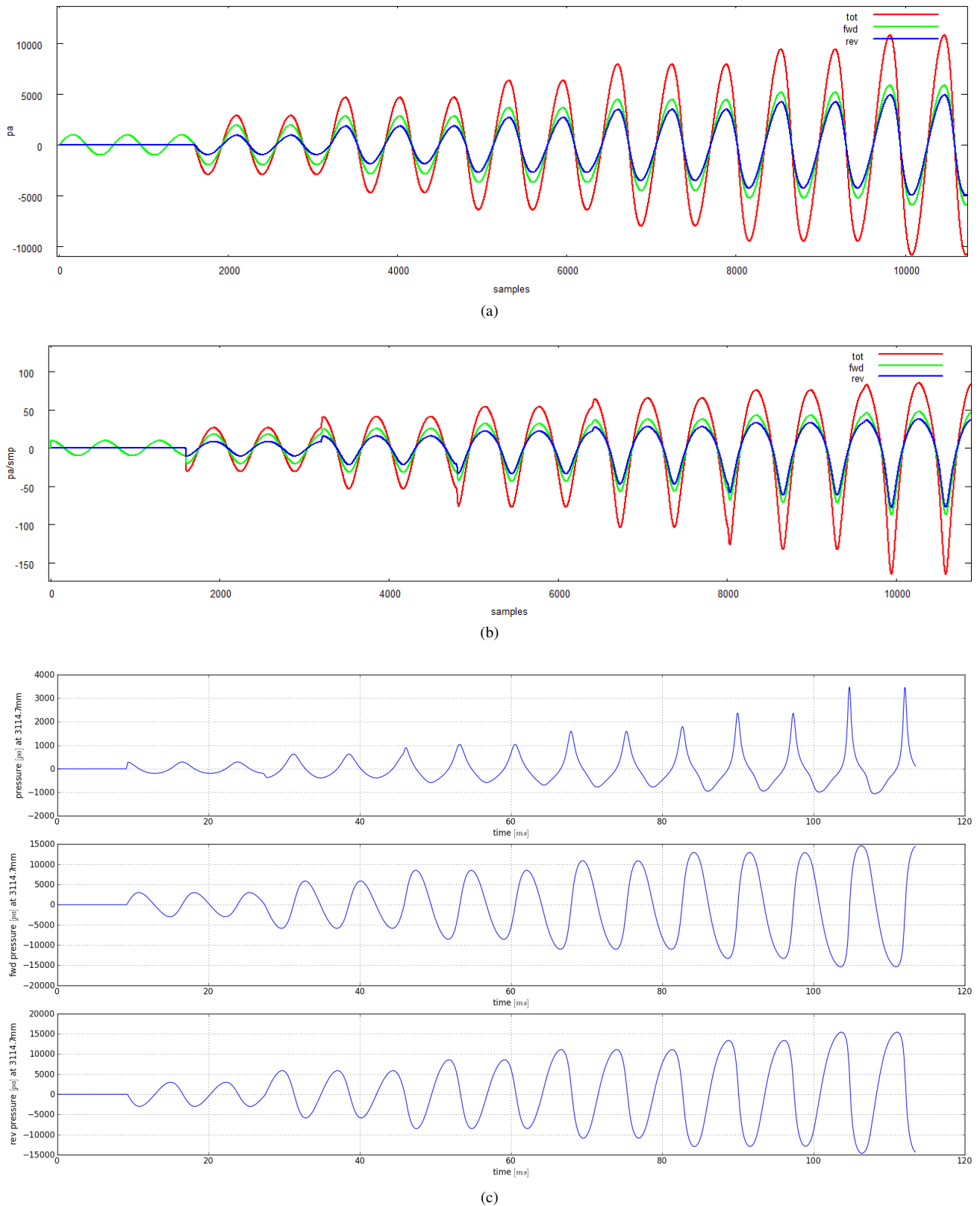


Figure 5. Sound pressure of standing, forward and reverse wave during starting transient of 3rd resonance, calculated at the stimulus location: (a) sound pressure in [pa] (b) derivative of sound pressure in [pa / sample] (c) total, forward and reverse pressure at the open end of the tube.

who used a siren with sinusoidally oscillating lip orifice to drive a French horn at fortissimo playing levels.

In order to get shock waves in musical instruments at low

frequencies it is required for the exciting sound flow signal to already contain a certain number of higher harmonics which in turn are being amplified and augmented by

a regime of still higher harmonics due to non-linear wave steepening and non-linear regeneration.

7. REFERENCES

- [1] P. M. Morse and U. Ingard, *Theoretical Acoustics*. Princeton, New Jersey: Princeton University Press, 1987.
- [2] L. Menguy and J. Gilbert, "Weakly nonlinear gas oscillations in air-filled tubes; solutions and experiments," *Acta Acustica united with Acustica*, vol. 86, no. 5, pp. 798–810, 2000.
- [3] A. Pierce, *Acoustics: An Introduction to its Physical Principles and Application*. Melville NY: Acoustical Society of America, 1989.
- [4] A. Hirschberg, J. Gilbert, R. Msallam, and A. P. J. Wijnands, "Shock waves in trombones," *Journal of the Acoustical Society of America (JASA)*, vol. 99, no. 3, pp. 1754–1758, 1996.
- [5] A. Benade, *Horns, Strings, and Harmony*. New York: Dover Publication, 1992.
- [6] A. Myers, M. Campbell, J. Gilbert, and R. W. Pyle, "The brassiness potential of chromatic instruments," in *10ème Congrès Français d'Acoustique*, 2010.
- [7] A. Myers, R. W. Pyle, J. Gilbert, and M. Campbell, "The influence of bore size on brassiness potential," in *Proceedings of the Second Vienna Talk on Music Acoustics*, W. Goebel, Ed. Vienna, Austria: Institute of Musical Acoustics (Wiener Klangstil), 2010, pp. 102–105.
- [8] R. Pyle, "Is it the player or is it the instrument," in *Proceedings of the Second Vienna Talk on Music Acoustics*. Univ. of Music, Vienna, Austria: edited by Werner Goebel, Inst. of Music Acoustics (IWK), 2010, pp. 113–116, invited paper.
- [9] L. Norman, J. P. Chick, D. M. Campbell, A. Myers, and J. Gilbert, "Player control of 'brassiness' at intermediate dynamic levels in brass instruments," *Acta Acustica united with Acustica*, vol. 96, no. 4, pp. 614–621, 2010.
- [10] L. Norman, J. Chick, M. Campbell, and A. Myers, "Embouchure control of brassiness at constant pitch and dynamic level in orchestral horn playing," in *Proceedings of the International Conference on Acoustics NAG/DAGA 2009*, Deutsche Gesellschaft für Akustik e.V. (DEGA). Rotterdam, Netherlands: M. M. Boone, TU Delft, 2009, pp. 862–865.
- [11] J. Gilbert, D. M. Campbell, A. Myers, and R. W. Pyle, "Differences between brass instruments arising from variations in brassiness due to nonlinear propagation," in *Proceedings of the International Symposium on Musical Acoustics*, Universitat Politècnica de Catalunya. Barcelona, Spain: Institut d'Estudis Catalans, 2007.
- [12] M. Ochmann, "Nonlinear resonant oscillations in closed tubes - an application of the averaging method," *Journal of the Acoustical Society of America (JASA)*, vol. 77, no. 1, pp. 61–66, 1985.
- [13] P. L. Rendón, F. Orduna-Bustamante, D. Narezo, and A. Pérez-López, "Nonlinear progressive waves in a slide trombone resonator," *Journal of the Acoustical Society of America*, vol. 127, no. 2, pp. 1096–1103, 2010.
- [14] M. W. Thompson and W. J. Strong, "Inclusion of wave steepening in a frequency-domain model of trombone sound production," *J. Acoust. Soc. Am.*, vol. 110, no. 1, pp. 556–562, 2001.
- [15] J. Gilbert, L. Menguy, and D. M. Campbell, "A simulation tool for brassiness studies (I)," *Journal of the Acoustical Society of America (JASA)*, vol. 123, no. 4, pp. 1854–1857, 2008.
- [16] D. Msallam and R. Caussé, "Physical model of the trombone including nonlinear effects. application to the sound synthesis of loud tones," *ACTA ACUSTICA/ACUSTICA*, vol. 86, no. 3, pp. 725–736, 2000.
- [17] R. Msallam, S. Dequidt, S. Tassart, and R. Caussé, "Physical model of the trombone including nonlinear propagation effects," in *Proceedings of the International Symposium of Musical Acoustics [ISMA 1997]*, vol. 2, Institute of Acoustics. St Alban (UK): MYERS, Arnold, 1997, pp. 419–424.
- [18] C. M. Cooper and J. S. Abel, "Digital simulation of 'brassiness' and amplitude-dependent propagation speed in wind instruments," in *Proc. of the 13th Int. Conference on Digital Audio Effects (DAFx-10)*, Graz, Austria, 2010.
- [19] "Acoustic Research Tool (ARTool) homepage," 2013, accessed 2013-04-06. [Online]. Available: <http://artool.sourceforge.net/>
- [20] C. B. Geyer and W. Kausel, "An open-source framework for time-domain simulations," in *Proceedings of the SMAC/SMC 2013*. Royal Swedish Academy of Music.
- [21] "Work group wg2 of the technical committee musical acoustics tcma of the eaa," 2013, accessed 2013-04-06. [Online]. Available: <http://www.eaa-fenestra.org/technical-committees/ma/workgroups/wg2/>
- [22] M. v. Walstijn and M. Campbell, "Discrete-time modeling of woodwind instrument bores using wave variables," *JASA, Journal of the Acoustical Society of America*, vol. 113, no. 01/2003, pp. 575 – 585, 2003.
- [23] W. Kausel, "Time resolved sound field visualization of french horn: Effect of wall vibrations and formation of shock waves," in *Proceedings of the MAN Summer Meeting on Wind Instrument Acoustics*. Edinburgh, GB: University of Edinburgh, 2009, p. <http://www.music.ed.ac.uk/euchmi/man/>.

CONTROL OF AN ARTIFICIAL MOUTH PLAYING A TROMBONE AND ANALYSIS OF SOUND DESCRIPTORS ON EXPERIMENTAL DATA

Nicolas Lopes

IRCAM-CNRS-UPMC, UMR 9912,
1 place Igor Stravinsky,
75004 Paris, France
nicolas.lopes@ircam.fr

Thomas Hélie

IRCAM-CNRS-UPMC, UMR 9912,
1 place Igor Stravinsky,
75004 Paris, France
thomas.helie@ircam.fr

René Caussé

IRCAM-CNRS-UPMC, UMR 9912,
1 place Igor Stravinsky,
75004 Paris, France
rene.causse@ircam.fr

ABSTRACT

This paper deals with a robotized artificial mouth adapted to brass instruments. A technical description of the robotic platform is drawn, including calibrations, initialization processes, and modes of control. An experimental protocol is proposed and the repeatability is checked. Then, experiments are conducted on a trombone for several types of quasi-static controls. Sound descriptors (fundamental frequency, roughness, energy) of measured acoustic signals are estimated and used to build cartographies indexed by the control inputs. An analysis reveals that several stable notes can easily be reached using a basic mapping with respect to these control inputs. However, a histogram of fundamental frequencies shows that some notes in the high range that can be played by musicians are not reached by the artificial mouth. It also reveals that some notes are difficult to play in the middle range. This exploration suggests some possible improvements of the machine that are finally discussed.

1. INTRODUCTION

Brass wind instruments are self-sustained musical instruments. Their self-oscillations are due to the non-linearity of the aero-elastic valve, namely, the jet coupled to the lips, which is loaded by the acoustic resonator. But, although the musician's control of the valve is crucial, it is very difficult to study this bio-physical system and to make "in vivo" measurements. For this reason, artificial mouths have been developed, see e.g. [1–5].

This paper deals with a robotized version of such systems. This robotization was initiated during the CONSONNES project [6]. It also involved mechatronic projects in an engineering school [7] and several internships [8–11]. Some first results and evolutions of the machine functionalities have been presented in [12–14]. In the last one, it has been showed that sequences of a few trumpet notes could be played with a simple open loop control, using a "hand-tuned" mapping based on a sound descriptor analysis¹.

¹ A movie can be downloaded on the following website:
<http://recherche.ircam.fr/anasy/helie/Brasstronics/FilmBrasstronics2011.avi>

In this paper, a systematic approach is proposed and based on (1) calibrations and initialization processes, (2) an experimental protocol and repeatability tests, and (3) sets of cartographies of sound descriptors (with respect to quasi-static inputs) for several modes of control. It allows characterizing notes that can be easily played, and to exhibit some limitations of the machine as well as some dyssymmetry between the lips.

This paper is organized as follows. Section 2 gives an overview of the robotic platform, its sensors and actuators. Section 3 is devoted to the calibration of some sensors and to the configuration of the machine. This includes initialization processes and feedback-loop controller settings. In section 4, an experimental protocol is proposed and the repeatability is checked. Then, the first experimental results obtained for a single control variable are presented in section 5. They allow the establishment of a partial but robust mapping based on a sound descriptor analysis. Section 6 extends these results to 2D cartographies for various control modes with multiple inputs. In particular, the playing frequencies of the set of all cartographies are examined and compared to the impedance peaks of the trombone. This analysis suggests some possible improvements of the machine that are discussed in section 7 with conclusions and perspectives.

2. TECHNICAL DESCRIPTION

The robotic platform is composed of three principal parts : (1) the air supply, (2) the artificial mouth with two lips and (3) the brass instrument with artificial fingers. It includes a set of height actuators, fourteen sensors, interfaced with a DSP setup and a computer to control the machine.

2.1 Mechanical parts

The mouth (M) is a $\simeq 80\text{cm}^3$ chamber, which is fed by a controlled air supply. It is ended by two vertical artificial lips (L_1 , L_2). Each lip is a cylindrical latex chamber filled with water. The brass instrument (in this work, a valve trombone) with its mouthpiece (MP) is fixed close to the mouth. The contact of the lips with the mouthpiece is ensured by controlling the position of the (mobile) mouth. Three artificial fingers can be used to push on the trombone valves. These components and their coupling are represented in figure 1.

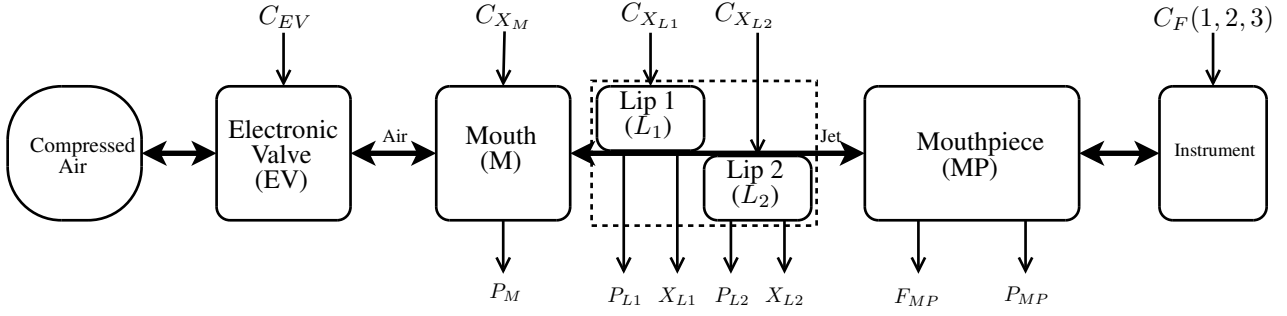


Figure 1. Block diagram of the robotized platform.

2.2 Actuators

The input airflow is controlled by an electronic valve (EV), here, a *Biç 1/2 rkert* product (Type 6022). The mouth displacement is driven by a SMAC linear actuator (LAL95-050-75F/LAA-5). The water volume inside each lip is provided by a hydraulic cylinder, also driven by a SMAC linear actuator (LAL35-015-75/LAA-5). These linear actuators are all moving coils that deliver a (Laplace) force which is proportional to the input voltage. Artificial fingers are built with simple (On/Off) electromagnets. Additionally, a horn loudspeaker can be plugged at the top of the mouth, for active acoustic control issues (see [15]). The results presented below do not involve this device but its use is considered in perspectives.

Note that the inputs of these actuators are all represented and labeled at the top of figure 1.

2.3 Sensors

The position of each linear actuator (X_M for the mouth, X_{L1} and X_{L2} for the lips, see figure 2) is measured by a built-in incremental encoder with a step of $5 \times 10^{-6}m$. The (static and acoustic) air pressure in the mouth (P_M) is measured by an Endevco sensor (8507-5). A second similar sensor (8507-2) is used for the mouthpiece (P_{MP}). The (static) water pressure (P_{L1} , P_{L2}) is measured at the top of each lip (same altitude) by two Kistler sensors (RAG-25R0.5BV1H). A SMD force sensor (S215) is localized between the mouthpiece and the instrument to measure the force (F_{MP}) applied by the lips to the mouthpiece.

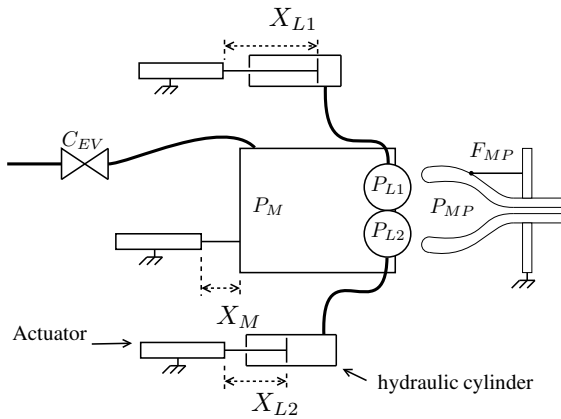


Figure 2. Sketch of the artificial mouth with its main actuators and sensors.

These sensors include built-in electronic signal conditioners except for P_M , P_{MP} and F_{MP} , which are conditioned with a low power instrumentation amplifier (INA 118). Moreover, we performed *home-made* calibrations of these sensors, except for the Endevco devices (P_M , P_{MP}) which received a factory calibration certificate. These calibrations are described in § 3.1.

Note that the outputs of the height sensors described above are all represented and labeled at the bottom of figure 1. The additional six sensors mentioned at the beginning of section 2 are: one high pressure sensor localized upstream of the electro-valve, three temperature sensors, one optical intensity sensor for estimating the opening area between the lips, and one microphone localized at the bell of the instrument. These additional sensors are not directly exploited in this paper.

2.4 Interface

The transducers associated with audio frequency ranges are connected to a sound card (see figure 3). Other transducers (with a lower frequency range) are connected to a dSpace[©] system, composed of an input/output interface and a Digital Signal Processor (DSP). The DSP is programmed using a dSpace software associated with a Matlab-Simulink-RTW[©] environment. It is used to design some (low-level) feedback loop controllers applied to the actuators. Real-time analysis of audio signals (fundamental frequencies, energy, etc) are performed by the MAX/MSP software. High-level controls of the robot (initializations, automated experiments, etc) are performed by Python scripts under the dSpace ControlDesk environment which communicates with MAX/MSP[©].

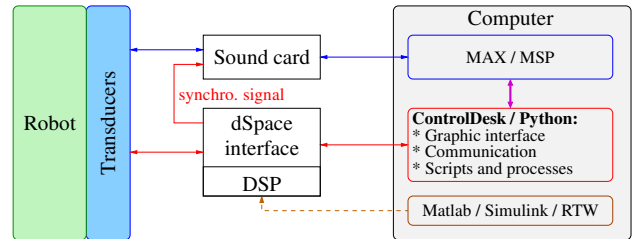


Figure 3. Overview of the interface.

3. CALIBRATION AND CONFIGURATION

The section deals with: (1) the calibration of the force sensor (F_{MP}) and the water pressure sensors (P_{L1} , P_{L2}), (2) the initialization of the *zero positions* of the linear actuators (X_{L1} , X_{L2} and X_M) so that this reference corresponds to a *robust reference state* of the lips, (3) the development and the adjustment of feedback loops to control linear actuators with respect to chosen command variables (possibly different from the natural one, the Laplace force).

3.1 Calibrations

The force and the water pressure sensors are low-frequency range sensors. We perform their calibration for quasi-static configurations. The force sensor is calibrated using gravity and reference masses (precision: $\pm 0.1g$) based on an incremental mass step of $50g$. For the water pressure sensors, a one-meter vertical water column is used, at the bottom of which the two sensors are simultaneously connected (same altitude, see figure 4). An incremental height step of five centimeters is used, corresponding to a $5 \times 10^2 Pa$ sampling precision for the pressure.

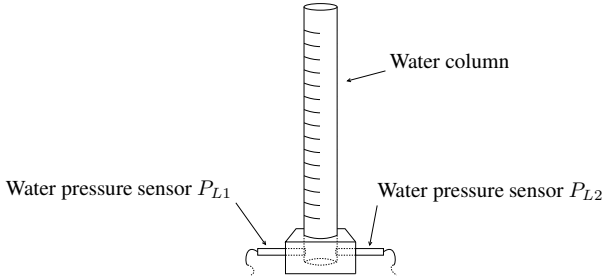


Figure 4. Schema of the water pressure sensors calibration. Static calibration based on a variable water column

The linearity of the three sensors was confirmed and the sensitivities measured.

3.2 Initialization processes

3.2.1 Zero positions for X_{L1} and X_{L2}

A lip (L) is a cylindrical latex chamber with natural volume V_L^{ref} . When the water volume V_L is smaller than V_L^{ref} , the latex is not stressed. In the opposite situation ($V_L > V_L^{ref}$), the latex stress makes the water pressure increase. This effect is measured by making the position X_L slowly increase from the forward to the backward trip points of the actuator². We observe on the measure (see figure 5) that, except in the vicinity of V_L^{ref} , these two behaviors can be approximated by one constant pressure for $V_L < V_L^{ref}$ and an affine function with slope σ (Pa/m) for $V_L > V_L^{ref}$. The zero position of $X_L = X_L^0 = 0$ is chosen and adjusted to correspond to the intersection point of the two straight lines.

Note that $X_L = 0$ does not match with V_L^{ref} . But, it splits the curve into two affine asymptotic behaviors and defines a more robust initialization (see table 1 in § 4.2 for

² The volume variation equals the position variation multiplied by the section of the hydraulic cylinder.

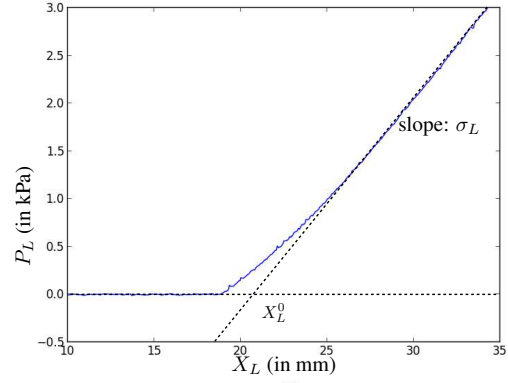


Figure 5. Measured water pressure P_{L1} with respect to the position X_{L1} when the second lip L2 is deflated and the mouthpiece is out of contact (Mouth in retracted position) and estimation of its piecewise linear approximation.

ten repetition tests). This initialization is performed for each lip, independently, and with no airflow.

3.2.2 Zero position for X_M

Once the zero positions of the lips are estimated, that of the mouth is adjusted using a similar principle. In this case, the measured quantity is the force (F_{MP}) (rather than the water pressure) while the lips are filled with their reference volume ($X_{L1} = X_{L2} = 0$). The measurements are displayed in figure 6.

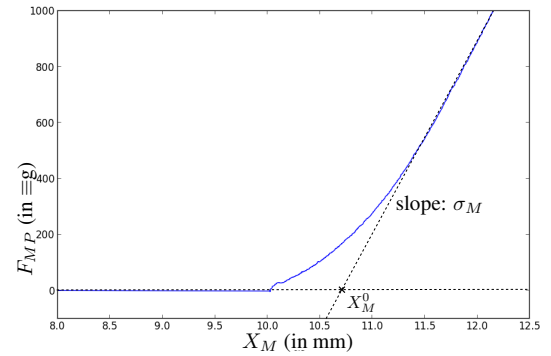


Figure 6. Measured force F_{PM} with respect to the position X_M when the lips are filled with their reference volume and estimation of its piecewise linear approximation.

Note that $X_M = 0$ does not match with the contact point between the lips and the mouthpiece, but this initialization is chosen for its robustness properties, as for $X_L = 0$.

For experiments, typical positions correspond to slightly crushed ($X_M > 0$) non over-inflated ($X_L < 0$) lips, such that the contact is established ($F_{MP} > 0$). In practice, this facilitates the formation of a buzz, or at least, that of an airflow path between the lips even for low static pressure in the mouth.

3.3 Feedback loop controllers and modes of control

The linear actuators are not naturally controlled with respect to the position but (proportionally to) the Laplace force. To achieve a control in position, we use standard tools of automatic control, here, some Proportional-Integral-Derivative (PID) controllers. Digital versions of these controllers are designed under the Matlab/Simulink/RTW environment and tuned following classical methods (see e.g. [16]). They are implemented in the DSP card of the dSpace system. For the control in position, typical performances are about 50ms. Other control types (in F_{MP} , P_{L1} or P_{L2}) are available and also based on PID controllers.

In this paper, several *modes of control* of the lips are considered. They consists of choosing, for each linear actuator, a control of *position type* or of *force/pressure type*. But, not all the combinations are compatible: positions are independent variables but, because of the contact between the lips and the mouthpiece, the force and the water pressures are linked. Here, we consider *modes of control* which include no more than one input of *force/pressure type*:

(X_M, X_{L1}, X_{L2}) this mode controls variables which do not depend of the system state;

(F_{MP}, X_{L1}, X_{L2}) this mode is well-adapted to control the contact quality between the lips and the mouthpiece;

(X_M, P_{L1}, X_{L2}) this mode (or its symmetrical version) allows the study of the effect due to the stress of one particular lip.

The feasibility of these three modes is confirmed and the repeatability is tested in section 4. Moreover, in sections 5 and 6, the two first modes are used, combined with the air-flow supply, to explore self-oscillations. They allow the automatic generation of cartographies of sound descriptors, following the curves or the surfaces detailed in figure 7.

4. EXPERIMENTAL PROTOCOL AND REPEATABILITY

In this section, we propose: (1) a protocol adapted to the quasi-static experiments presented in section 5 and section 6, (2) a repeatability test.

4.1 Protocol

An experiment is processed choosing a mode of control, an exploratory subspace, and following a precise protocol. The subspace is explored with quasi-static commands. To ensure quasi-static states, waiting times are added between measurements. For each measured point, every data from all sensors (Temperatures, Pressures, Positions, etc) are recorded and saved. Moreover, acoustic signals are automatically analyzed using tools provided by the MIR toolbox [17, 18]. For all measured points, sound descriptors such as fundamental frequency (if any), sound energy and roughness are estimated and saved. The protocol consists of the following steps:

1. Measurement of the idle-state of the machine

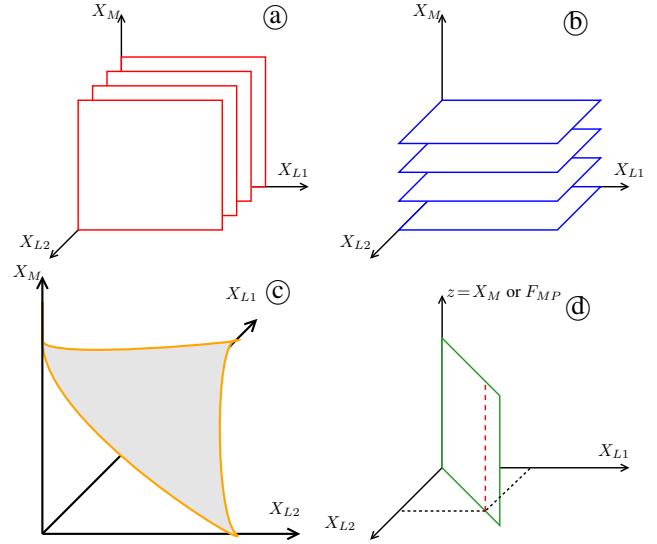


Figure 7. *Modes of control* of the three linear actuators combined with *partitions* of 3D spaces Subfigures (a)-(b) illustrate two types of partition of the 3D space into 2D planar subspaces, for the *mode of control* (X_M, X_{L1}, X_{L2}) . Subfigure (c) illustrates a F_{MP} -constant surface. It is obtained with the mode of control (F_{MP}, X_{L1}, X_{L2}) where the first input is a fixed value. Subfigure (d) describes a symmetrical control of the lips ($X_{L1} = X_{L2}$) for a fixed value X_{L1} (1D-space: red dashed straight line) or on a range (2D-space: green plane). The very first exploration in § 5 is obtained by using the 1D control subspace in (d) with $z = F_{MP}$, and the 2D cartographies in § 6 by using the 2D spaces described in (a)-(c) and (d) with $z = X_M$.

2. First initialization process: measurement of X_M^0 , X_{L1}^0 , X_{L2}^0 , σ_M , σ_{L1} and σ_{L2} .
3. For each desired point of measurement in the subspace:
 - (a) The chosen position (or force/pressure) command are set up;
 - (b) First waiting: a 0.5s waiting is imposed to ensure actuators positions;
 - (c) Breath activation: C_{EV} goes from 0 (close position) to a reference (here, 35% of the maximal aperture);
 - (d) Second waiting: a 1s waiting is imposed to ensure a quasi-stationary regime;
 - (e) Measurements are recorded and saved.
4. Second initialization process.
5. Measurement of the static pressures and the temperatures of the machine.

Parameters that are estimated during steps 2 and 4 are saved. They are compared to validate the constancy of the lips behavior. More precisely, the deviations of σ_M , σ_{L1} and σ_{L2} characterize the latex fatigue due to an experiment. The deviations of X_M^0 , X_{L1}^0 and X_{L2}^0 allow the detection of anomalies such as water leaks.

Observed variable Θ	Operating Range (OR)	$Error(\Theta) * 100/OR$ (in %) by mode of control		
		(X_M, X_{L1}, X_{L2})	(F_{MP}, X_{L1}, X_{L2})	(X_M, P_{L1}, X_{L2})
X_M	$[0, 5 \times 10^{-3}]$ (in m)	0.013%	0.26%	0.020%
X_{L1}	$[0, 3 \times 10^{-2}]$ (in m)	$5.48 \times 10^{-3}\%$	$6.31 \times 10^{-3}\%$	1.8%
X_{L2}	$[0, 3 \times 10^{-2}]$ (in m)	$6.12 \times 10^{-3}\%$	$6.13 \times 10^{-3}\%$	$6.20 \times 10^{-3}\%$
F_M	$[0, 1500]$ (in g)	1.18%	0.57%	1.34%
P_{L1}	$[0, 15]$ (in kPa)	1.56%	0.89%	0.90%
P_{L2}	$[0, 15]$ (in kPa)	1.57%	0.96%	1.16%
P_M	$[0, 15]$ (in kPa)	3.5%	1.45%	2.93%
Average Error		1.21%	0.59%	1.17%

Table 1. Errors relative to the operating range (OR), measured with the repeatability tests ($N_i = 10$, $N_k = 4 \times 4 \times 4 = 64$) for three modes of control. Values in bold correspond to the controlled variables.

4.2 Repeatability tests

To evaluate the repeatability of an experiment based on the protocol seen above, we perform a test for each mode of control presented in section 3. N_i identical experiments are processed. Experiments explore a set of N_k points that are fixed and distributed in the 3D-space. Between each experiment, the initialization parameters are manually disturbed: some water is added or removed in the “water circuit” and the instrument is slightly displaced (by hand).

Measurements are compared and analyzed, based on the following definitions:

1. A variable Θ measured at point k during experiment i is denoted Θ_k^i where Θ can be a *measurement* on a sensor with a low frequency range (F_{MP} , X_M , etc), or a *parameter estimated on a stationary signal* (typically, a *sound descriptor*, see below).
2. For a given point k , the expected value $E(\Theta_k)$ is estimated by its average on the N_i experiments, namely,

$$E(\Theta_k) \approx \frac{1}{N_i} \sum_{i=1}^{N_i} \Theta_k^i.$$

3. The standard deviation is correspondingly estimated by

$$s(\Theta_k) = \sqrt{E(\Theta_k^2) - E(\Theta_k)^2}.$$

4. The standard deviation of a variable Θ , averaged on the N_k experimental points, is denoted

$$Error(\Theta) = \frac{1}{N_k} \sum_{k=1}^{N_k} s(\Theta_k).$$

Results are presented in table 1 for the measurements of positions, water pressures and the force F_{MP} . In this table, to make the comparison between variables easier, the error is normalized with the (amplitude of the) operating range. The results show that the protocol and the initialization process are accurate enough to guarantee and quantify the repeatability for quasi-static experiments. Note that the mode of control based on the force (F_{MP}, X_{L1}, X_{L2}) appears to be, globally, the most efficient: it reduces the average error. Moreover, for this mode of control, the last step of the initialization process is not required.

5. FIRST EXPERIMENT WITH A SINGLE CONTROL VARIABLE

5.1 Global considerations

At low order, the musician’s lips can be approximated by mass-damper-spring mechanical systems [4]. Here, these 3×2 macro-parameters are related to the 3 control inputs (see figure 7), making them linked together. For instance, for a fixed position X_M (figure 7, mode ⑥), the variation of the (oscillating part of the) mass M_k of the lip Lk nearly equals that of the water displaced by X_{Lk} . The variations of stiffness K_k and damping D_k are more complex to model. Their measurement is not straightforward, even at a static equilibrium. This is why in e.g. [3, 4], these parameters are (at least partially) estimated, by analyzing the natural frequency of buzzes. In this paper, we simply explore the variability of regimes by modifying the control inputs: several frequencies can actually be reached because the sensitivities of parameters (M_k, D_k, K_k) w.r.t the control inputs are not proportional. The experiment described below makes a first exploration of regimes for a simple 1D control.

5.2 Description of the experiment

The 1D control is chosen so that the lips are in a symmetrical configuration. In order to makes this configuration as robust as possible, the independent variables $X_{L1} = X_{L2}$ are kept constant and fixed to $-15mm$, and the mouth control is chosen to be F_{MP} which increases from $100g$ to $1000g$, following the protocol described in § 4.1.

5.3 Measurements and observations

Figure 8 shows the measured force F_{MP} and signal descriptors [17, 18] of the acoustic pressure measured in the mouthpiece P_{MP} . The three sub-figures respectively display: (1) the *force*, (2) the *fundamental frequency* f_0 of P_{MP} , estimated by the YIN algorithm [18], and (3) the *energy* of the signal. An additional curve representing the *roughness* is superposed to the estimated fundamental frequency. The roughness is defined without unit: a high value means that the signal is not harmonic.

This figure validates that several fundamental frequencies are reachable. The analysis of sound descriptors clearly allows the extraction of connected areas of self-oscillating

regimes for which the roughness is low and the energy is high. These areas correspond to some “*stable notes*”. The complementary areas basically correspond to non-oscillating or complex signals (multiphonics, chaos, etc). Complex signals are mainly located in thin transition areas between stable notes.

5.4 First conclusion

This experiment shows that the robotic platform is able to produce various self-oscillating regimes including *stable notes*. For *stable notes*, it appears that, to a large extent, the higher the force F_{MP} , the higher the fundamental frequency. Moreover, stable notes and their areas are sufficiently reproducible to map some control input values to fundamental frequencies. Next section extends this exploration to the case of two control inputs.

6. TWO-DIMENSIONAL CARTOGRAPHIES

In this section, four experiments are set up, based on the modes of control ③ to ⑥ in figure 7.

6.1 Description of the experiments

Experiments are conducted following the protocol in § 4.1. For each experiment, the control inputs are specified below with their exploration range and the incremental step. Variables are sorted as follows: the first, second and third ones respectively correspond to (1) a *constant* value (except in case ⑥), (2) a slow increasing sweep, and (3) a faster (but still quasi-static) increasing sweep. Parameters are:

Mode ③: $X_{L2} = -15mm$; $1mm \leq X_M \leq 4mm$ (step: $0.1mm$); $-30mm \leq X_{L1} \leq 0mm$ (step: $0.5mm$);

Mode ④: $X_M = 3.5mm$; $-30mm \leq X_{L2} \leq 0mm$ (step: $0.5mm$); $-30mm \leq X_{L1} \leq 0mm$ (step: $0.5mm$);

Mode ⑤: $F_{MP} \equiv 500g$; $-30mm \leq X_{L2} \leq 0mm$ (step: $0.5mm$); $-30mm \leq X_{L1} \leq 0mm$ (step: $0.5mm$);

Mode ⑥: $1mm \leq X_M \leq 4mm$ (step: $0.1mm$); $X_{L1} = X_{L2}$ with $-30mm \leq X_{L1} \leq 0mm$ (step: $0.5mm$).

6.2 Measurements: comments and observations

The results of these four experiments are given in figure 10. For each mode of control, ③-⑥, connected areas associated with “*stable notes*” can still easily be extracted. However, these 2D cartographies makes some complexity appear: they include non convex areas and a large variety of sizes, specially for modes ③ and ④. The experiment for mode ⑥ is an extension of that shown in section 5. We can observe wide diagonal bands of *stable note* areas so that mode ⑥ makes the note selection simpler than other modes. Indeed, a preferred axis can be selected (such as X_L proportional to X_M) in order to explore areas of stable notes with increasing frequencies.

Sub-figure ③ displays a few number of huge or confined areas, meaning that the force is well adapted to stabilize some specific notes (here, denoted N_1 and N_2).

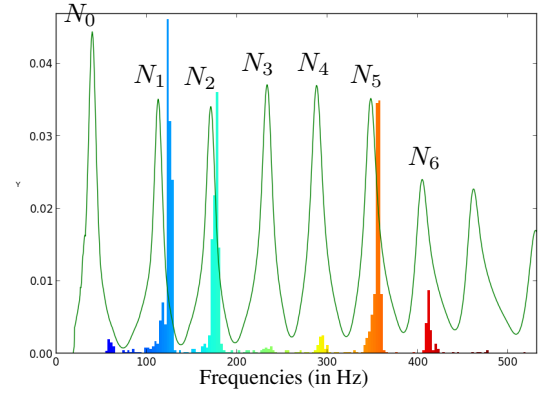


Figure 9. Normalized histogram of fundamental frequencies superposed to the modulus of the input impedance measured with the BIAS system [20].

However, note that, the symmetry with respect to lips 1 and 2 is not satisfied, although initializations are correct (and their reproducibility is checked). This can also be observed in sub-figure ⑥. This issue will require a special attention in a future work.

Moreover, the histogram in figure 9 shows that the seven first notes of the instrument (N_0 to N_6) can be played. We can notice that playable frequencies perfectly fit with the trombone impedance (green solid line curve) but are located at the right of each peak (as expected [19]).

7. CONCLUSIONS AND PERSPECTIVES

In this paper, a robotized artificial mouth dedicated to playing the brass instruments is described with its mechatronic parts, its calibration, some robust initialization processes and several modes of control. Repeatability tests (ten acquisitions for each mode of control) is performed to estimate the standard deviation on the measured variables which characterize the lips in quasi-static configurations. Globally, the relative deviation proves to be less than 0.9% for the controlled variables and less than 3.5% for the uncontrolled ones. A protocol for experiments is proposed. In particular, it records the temperature during a (possibly long) experiment and characterizes the latex fatigue. Experiments on a trombone are designed to build cartographies for several modes of (quasi-static) control. These cartographies display sound descriptors (fundamental frequency, roughness, energy) as a function of the control input values. This exploration reveals that several stable notes can easily be reached using a basic mapping but that notes in the high range of the instrument cannot be played by the artificial mouth.

To cope with this limitation, a perspective is to modify the machine in order to make it able to play notes in the high register. A solution could consist of controlling the acoustic impedance of the mouth. A partial but encouraging result based on an acoustic active control can be found in [15]. In this work, the RMS pressure amplitude is analyzed with respect to the (controlled) phase difference between the pressure in the mouth and in the mouth-

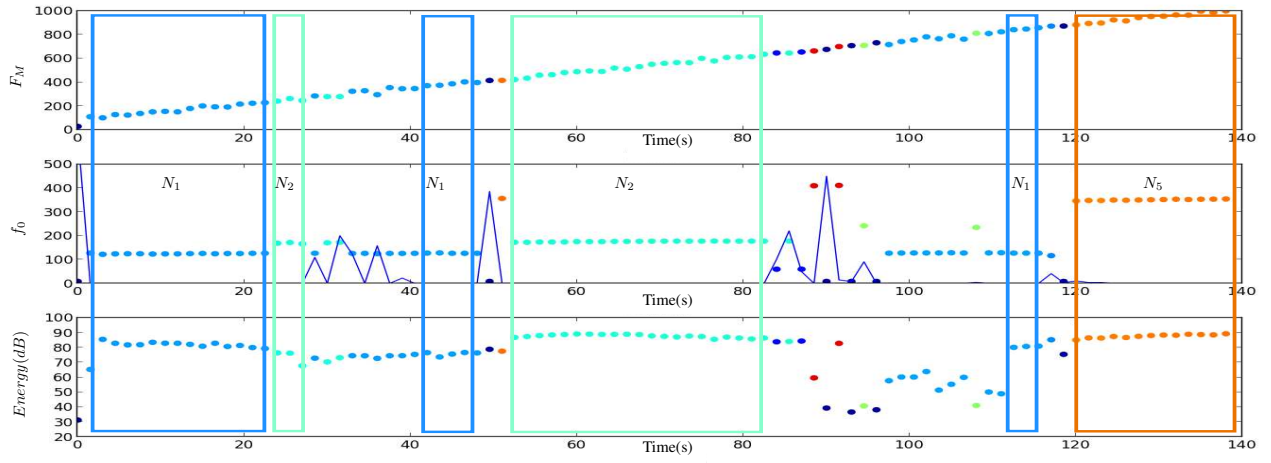


Figure 8. (First experiment: $X_{L1} = X_{L2} = -15mm$ and F_{PM} increasing from 100g to 1000g) Top: F_M w.r.t. time. Middle: fundamental frequency estimated on the acoustic signal P_{MP} (\cdot) and roughness ($-$). Bottom: signal energy. In this figure, the color map corresponds to the frequencies and areas with stable notes are emphasized with boxes.

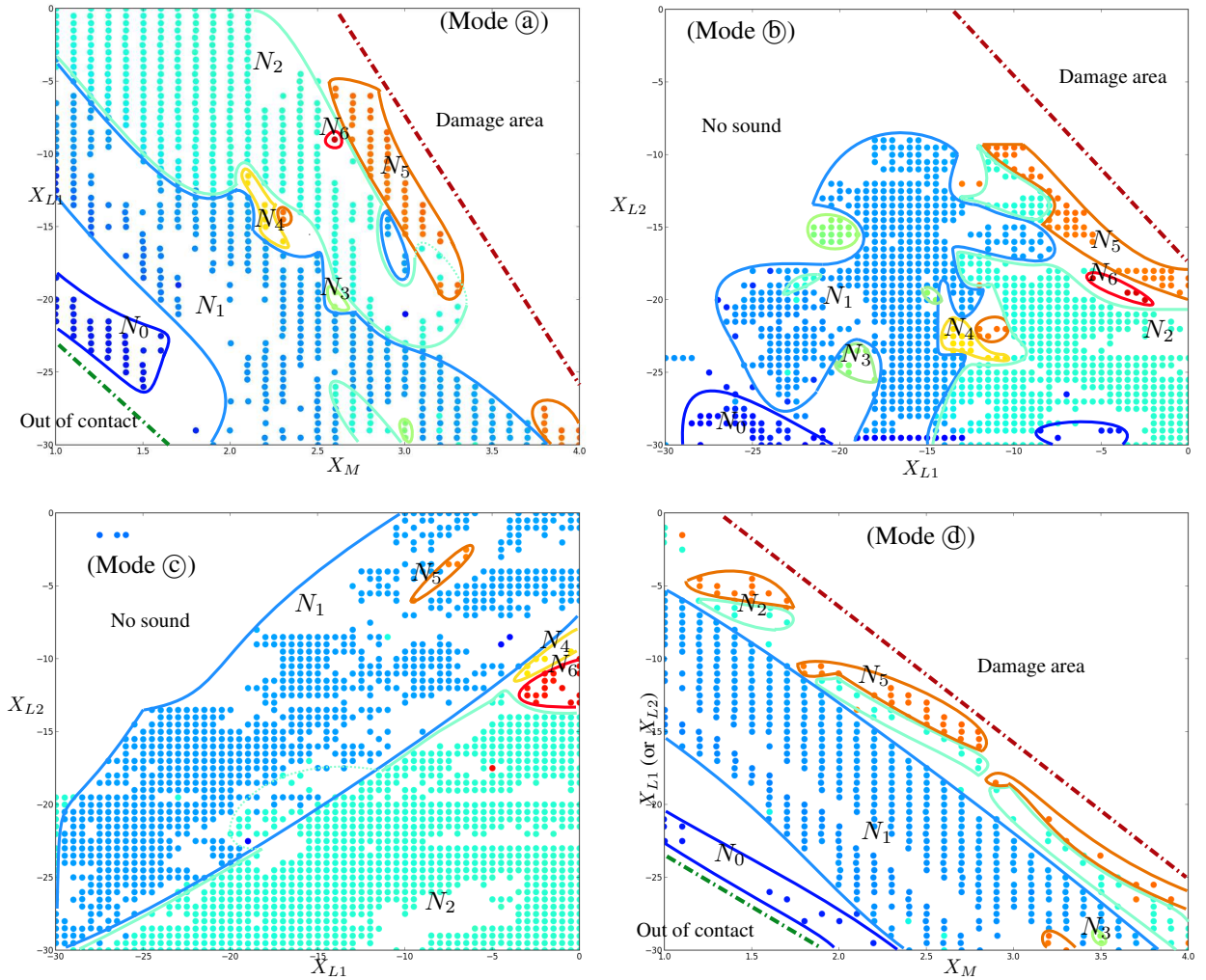


Figure 10. 2D-cartographies for various control modes. Top left: mode (X_M, X_{L1}, X_{L2}) with constant X_{L2} . Top right: mode (X_M, X_{L1}, X_{L2}) with constant X_M . Bottom left: mode (F_{MP}, X_{L1}, X_{L2}) with constant F_{MP} . Bottom right: mode $(X_M, X_{L1} = X_{L2})$. The color map corresponds to the frequencies (see figure 9) and stable note areas are surrounded. The red dotted line represents the security limit ($P_{L1} > 15kPa$ or $P_{L2} > 15kPa$) and the green dotted line represents the contact between the lips and the mouthpiece.

piece. A second perspective concerns the study of some macro-mechanical parameters of the lips (effective mass M , damping D and stiffness K) with respect to the control inputs. However, note that the (3D) control inputs do not allow the setting of the (3×2) macro-parameters. A first approach is to hold up one lip (its 1D control is then used to ensure the air-tightness). The 2D remaining control inputs can then be used to tune two of the three macro-parameters of the other (vibrating) lip, typically, M and K . A third perspective is to implement some high-level feedback loop controllers driven by the fundamental frequency and acoustic the energy.

More generally, a complete dynamical system modeling of the machine is under study. The long-term goals of this research are to: (1) derive state observers, (2) design efficient high-level controllers, and (3) build inversion processes to make the machine play target sound waves.

Acknowledgments

Authors wish to thank the French National Research Agency and the project CAGIMA for supporting this work and Alain Terrier and G   ard Bertrand for their technical assistance.

8. REFERENCES

- [1] C. Vergez and X. Rodet, "Comparison of real trumpet playing, latex model of lips and computer model," in *ICMC: International Computer Music Conference*, Thessaloniki Hellas, Greece, Septembre 1997, pp. 180–187.
- [2] —, "Experiments with an artificial mouth for trumpet," in *ISMA: International Symposium of Music Acoustics*, Leavenworth, Washington state USA, Juin 1998, pp. 153–158.
- [3] J. Gilbert, S. Ponthus, and J.-F. Petiot, "Artificial buzzing lips and brass instruments: experimental results," *J. Acoust. Soc. Am.*, vol. 104, no. 3, pp. 1627–1632, 1998.
- [4] J. Cullen, J. Gilbert, and D. M. Campbell, "Brass instruments : linear stability analysis and experiments with an artificial mouth," *Acta Acustica united with Acustica*, vol. 86, no. 3, pp. 704–724, 2000.
- [5] J.-F. Petiot, F. Teissier, J. Gilbert, and M. Campbell, "Comparative analysis of brass wind instruments with an artificial mouth: First results," *Acta Acustica united with Acustica*, vol. 89, no. 6, pp. 974–979, 2003.
- [6] J. Kergomard, "Projet consonnes: Contr  le des sons naturels et synth  tiques," Agence Nationale de la Recherche (ANR-05-BLAN-0097-01), 2005-2009, <http://www.consonnes.cnrs-mrs.fr/>.
- [7] "Mechatronics projects of the engineering school of mines paristech," <http://www.mecatrou.fr/>.
- [8] B. V  ricel, "Commande et interfa  age d'un robot musicien," 2009.
- [9] —, "Confrontation th  orique/exp  rimentale de caract  ristiques d'excitation dans le jeu des cuivres," 2010. [Online]. Available: <http://articles.ircam.fr/textes/Vericel10a/>
- [10] N. Lopes, "Cartographie de param  tres de jeu de trompettiste: mise en correspondance automatique du son produit avec les param  tres de contr  le d'une bouche artificielle asservie," 2011. [Online]. Available: <http://articles.ircam.fr/textes/Lopes11a/>
- [11] —, "Mod  lisation, asservissement et commande d'une bouche artificielle robotis  e pour le jeu des cuivres," 2012. [Online]. Available: <http://articles.ircam.fr/textes/Lopes12a/>
- [12] D. Ferrand, T. H  lie, C. Vergez, B. V  ricel, and R. Causs  , "Bouches artificielles asservies:   tude de nouveaux outils pour l'analyse du fonctionnement des instruments    vent," in *Congr  s Fran  ais d'Acoustique*, vol. 10, Lyon, France, Avril 2010. [Online]. Available: <http://articles.ircam.fr/textes/Ferrand10a/>
- [13] T. H  lie, N. Lopes, and R. Causs  , "Robotized artificial mouth for brass instruments: automated experiments and cartography of playing parameters," in *PE-VOC - Pan European Voice Conference*, vol. 9, Marseille, France, 2011, pp. 77–78.
- [14] —, "Open-loop control of a robotized artificial mouth for brass instruments," in *Acoustics 2012 (ASA)*, Hong Kong, China, 2012, pp. 1–1.
- [15] V. Fr  our, N. Lopes, T. H  lie, R. Causs  , and G. Scavone, "Simulating different upstream coupling conditions on an artificial trombone player system using an active sound control approach," in *International Conference on Acoustics*, vol. 21, 2013, p. 5p.
- [16] K. Astr  m and T. H  gglund, *PID Controllers: Theory, Design, and Tuning*, 2nd ed. International Society for Measurement and Control, 1995.
- [17] O. Lartillot and P. Toivianen, "A Matlab Toolbox for Musical Feature Extraction from Audio," in *the 10th International Conference on Digital Audio Effects (DAFx07)*, 2007, pp. 237–244.
- [18] A. de Cheveign   and H. Kawahara, "Yin, a fundamental frequency estimator for speech and music," *J. Acoust. Soc. Amer.*, vol. 111, pp. 1917–1930, 2002.
- [19] M. Campbell, "Brass instruments as we know them today," *Acta Acustica united with Acustica*, vol. 90, pp. 600–610, 2004.
- [20] G. Widholm, H. Pichler, and T. Ossmann, "Bias: A computer-aided test system for brass wind instruments," in *Audio Engineering Society Convention*, vol. 87, 1989, paper 2834.

Muscle Activity in Playing Trumpet: The Dependence on the Playable Pitch Region and the Experience of a Non-trumpet Brass Instrument Player

Shogo Matsukata

Nihon University
matsukata@kthrlab.jp

Hiroko Terasawa

University of Tsukuba / JST, PRESTO
terasawa@slis.tsukuba.ac.jp

Masaki Matsubara

University of Tsukuba
masaki@slis.tsukuba.ac.jp

Tetsuro Kitahara

Nihon University
kitahara@kthrlab.jp

ABSTRACT

In this paper, we investigated the relationship between the muscle activity around the lips and the sounds of playing the trumpet. While one is playing the trumpet, it is important but also difficult to keep an embouchure. Previous studies have investigated the difference between muscle activity for register and proficiency, but these studies have not considered the player's playable register and the influence of experience playing other instruments. Thus, in this study we aimed to solve these problems by examining the surface electromyograms (EMGs) around the lips: the orbicularis oris superioris, orbicularis oris inferioris, depressor anguli oris, and levator anguli oris. As a result, we found that (1) muscle activity is greater in the high register than the low register, but muscle activity in the high register are not significantly greater than those in the low register for the wide range of the playable register in the fixed register (FR); and (2) a novice trumpeter has greater muscle activity in the upper lip than in the lower lip, while an expert shows no difference. The muscle activity is the same as for the novice player when a non-trumpet brass player plays the trumpet.

1. INTRODUCTION

The trumpet is very popular in brass bands, jazz bands and orchestras, and it is well known as a leading instrument. In order to vibrate the lips properly in trumpet playing, trumpeters regularly maintain an embouchure, which is the use of facial muscles and the shaping of the lips to the mouthpiece of the instrument [1], [2].

However, keeping the embouchure is difficult because muscle fatigue occurs if the trumpeters try to play forcibly [3]. For that reason, trumpeters often struggle to play higher notes when performing for a long time [1]. In addition, Papsin has said that the embouchure of a brass player is in-

tegral to production of sound and manipulation of tonality and muscle [1].

Several studies have measured muscle activity during musical performance [4], [5]. In particular, White investigated the muscle activity while playing the trumpet [5]. White did a survey of 18 trumpeters and found the following characteristics of muscle activity around the lips during performance:

1. Muscle activity is greater in the high register (range of pitch) than in the low register.
2. Novice trumpeters have greater muscle activity in the upper lip than in the lower lip, while experts show no difference.

White did not consider the players' playable register. In fact, for statistical analyses the sustained tones G3 and C4 were designated as the low register and G5 and C6 as the high register. However, the playable register depends on the player, so we classified it into two registers: the register of each player's playable register (individual register) and the common register among all players (fixed register).

In addition, White investigated trumpet players but neglected the years of experience playing non-trumpet brass instruments. Other brass instruments (e.g., the horn and trombone) have similar sounding mechanisms to the trumpet's, so the basic muscle activity may be close among these instruments. In other words, we asked: what are the muscle activity when non-trumpet players play the trumpet?

Thus, the following questions are still open:

1. The difference of muscle activity between the individual register and the fixed register.
2. The difference in muscle activity between trumpet experience and other instrumental experience.

In this paper, we investigate these two questions by measuring muscle activity during playing the trumpet using surface electromyographs. Specifically, we conduct the following experiments:

1. We compare EMG signals during playing the trumpet in the high and low registers under two different

definitions of “high register.” In the first definition, the highest register playable for each subject is regarded the high register. In the second definition, the same register for all subject is used as the high register. Through comparing the results with these two definitions, we investigate the dependency of the muscle activity on the pitch region that the player can play.

2. In addition to novice and expert trumpeters, we measure the EMGs in playing the trumpet by subjects who have long experience of playing the horn but no experience with the trumpet. We then investigate which of the novices’ and experts’ muscle activity are close to those of the horn players’.

2. EXPERIMENTAL PREPARATION

2.1 Subjects

The subjects were four brass instrument players ranging in age from 19 to 23 years (see Table 1). Their years of experience playing the trumpet ranged from 0 to 7. We regarded Subject A as an expert trumpet player and Subject B as a novice. In addition, in order to investigate the muscle activity of non-trumpet players performing on the trumpet, we included two horn players (Subjects C and D).

Table 1. Subjects information

	Age	Gender	Years of Experience
Subject A	22	male	Trumpet 7
Subject B	19	female	Trumpet 0.5
Subject C	22	female	Horn 12 (Trumpet 0)
Subject D	23	male	Horn 8 (Trumpet 0)

2.2 Muscles to be measured

We measured the following muscles (Fig. 1), which are the same muscles as in White’s study [5]:

- Orbicularis Oris Superioris (OOS)
- Orbicularis Oris Inferioris (OOI)
- Depressor Anguli Oris (DAO)
- Levator Anguli Oris (LAO)

2.3 Measureing Instruments Used

We used a Wireless EMG Logger (Logical Products, Inc.) and an EMG sensor (Oisaka Electronic Equipment Ltd.) for measuring EMGs, and a C519M (AKG) for recording sounds (Fig. 2). All subjects used a YTR-4335GS (Yamaha Corp.) as a trumpet in order to reduce the instrument-dependent effects.

2.4 EMG signal processing

While the subjects play the trumpet, we simultaneously measured their muscle activity and recorded the sounds. The EMG and audio signal are then analyzed with the following steps:

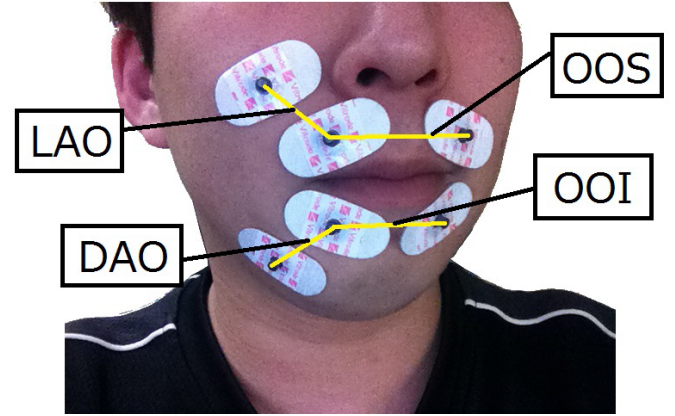


Figure 1. Muscles to be measured.

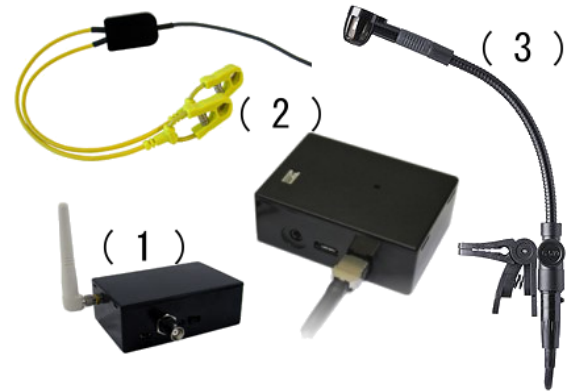


Figure 2. Instruments used in the experiments. (1) Wireless EMG Logger, (2) EMG sensor, (3) Microphone (C519M)

1. The onset and offset times of each note are calculated from the audio signal. The sets of the onset and offset times are defined as follows:

$$\text{Onset} = \{t | A(t) < 0.1A_{\max} \wedge A(t+1) > 0.1A_{\max}\},$$

$$\text{Offset} = \{t | A(t) > 0.1A_{\max} \wedge A(t+1) < 0.1A_{\max}\},$$

where $A(t)$ is the amplitude at time t of the audio signal and A_{\max} is its maximum (Fig. 3 left side).

2. From the duration (from onset to offset) of each note, the steady state is extracted. The steady state is defined as the section all frames in which have an amplitude greater then 75 % of the maximum amplitude within the duration of that note.
3. For each note, the EMG signal corresponding to the sustain section is extracted.
4. The root-mean-square (RMS) value of the extracted EMG signal for each note is calculated (Fig. 3 right side).

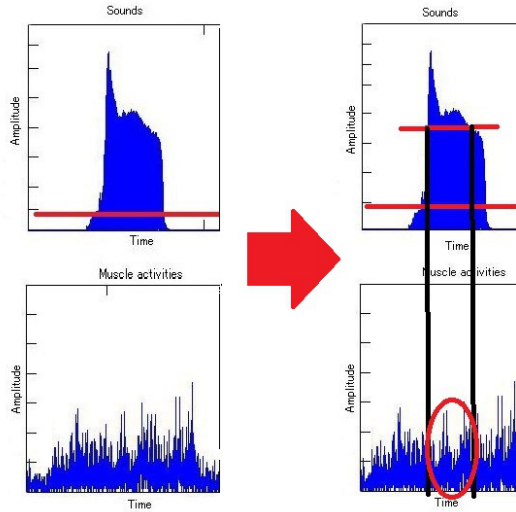


Figure 3. EMG signal processing: The steady state of the sound was extracted by setting a threshold at 10% of the maximum amplitude (left side). Then root-mean-square of the steady state was calculated (right side).

3. EXPERIMENT 1: THE RELATIONSHIP BETWEEN THE REGISTER AND MUSCLE ACTIVITY

In this section, we investigated the muscle activity focused on the playable register. While investigated the muscle activity in fixed register but did not consider the muscle activity in players' playable register. Thus, we examined the difference between the muscle activity of the players' playable register and the fixed register.

3.1 Register Settings for the Experiment

Subjects were asked to play notes one by one in the B \flat major scale. Each note has a duration of 1–2 seconds. Fig. 4 shows the pitch region that each subject could play. We defined two types of registers in order to consider the players' proficiency: individual register (IR) and fixed register (FR). The red circle shows the individual register (defined as the highest and lowest three notes that each subject could play), and the blue square shows the fixed register (defined as the highest and lowest three notes that all subjects can commonly play). The performance in the same condition by the same subject was repeated three times.

3.2 Results

Table 2 lists the results of the significant-difference test between the muscle activity of the high and low registers in IR. We used the Mann-Whitney U test and the Bonferroni correction in the statistical analysis. The marks in the table indicate the p value in which “-” is not significant, “*” is $p < 0.05$, “**” is $p < 0.01$, and “***” is $p < 0.001$. Thus, the muscle activity in the high register are observed significantly greater than those in the low register for all subjects.

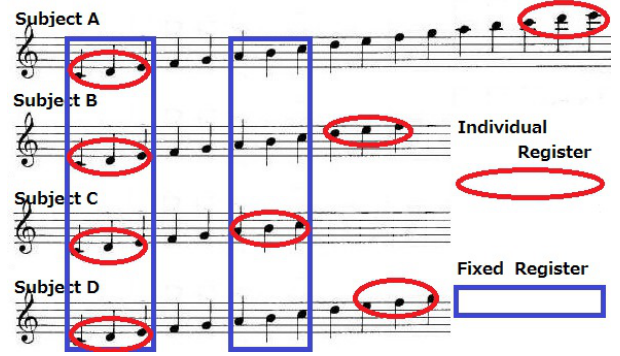


Figure 4. Individual register (red circle) and fixed register (blue square). Each staff represents the pitch region that each subject can play. From this figure, we can see a difference of the pitch regions among subjects.

Table 2. Differences of muscle activity between the highest three notes and lowest three notes in IR.

(-: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, Bonferroni adjusted)

Muscle	Subj. A	Subj. B	Subj. C	Subj. D
OOS	5.3×10^{-7} , ***	7.3×10^{-7} , ***	2.28×10^{-6} , ***	3.82×10^{-7} , ***
OOI	5.6×10^{-6} , ***	4.7×10^{-5} , ***	3.8×10^{-7} , ***	3.8×10^{-7} , ***
DAO	1.2×10^{-3} , **	8.2×10^{-5} , ***	4.1×10^{-5} , ***	4.1×10^{-5} , ***
LAO	2.8×10^{-3} , **	8.2×10^{-5} , ***	4.1×10^{-5} , ***	4.1×10^{-5} , ***

Table 3 lists the results of the significant-difference test between the muscle activity of the high and low FR. In the same way as the IR, we used the Mann-Whitney U test and the Bonferroni correction in our statistical analysis. From the table, muscle activity in the high register were significantly greater than those in the low register for Subjects C and D. However, there was no significant difference of the muscle activity from Subject A. Thus, the muscle activity in the high register do not seem to be significantly greater than those in the low register for the wide range of the playable register in FR.

Table 3. Differences of muscle activity between the highest three notes and lowest three notes in FR.

(-: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, Bonferroni adjusted)

Muscle	Subj. A	Subj. B	Subj. C	Subj. D
OOS	5.0×10^{-4} , ***	2.2×10^{-5} , ***	2.2×10^{-6} , ***	21.9×10^{-6} , ***
OOI	2.7×10^{-2} , -	2.5×10^{-3} , **	3.8×10^{-7} , ***	3.2×10^{-7} , ***
DAO	3.4×10^{-1} , -	7.7×10^{-2} , -	4.1×10^{-5} , ***	4.1×10^{-5} , ***
LAO	7.9×10^{-1} , -	4.9×10^{-4} , ***	4.1×10^{-5} , ***	4.1×10^{-5} , ***

4. EXPERIMENT 2: THE RELATIONSHIP BETWEEN PROFICIENCY AND MUSCLE ACTIVITY

We examined the muscle activity in proficiency between trumpeters and other brass-instrument players. White et al. [5] argued that novice trumpeters have greater muscle activity in the upper lip than the in the lower lips, while experts show no difference, but how the muscle activity behaves when experts of brass instruments other than the trumpet play the trumpet has remained an open question. Therefore, we investigated the muscle activity when such players play the trumpet using the same data as those for Experiment 1.

4.1 Proficiency Settings for the Experiment

As seen from Table 4, Subject A was an expert on the trumpet, while Subject B was a novice. Subjects C and D were experts on the horn but had no experience of playing the trumpet. We used the eight tones of the B \flat major scale from B \flat 3 to B \flat 4, and we measured the muscle activity of the upper lips (OOS) and lower lips (OOI) of the subjects. In the same way as Experiment 1, we used the Mann-Whitney U test and the Bonferroni correction in the statistical analysis.

4.2 Results

Table 4 shows the results of the significant-difference test between the upper lip and lower lip. It can be seen that the muscle activity between the OOS and OOI were significantly different for the novice, while those for the expert were not significantly different. In addition, those for Subjects C and D were also significantly different. Thus, in this condition, non-trumpet brass players behave in the muscle activity like novice trumpeters.

Table 4. Differences of muscle activity between the upper and lower lips (OOS vs OOI).

(-: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, Bonferroni adjusted)

Subj. A	Subj. B	Subj. C	Subj. D
2.4×10^{-1} , -	3.0×10^{-9} , ***	2.5×10^{-3} , **	6.2×10^{-8} , ***

5. DISCUSSION

The results of our experiments can be summarized as follows:

1. Muscle activity is greater in the high register than in the low register, but the muscle activity in the high register are not significantly greater than in the low register for the wide range of the playable register in FR.
2. Novice trumpeters have greater muscle activity in the upper lip than the lower lip, while experts show no such difference. In addition, horn experts with no trumpet experience show muscle activity similar to novice trumpet players.

5.1 The Relationship Between Register And Muscle Activity

Table 2 shows that muscle activity is greater in the high register than in the low register in IR. From this experiment, we obtained a result similar to the claim of White et al. [5]. In fact, Fig. 5 shows that there is a significant difference of muscle activity in IR (surrounded by red square), and there is a no significant difference of muscle activity in FR (surrounded by blue square).

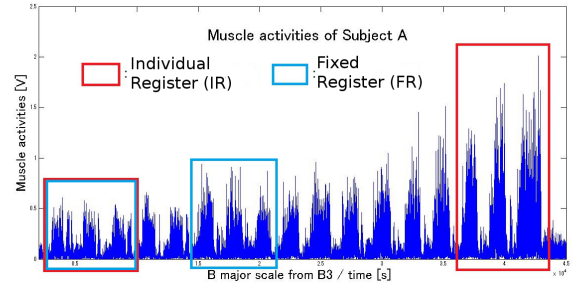


Figure 5. The muscle activity recorded OOI by Subject A. IR surrounded by red square and FR surrounded by blue square.

5.2 The Relationship Between Proficiency And Muscle Activity

Table 4 shows that experts are balanced and novices are not balanced in the upper and lower lips, which is similar to White et al.'s result [5]. Fig. 6 shows the muscle activity in the upper lip (upper figure) and the lower lip (lower figure) of an expert (Subject A).

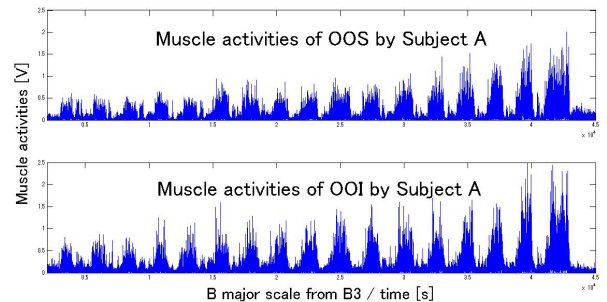


Figure 6. Muscle activity of an expert performance, measured at upper lip (OOS, above) and at lower lip below (OOI, below). Expert players have balanced the muscle activity in upper and lower lips.

Subject A has balanced muscle activity in the upper and lower lips. Fig. 7 shows the muscle activity in the upper lip (upper figure) and the lower lip (lower figure) of a novice (Subject B).

Subject B does not have balanced the muscle activity in the upper and lower lips. Subjects majoring in the horn have similar muscle activity to the novice's. In other words, we considered that the muscles are used in different ways in playing the trumpet and the horn.

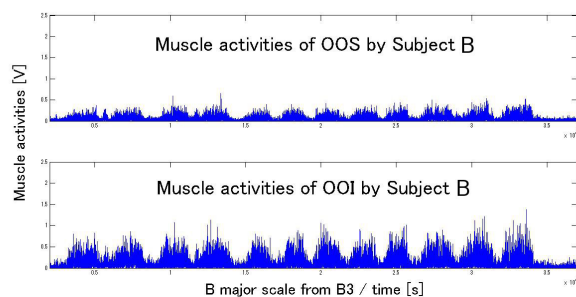


Figure 7. Muscle activity of a novice performance, measured at upper lip (OOS, above) and at lower lip below (OOI, below). Novice players have not balanced the muscle activity in upper and lower lips.

6. CONCLUSIONS AND FUTURE WORK

We focused on the fact that keeping the embouchure is difficult and influences the muscles around the lips for trumpet playing. In this study, we investigated the relationship between the muscle activity around the lips and the sounds of playing the trumpet. As a result, we found the following.

1. Muscle activity is greater in the high register than in the low register, but muscle activity in the high register are not significantly greater than in the low register for the wide range of the playable register in FR.
2. A novice trumpeter has greater muscle activity in the upper lip than in the lower lip, while experts show no difference. In addition, the muscle activity of player of a brass instrument other than the trumpet seems to be similar to that of the novice trumpet player.

As future work, we should investigate a growth process of moving from novice to expert because there is a difference in muscle activity around the lips between the novice and the expert. We will investigate whether the novice decreases the upper lip muscle activity or increases the lower lip muscle activity when acquiring advanced skill. In addition, referring to attempts to achieve real-time feedback of EMG for performance support and rehabilitation [6], [7], [8], [9], we will develop a system to support trumpet practice through real-time analysis and feedback of data on muscle activity in the lips.

7. REFERENCES

- [1] Papsin BC, Maaske LA, McGrail JS, "Orbicularis oris muscle injury in brass players," *The Laryngoscope*, vol. 6, no. 106, pp. 757–760, 1996
- [2] Iltis PW, Michael WG, "EMG Characterization of Embouchure Muscle Activity: Reliability and Application to Embouchure Dystonia," *Medical Problems of Performing Artists*, vol. 1, no. 20, pp. 25–34, 2005
- [3] Kourakata I, Moriyama K, Hara T, "Identification of Control Parameters for Brass Player's Embouchure by

Measuring Contact Pressure on the Teeth Buccal Surface," *JSME international journal. Series C, Mechanical systems, machine elements and manufacturing*, vol. 4, no. 44, pp. 1142–51, 2001

- [4] Bianco T, Freour V, Cossette I, Bevilacqua F, Causse R, "Measures of Facial Muscle Activation, Intra-oral Pressure and Mouthpiece Force in Trumpet Playing," *Journal of New Music Research*, vol. 1, no. 41, pp. 49–65, 2012
- [5] White RE, Basmajian VJ, "Electromyographic Analysis of Embouchure Muscle Function in Trumpet Playing," *Journal of Research in Music Education Winter*, vol. 4, no. 22, pp. 292–304, 1974
- [6] Morasky RL, Reynolds C, Sowell LE, "Generalization of lowered EMG levels during musical performance following biofeedback training," *Biofeedback Self Regul.*, vol. 2, no. 8, pp. 207–216, 1983
- [7] Tsubouchi Y, Suzuki K, "BioTones: a wearable device for EMG auditory biofeedback," *Conf Proc IEEE Eng Med Biol Soc. 2010*, pp. 6543–46, 2010
- [8] Matsubara M, Terasawa H, Kadone H, Suzuki K, Makino S, "Sonification of muscular activity in human movements using the temporal patterns in EMG," *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*, vol. 6, no. 3, pp. 1–5, 2012
- [9] Andy H, Sandra P, "Use of Sound for Physiotherapy Analysis and Feedback," in *The Sonification Handbook*, Hermann T, Hunt A, Neuhoff J, Editors., Logos, Verlag, Berlin, pp. 528–531, 2011.

Pitch bending technique on early horns by manipulation of the embouchure: A comparison between measured and predicted data.

Lisa Norman

University of Edinburgh
Acoustics and Audio Group
Lisa.Norman@ed.ac.uk

Jonathan Kemp

University of St Andrews
Department of Music
jk50@st-andrews.ac.uk

John Chick

University of Edinburgh
Acoustics and Audio Group
John.Chick@ed.ac.uk

Murray Campbell

University of Edinburgh
Acoustics and Audio Group
d.m.campbell@ed.ac.uk

ABSTRACT

Brass players sometimes adopt a technique whereby they adjust their embouchure in order to alter or bend the pitch of a note away from the centre of the resonance. The ease and control with which this can be achieved is an important factor in assessing the playability of a brass instrument. A good instrument will have well defined resonances, but experienced players do not like instruments with notes that are too ‘stiff’, and which lack sufficient flexibility for musical expression. The need for the ability to bend the pitch of a note is particularly important for natural trumpets and horns used in the baroque era, when instruments did not have valves and players were required to bend the pitch of some resonances (e.g. the 11th and 13th) by a significant fraction of a semitone. The instrument and player form a complex and closely coupled system. Using experimental data from playing tests on early orchestral horns, and comparing these results with those from a recently developed time domain model, it is possible to begin to identify features of an instrument and its interaction with a player which make it more or less susceptible to this type of manipulation.

1. INTRODUCTION

Arguably the most fundamental aspect of the acoustic design of a brass instrument is the creation of a bore profile whose acoustic resonances are harmonically related. The harmonicity of this series is important as it allows for a strong coupling between the resonances of the air column within the instrument, also known as a cooperative regime of oscillation [1]. This interaction helps to support the oscillation of the lip, a key function of the player’s embouchure.

Although the internal shape of the instrument determines the frequencies at which the lips will most readily vibrate, the player still has the ability to exert some influence on the finer tuning of these resonances by ‘bending’ the pitch of notes using embouchure manipulation.

Pitch bending is commonly referred to by players as ‘lip-ping’, but the use of this term is perhaps slightly misleading. The technique involves the manipulation of the whole

of the mouth and embouchure to change the pitch of the note; the lips, the tongue, the oral cavity, and mouthpiece pressure may all be brought into play when adopting this technique, and different musicians achieve this effect in different ways.

1.1 Historical context

Horns and trumpets from the seventeenth and eighteenth centuries do not have valves (the invention of which extended the range of notes available to the player), and so composers at this time were restricted to using predominantly pitches based on the natural resonances of the instrument.

Most of these pitches correspond to the notes of a diatonic scale whose tonic is the nominal pitch of the instrument, but the seventh, eleventh and thirteenth do not fit comfortably within this tuning system. Wayward pitches such as these were most likely problematic for musicians of the day, and they remain an issue for modern brass players interested in performing on natural instruments. In order to fine tune these pitches it is likely that a pitch bending technique was adopted.

Players of modern orchestral instruments are also interested in pitch bending, but at the more subtle level of musical inflection. Although significant for players in evaluating instruments, this usage is modest by comparison with baroque horn and trumpet playing. The focus of this paper is on the higher levels of pitch bending required for baroque horn playing.

1.2 The 11th resonance

Baroque composers do not seem to have made a particular effort to avoid these problematic pitches in their horn and trumpet writing. The written F^{\natural} and F^{\sharp} , see Figure 1, are two notes frequently used by composers around this time, and yet both notes share the same natural resonance, the eleventh, which falls somewhere in between the two pitches. Players would have been required to distinguish between these two notes and therefore would have had to ‘bend’ the pitch of the eleventh resonance up or down accordingly.

1.3 Harmonicity

There has been some discussion amongst players and scholars suggesting that early original brass instruments may have been better suited to the demands of eighteenth-century



Figure 1. Two notes, written F^\sharp and F^\sharp , which are both played using the 11th resonance.

repertoire than their modern counterparts, particularly with regard to pitch bending potential. This is arguably because the resonant modes on these early horns and trumpets are not as harmonically aligned as on either modern instruments, or even copies of early eighteenth-century instruments [2,3], and that this in some way aids the application of early performance techniques.

A related study exploring the acoustical properties of different trumpet mouthpieces, [4] has suggested that slight inharmonicity might in fact be a desirable quality in any brass instrument. From the point of view of pitch bending, misalignment in the series of resonant modes would result in weaker support from higher resonances and could thus, in theory, provide the player with greater flexibility of pitch, i.e. greater pitch bending potential.

1.4 Structure

Section 2 describes the experimental set up and the process for carrying out player tests while Section 3 includes a brief description of the computer model and how the playing properties of a brass instrument have been predicted. The results from measured and predicted data are discussed in Sections 4 and 5 respectively, and conclusions are proposed in Section 6.

2. EXPERIMENTAL SETUP

The horns used for the playing tests included five early orchestral horns from the Edinburgh University Collection of Historic Musical Instruments. Figure 2 shows a photograph of one of these early instruments, a horn from the mid-eighteenth century, made in England by Nicholas Winkings.

For the purpose of this investigation, all horns were crooked or pitched in D, a common key for horn players of the eighteenth century to play in,¹ and the tests were carried out without the hand in the bell in order to limit potential variables.

The photograph in Figure 3 shows the experimental setup for recording the playing tests. The radiated sound from the bell of the instrument was measured using a Brüel and Kjær 4192 pressure-field microphone, located one bell diameter from the plane of the bell of the instrument (approximately 230 mm depending on instrument used), on axis.

In addition to the radiated sound, the mouthpiece pressure was recorded using a 106B PCB Piezotronics dynamic pressure transducer located near the throat of a specially

¹ A horn crooked in D has a tube length of approximately 4.42 m, corresponding to a nominal pitch of D_1 .



Figure 2. Photograph of a horn by the London maker Nicholas Winkings, from the mid-eighteenth century (GB.E.u, 2492).



Figure 3. Experimental set up, showing the positions of the microphones and sound-level meter.

modified mouthpiece as shown in Figure 4. The signals from both transducers were recorded at a sampling rate of 44.1 kHz.

Players were given a familiarisation period with each instrument before starting the recording process. Pitch bending capabilities were explored on a number of different resonances including the fourth, tenth and eleventh. In each

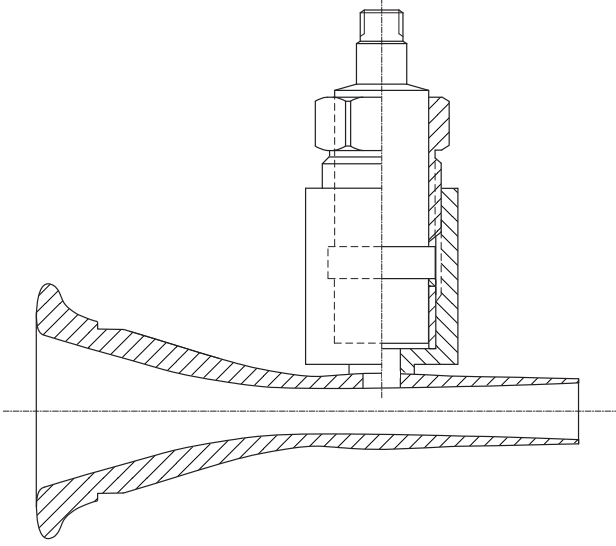


Figure 4. Diagram of a horn mouthpiece with a specially made fitting for a microphone.

case, players were asked to bend the pitch of the resonance downward and upwards as far as they could, until the note suddenly changed to the neighbouring resonance.

Typically the sound samples lasted for about two to three seconds. The test was played at a reasonably loud dynamic level, approximately *forte*, and was repeated at least four times. To assist the players with keeping the sound at a constant dynamic level over the duration of the pitch bend, they were provided with a sound-level meter (set to A-weighting). The players were asked to make no attempt to aim for a specific pitch, only to test the extremes of the flexibility of the pitch on the particular resonance.

The reason for testing the pitch bending capability of the fourth resonance is because it is relatively low in the playing range of the horn and therefore is easier to manipulate with the embouchure. The eleventh resonance was chosen because, although it is higher in the series and more difficult to bend, it is a particularly problematic pitch as mentioned earlier, which typically lies somewhere in between a written F^b and F^\sharp . This resonance would certainly have been subjected to some form of pitch manipulation as composers include both these notes in their compositions.

After the playing tests, the initial analysis of the recordings was carried out using the pitch detection software Praat [5], a program originally developed for speech analysis, but found to work well for analysis of instrumental sounds. Further analysis was carried out by looking at the acoustic input impedance for each horn and mouthpiece combination. The equipment used to measure the acoustic input impedance of the various horns is the commercially available Brass Instrument Analysis System (BIAS) [6].

3. COMPUTATIONAL MODELLING

The use of computer simulation to assess the playing characteristics of an instrument has a significant advantage over player tests because it is non invasive, requires little con-

tact with the instrument, and removes player subjectivity. Brass instruments are perishable and owners of collections are becoming increasingly sensitive to issues of excessive handling and contact.

It is already the case that many eighteenth-century horns are no longer in playing condition due to large dents, leaks, or cracks in the metal. A thick layer of dust inside the tubing or a trapped foreign object could alter the acoustics of an instrument, so in the case of instruments that have already experienced significant wear and tear, and are no longer in playing condition, extrapolated measurements of their bore profiles can be used as the input to computational models in order to predict the acoustic behaviour of the instruments.

The vibration of the lips of brass players can be modelled by considering the lip to be a single mass able to move in two dimensions and obeying a forced damped harmonic oscillator equation which is given by Adachi and Sato [7] as:

$$\frac{\partial^2 \vec{\xi}}{\partial t^2} + \frac{\omega_0}{Q} \frac{\partial \vec{\xi}}{\partial t} + \omega_0^2 (\vec{\xi} - \vec{\xi}_{equil}) = \frac{2b}{m_{lip}} \left((p_m - p_0) (\vec{\xi} - \vec{\xi}_{joint})^\perp + p_{lip} \vec{e}_y d \right). \quad (1)$$

Here m_{lip} is the mass of the lip, ω_0 is equal to the natural angular frequency of the lip, Q is the quality factor of the lip, b is the lip width (perpendicular to the axial direction, x , and the vertical direction, y) and d is the lip thickness in the x direction. The vectors are two dimensional along x and y , with $\vec{\xi}$ giving the displacement of the bottom corner of the lip, $\vec{\xi}_{joint}$ giving the fixed position of the top corner of the lip and $\vec{\xi}_{equil}$ being the equilibrium position of the bottom corner of the lip in the absence of air pressure differences. Bernoulli forces must be taken into account in calculating the pressure beneath the lip, which is denoted by p_{lip} while p_m is the pressure in the mouth, p_0 is the pressure in the mouthpiece immediately beside the lip and \vec{e}_y is the unit vector in the y direction.

The lip model was run with the natural frequency of the lip, ω_0 ramping linearly, either upwards or downwards through the appropriate range for the transient being modelled. The value of Q was multiplied by a factor of $0.25\sqrt{3}$ and values of ω_0 multiplied by a factor of $\sqrt{3}$ when the vertical displacement of the lip was negative (corresponding to the lips overlapping) [9].

Parameters chosen were mainly the same as those found in Adachi and Sato [7], with the lips assumed to be just touching in the absence of an air pressure difference (vertical equilibrium of zero). The lip mass (in kg) depends on lip natural frequency, f_{lip} , according to the empirical formula taken from Adachi and Sato [7]:

$$m_{lip} = \frac{1.5}{f_{lip}(2\pi)^2}. \quad (2)$$

The static mouth pressure was taken to be 6 kPa. This model was then coupled to the Finite Difference Time Domain (FDTD) model of Bilbao [8] which simulates lossy linear wave propagation based on the bore profile. Full details of numerical model, including the FDTD model for

wave propagation in the bore, the method of coupling the lip model and values of the constants used are found in Kemp et al [9]. The results in this case will be the same as that which would be obtained from measuring or computing the instrument time domain reflectance and coupling this to the lip model (within the accuracy of computation or measurement).

4. MEASURED RESULTS

The results from player tests revealed a subtle but measurable variation with regard to the pitch bending potential of specific notes on certain instruments. A horn by the maker Hofmaster (see Figure 5 for a photograph) was relatively amenable to pitch bending on the 11th resonance. Figure 6



Figure 5. Photograph of a horn by the London maker Christopher Hofmaster, from the mid-eighteenth century (GB.E.u, 3297).

shows an example of the change in pitch as the 11th resonance on the Hofmaster horn is bent downwards till the note drops to the the resonance immediately lower.

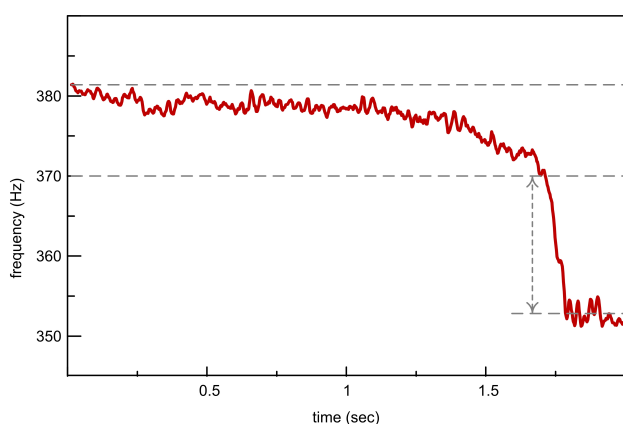


Figure 6. Graph showing the change in pitch as the eleventh resonance is lipped downwards until it drops to the tenth resonance, played on a horn by the maker Hofmaster.

Compare this with Figure 8 which shows an example from a different horn; this graph shows the pitch contour for

an instrument by the maker Sandbach (see Figure 7 for a photograph) which proved to be less flexible on this resonance.²



Figure 7. Photograph of a horn by the London maker Sandbach, from the early nineteenth century (GB.E.u, 203).

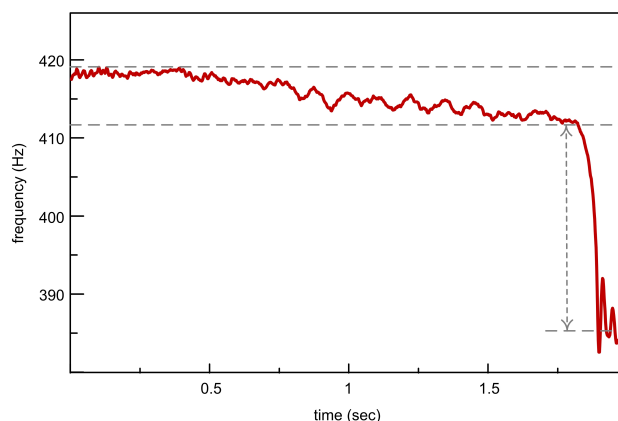


Figure 8. Graph showing the change in pitch as the eleventh resonance is lipped downwards until it drops to the tenth resonance, played on a horn by the maker Sandbach.

The comparison of upward pitch bends on both of these horns showed a similar response to the downward pitch bend results; the Hofmaster instrument responded more readily to embouchure manipulation technique than the Sandbach horn. The maximum range of flexibility on the eleventh resonance, summing the flexibility in the upward and downward directions and taken as an average from all relevant sound clips, was 82 cents for the horn by Hofmaster and 56 cents for the horn by Sandbach.

As an example, the grey lines in figure 6 show the note being lipped in the downward direction from around 382 Hz to around 370, a change of 55 cents to 2 s.f. before the note transitioned rapidly to the tenth resonance (producing a total change in pitch of around 136 cents during the course of the recording). In the case of the Sandbach horn,

² Note: these two horns are from different time periods and therefore have slightly differing nominal values of concert pitch

the example shown in figure 8 shows the frequency produced by the player varying from around 419 Hz to around 411 Hz, a change of 33 cents before the rapid transition to the tenth resonance (giving a total change of 146 cents during the course of the recording). It should be noted that the pitch range analysis based on these graphs is an estimation; sometimes there is not a clearly defined point at which the note appears to make the sudden transition to the neighbouring resonance. This is particularly apparent on the the graph in Fig 6. If the pitch bend was to be carried out at a slightly slower speed, the point at which the transition occurs might be more apparent.

5. PREDICTED RESULTS

The results shown in Section 4 were then compared to predictions from the time domain computer model described in Section 3. Predictions of pitch bending by the model are shown in Figure 9, based on bore profile measurements from the Hofmaster horn, and Figure 10, based on those from the Sandbach horn.

Since the time scale shown (in seconds) for the simulated data transitions is smaller than that for the playing tests, the slope corresponding to the rapid transitions between resonances is shallower, and the exact point of transition harder to establish. A reasonable estimate for the modelled transition for the Hofmaster horn shown in figure 9 gives the frequency changing from 385 Hz to 373 Hz before the rapid transition, corresponding to a downward pitch bend of 55 cents to 2 s.f. (agreeing with the pitch range shown in the player results shown in figure 6). In the case of the Sandbach horn, the modelled transition shown in figure 10 includes a pitch bend from 420 Hz to 410 Hz, giving a variation of 42 cents to 2 s.f., somewhat larger than the 33 cents observed in the player generated note transition shown in figure 8 but agreeing with the conclusions of playing tests with regard to the fact that lipping note pitch is easier on the Hofmaster horn.

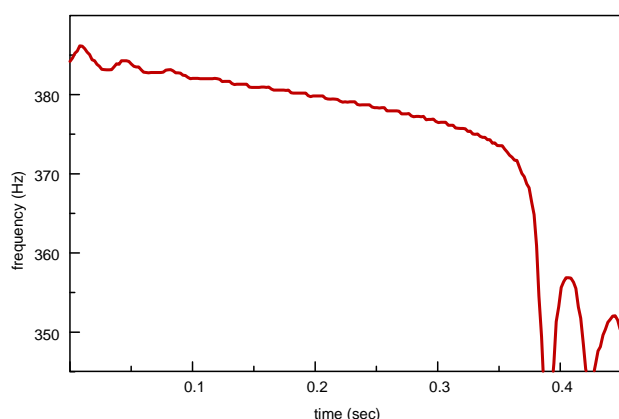


Figure 9. Graph showing the predicted change in pitch, comparable to the measured results in Figure 6, for a horn by the maker Hofmaster.

Lipping notes above or below the input impedance peak for an instrument was achieved by adjusting the value of f_{lip} in the two dimensional lip model. This process fea-

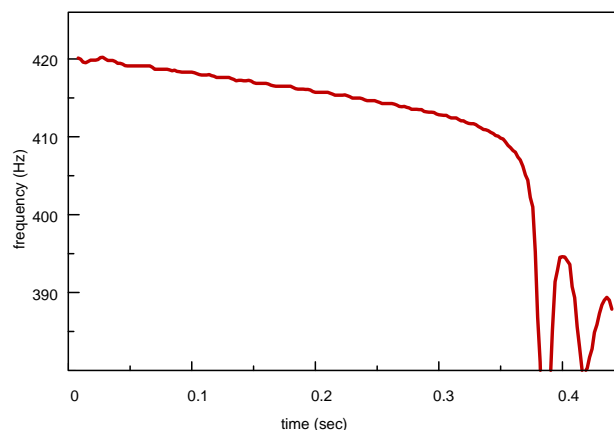


Figure 10. Graph showing the predicted change in pitch, comparable to the measured results in Figure 8, for a horn by the maker Sandbach.

tures hysteresis in that different playing frequencies are possible when the lip frequency is adjusted in an upward or downward direction. The degree to which notes can be bent was also found to vary with the quality factor of the lip resonance.

Figure 11 shows the pitch contour obtained from analysing the pressure in the mouthpiece, taken from the computer model where the lip frequency has been simulated to produce an upward pitch bend over the 10th resonance based on bore profile measurements of a horn by the maker Winkings. Results are plotted for both $Q = 3$ and $Q = 5$.

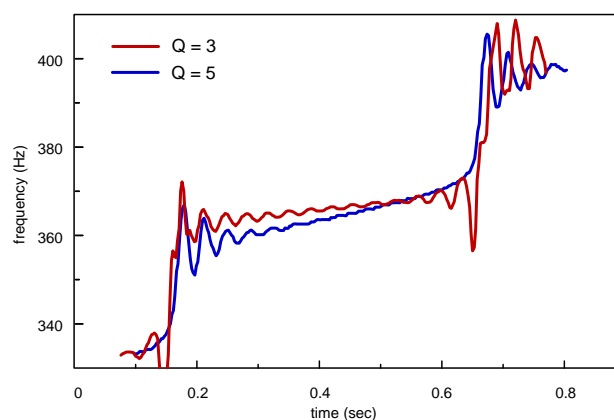


Figure 11. Graph showing a comparison of predicted change in pitch of the 10th resonance with two different values of Q for the lip.

It may be noted that larger values of lip quality factor produce a greater range of sounding frequencies when increasing the lip frequency, while the different quality factors produce very similar ranges of sounding frequency when decreasing the lip frequency.

6. CONCLUSIONS

The results of playing tests and modelling showed broadly similar behaviours for note transition phenomena. While players may vary a large number of control parameters to

perform lipping, there was nonetheless a similar range of frequencies available during lipping in the physical model when adjusting only the lip frequency (and frequency dependent lip mass). Since the quality factor associated with the lip model has an impact on how far a note may be lipped in an upward frequency direction, further work may involve studying whether conclusions can be reached over the quality factor associated with the lip model in different playing ranges, in order to provide input parameters for realistic physical modelling synthesis of such phenomena.

Playing tests confirmed that the mid-eighteenth horn by Hofmaster was easier to lip than the nineteenth century Sandbach horn, at least for the notes studied in this project. The results of FDTD time domain simulations using the measured bores of these instruments were in qualitative agreement with the playing tests, demonstrating that the difference in lipping behaviour was not dependent on specific human players and their expectations. Future work will attempt to identify the features of the instrument bores, and their associated input impedance curves, which are responsible for these musically significant differences in playability.

7. REFERENCES

- [1] A. Benade, "The physics of brasses," *Scientific American*, pp. 24–35, 1973.
- [2] A. Fromme, "Performance technique on brass instruments during the seventeenth century," *Journal of Research in Music Education*, vol. 20, no. 3, pp. 329–343, 1972.
- [3] D. Smithers, K. Wogram, and J. Bowsher, "Playing the baroque trumpet," *Scientific American*, vol. 254, no. 4, pp. 108–115, 1986.
- [4] E. Poirson, J.-F. Petiot, and J. Gilbert, "Study of the brightness of trumpet tones," *Journal of the Acoustical Society of America*, vol. 118, no. 4, pp. 2656–2666, 2005.
- [5] P. Boersma, "Praat, a system for doing phonetics by computer," *Glott International*, vol. 5, no. 9/10, pp. 341–345, 2002.
- [6] G. Widholm, "Bias 5.1 manual," Vienna: IWK (MA), 2001.
- [7] S. Adachi and M. Sato, "Trumpet sound simulation using a two-dimensional lip vibration model," *Journal of the Acoustical Society of America*, vol. 99(2), pp. 1200–1209, 1996.
- [8] S. Bilbao, "Time domain simulation of brass instruments," in *6th Forum Acusticum, Aalborg, Denmark*, 2011.
- [9] J. Kemp, S. Bilbao, J. McMaster, and R. Smith, "Wave separation in the trumpet under playing conditions and comparison with time domain finite difference simulation," 2013, submitted to the *Journal of the Acoustical Society of America* and is in press.

Percussions

THE ROLE OF DAMPING IN STEEL PAN MANUFACTURE

Claire Y Barlow
University of Cambridge
cyb@eng.cam.ac.uk

Soren E Maloney
University of Cambridge
Department of Engineering
Trumpington Street
Cambridge CB2 1PZ UK
soren.maloney@gmail.com

Jim Woodhouse
University of Cambridge
jw12@cam.ac.uk

ABSTRACT

Results are presented of systematic studies of vibration damping in steel of a type, and processed by a route, relevant to Caribbean steel pans. Damping is likely to be a significant factor in the variation of sound quality between different pans. The main stages in pan manufacture are simulated in a controlled manner using sheet steel, cold-rolled to a prescribed level of thickness reduction then annealed at a chosen temperature in a laboratory furnace. Small test strips were cut from the resulting material, and tested in free-free beam bending to deduce the Young's modulus and its associated loss factor. It is shown that the steel type, the degree of cold working and the annealing temperature all have a significant influence on damping. Furthermore, for each individual specimen damping is found to decrease with rising frequency, approximately following a power law. Comparison with samples cut from a real pan show that there are further influences from the pan's geometrical details.

1. INTRODUCTION

Caribbean steel pans are a comparatively new musical instrument, dating in their modern form from around 1945 and invented by the native inhabitants of Trinidad in response to a ban on their traditional instruments. The original instruments were made from old oil drums, readily available on the island; the drums often still used by contemporary pan-makers are no more than commercially available oil drums.

Steel pan manufacturers have traditionally used whatever drums the manufacturers were able to sell them. However, from time to time they encountered drums which did not produce satisfactory pans. Sometimes they would tear or split – so in some way demonstrating lack of formability or ductility. This is a familiar metallurgical problem, but one that will not be pursued further here. A more subtle failure mode for the pans was that they would sometimes produce an unsatisfactory sound – once described to the authors by pan-maker Dudley Dickson as

sounding 'like lead'.

This description may point to a problem involving abnormally high damping. It is natural to ask whether significant variations in damping can arise from the properties of the material, either in terms of the chemical makeup of the steel or as a result of alterations made during the manufacturing process of a steel pan. An earlier investigation looked at chemical composition, with no very clear conclusions emerging [1]. Between samples of steel found good and bad by pan-makers, there were differences in distribution and numbers of precipitates, particularly manganese sulphide, but this was not obviously linked with damping. This paper will give some key results from a comprehensive study looking at the effect of metalworking processes on the vibration characteristics of steels, with particular emphasis on damping [2].

2. PAN MANUFACTURE

The manufacture stages for pans are shown in Figure 1, comprising a series of traditional metalworking processes of cold-working and annealing. The way in which these are done by pan-makers working in small-scale operations is labour-intensive and is very much a craft. The first stage is to "sink" or dish the top of the oil drum to a depth of 100-200mm, using a sledge or pneumatic hammer. Then the notes are marked out and delineated by a groove made using a punch. The drum is then "fired" or "burned" – traditionally, this means being heated over a bonfire, but propane gas torches are also used.

The separate note areas are then tuned: this involves hammering to produce small changes to the geometry. Some local annealing with a blowtorch may be used during the tuning process. Initial hammering is from the back of the dished pan, so that the note areas become flat or slightly arched upwards. The characteristic sound of the pan results from vibration mode confinement within each note area, controlled by the geometrical changes in the tuning process. Notes can be visualised, approximately, as locally flat regions set within the spherical curvature of the dished pan. Bending vibration with relatively long wavelengths within this flat region cannot propagate into the curved surrounding region because of the stiffening effect of the curvature. A fundamental mode is always tuned, and 1–3 further modes may also be tuned to harmonically related frequencies within a single note (see for example [1,3]).

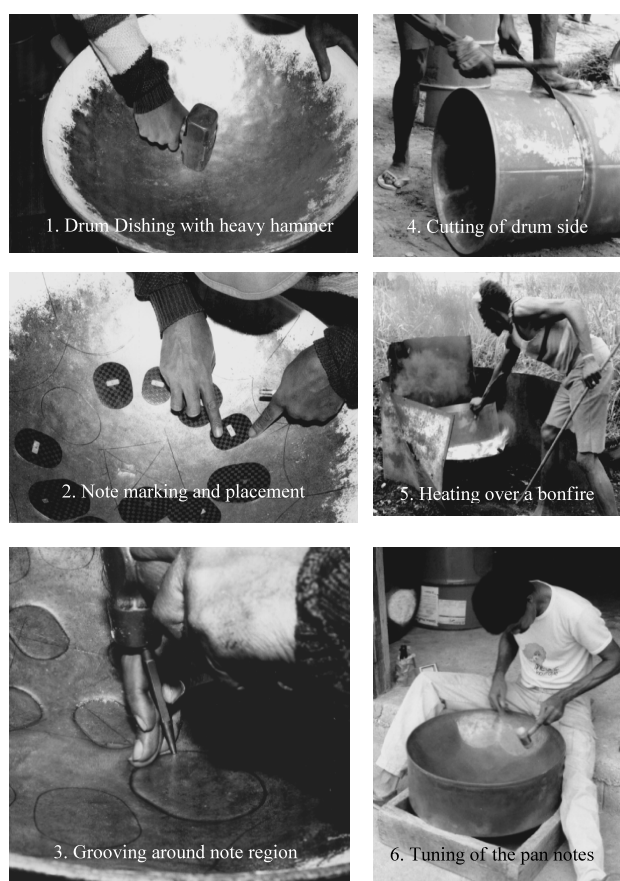


Figure 1. The main stages in pan manufacture [4].

The dishing process involves stretching the steel by different amounts across the diameter of the pan surface, so producing a range of thicknesses. Notes delineated in different parts of the pan may be expected to have a range of mechanical histories and so may behave in different ways when subjected to annealing heat treatments. The tuning process introduces a further stage of local stretching and, perhaps, heat treatment. This motivates the studies to be reported in this paper: how is the damping of relevant types of steel influenced by the strain induced by cold-working, and by annealing heat treatment?

3. DAMPING IN STEELS

The influence of features of metallurgical structure on damping in steels has been investigated by a number of researchers. Damping is a consequence of processes that absorb energy within the metal structure when it is subjected to small cyclical strains as the object vibrates. The main mechanism relevant to the present question, probably, is associated with mobile dislocations and other atomic-level transitions. In response to a sustained stress above a critical level, dislocations can move progressively, leading to ductile plastic deformation. In response to small cyclical stresses during vibration, the atoms at the tips of the dislocations can hop back and forth across the dislocation front. There is no cumulative deformation, but energy is lost with each hop into lattice vibration, i.e. to heat. In consequence, it tends to be the case

that ductile metals like lead have high damping, whereas harder, more brittle metals have lower damping. This immediately suggests that pan makers may have to reach a compromise: they need sufficient ductility to form the pan and keep it in tune, but they don't want too much damping in the final product.

Metallurgical features implicated in energy dissipation are the number of dislocations, the number and type of point defects (vacancies, interstitial solute atoms such as carbon, hydrogen or nitrogen), the grain size, and the type and distribution of phases present. A further factor may be residual internal elastic stresses in the steel, resulting from inhomogeneous plastic deformation. Finally, damping may be affected if the steel is coated with chrome, or indeed with any other surface finish. A consequence of all these factors, which are not independent, is that understanding exactly what is happening in steel pans is not straightforward.

The main effect of room-temperature plastic deformation, cold working, is to increase the number of dislocations in the metal; there will also be an increase in the number of vacancies. Damping increases with the number of dislocations (e.g. Eshelby [5]), so cold deformed steel is expected to have increased damping. Internal residual elastic stresses are cited as a cause of increased damping by Zener [6]. Annealing the metal, heating it, reduces the number of dislocations, reduces or removes internal elastic stresses and also reduces the number of vacancies. Annealing is typically associated with a reduction in damping [6]. However, annealing at a temperature high enough to cause full recrystallisation of the steel has been found to result in increased damping, perhaps because of an increase in grain size [7].

Increasing the amount of carbon in the steel is associated with changes in damping. A particular interaction between interstitial carbon atoms and dislocations, known as strain ageing, results in pinning of dislocations. This might be expected to result in reduction in damping, but some authors have cited evidence for strain ageing being associated with increased damping [8]. In the light of these uncertainties in the literature, direct experimental measurements are needed for the types of steel and processing history relevant to pan-making.

4. EXPERIMENTAL STUDIES

Damping is not easy to measure reliably. For many measurements of physical properties of materials, the errors arising from the uncertain details of experimental procedure are random, so that an average over many tests will improve the accuracy of the result. For damping this is not the case: virtually all flaws in experimental conditions lead to systematically *increased* damping because additional energy is dissipated by, for example, the way the specimen is held, or by the effects of any attached instrumentation, or by interaction with the surrounding air. None of these effects ever lead to reduced damping, and averaging many tests does not help to identify such faults. The only safe approach is to try alternative ways of support, instrumentation and so on, and to place most

trust in whichever gives consistently the lowest values of damping.

The approach to be used in this work is based on vibration testing of thin strips of steel. The strips are excited so as to vibrate in bending, and the results are interpreted via Euler-Bernoulli bending theory according to which the only relevant elastic modulus is the Young's modulus in the direction parallel to the strip's long axis. That Young's modulus will have a complex value $E(1+i\eta)$, at a given frequency, and the dimensionless loss factor η characterises the damping. Results will be presented in terms of the "quality factor" $Q=1/\eta$, so that high Q-factor denotes low damping, and vice versa.

It is difficult to extract strips big enough for these measurements from actual pans. The length scales of note regions in a pan do not allow conveniently large specimens to be obtained, and in any case we will not usually know the exact strain and heat-treatment history of a particular patch on a particular pan. It is far better to use specimens prepared for the purpose using precisely controlled processing steps, although some corroborative evidence was also gained from cutting specimens from notes on actual pans.

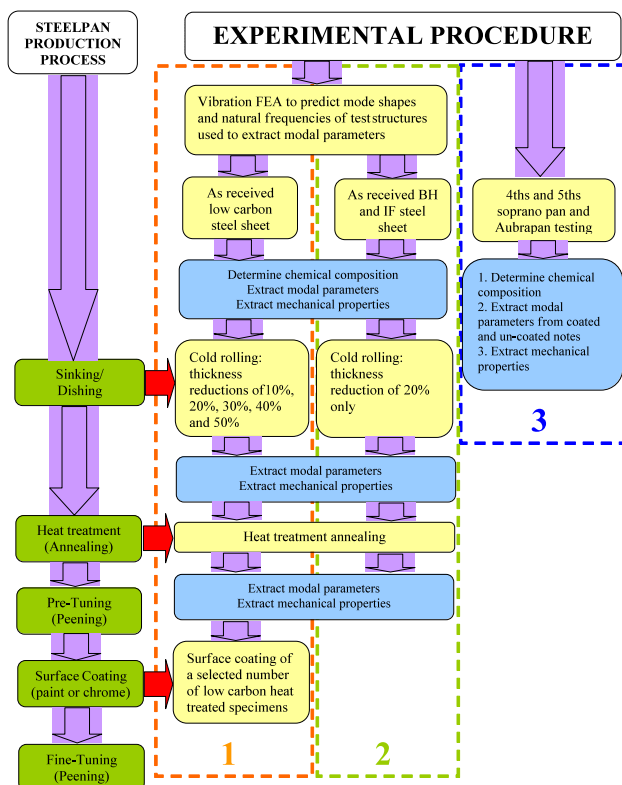


Figure 2. Mapping of the laboratory experiments to the stages of pan manufacture.

Controlled tests on the kind of low-carbon mild steel used in pans were carried out using strips of sheet steel. Samples of two specialized ultra-low carbon steels were also tested: interstitial-free (IF) and bake-hardenable (BH). These were chosen not because they are currently used in pan-making but to try to separate out some of the complex factors implicated in damping. IF steels contain very low amounts of free carbon and nitrogen and are

highly formable, but at the expense of relatively low strength. BH steels were designed for making automotive body panels, and contain low, controlled amounts of carbon (typically 0.002%). The material is initially formable to allow the panels to be pressed easily, but is then strengthened during the paint stoving operation which bakes the paint on the panels.

The steel samples were subjected to cold rolling (with thickness reduction ranging from 10 to 50%) to simulate the dishing deformation. The rolling direction was chosen to be the same as that of the sheet steel: the mechanical properties are anisotropic so it is important to be clear about the orientation of the sample with respect to the rolling direction. Annealing was carried out at a range of temperatures spanning what is used in practice. The samples were wrapped in stainless steel foil to inhibit surface oxidation, and heated in a laboratory fluidised-bed furnace for a period of 10 minutes at the desired temperature. Samples were then either allowed to air-cool, or quenched in water. The correlation between the test regime and the pan manufacturing process is illustrated in Figure 2.

Beam specimens for vibration testing were cut to size 150×15 mm by guillotine, after all processing was complete. Two small holes were drilled in each strip near a pair of corners along one long side, and the strip was suspended from these holes using fine nylon thread: preliminary tests revealed that this method gave the lowest damping of all methods tried. A miniature impact hammer (PCB 086D80) was used to excite vibration. The hammer was held in a pendulum fixture to ensure that the same point was hit every time. Vibration response was measured with a laser-Doppler vibrometer (Polytec OFV-056). To obtain good reflection of the laser beam, retro-reflective dust was applied to the beam: it was found that reflective tape was unsuitable as it resulted in unacceptable levels of additional damping.

Transfer functions were measured using 15 repeat impacts and then performing an average using a PC-based data-logger and software in Matlab. The associated coherence function was monitored as a check on data quality. Transient time histories were also recorded. Modal frequencies were extracted from the transfer function peaks, and Q-factors were determined from fitting time-decay envelopes for each mode, obtained by processing the transient time histories into a sonogram (or spectrogram): see [9] for more detail. Modes up to 4 kHz were analysed: this frequency range covers the main sound production from steel pans.

5. RESULTS AND DISCUSSION

It was found that almost every factor investigated had some influence on damping behaviour. Different compositions of steel behave differently; for the mild steel, the degree of reduction by rolling and the temperature of annealing made a difference; for every individual specimen, damping was found to be quite a strong function of frequency; surface coatings made a difference; and fi-

nally, tests on samples extracted from a real pan showed a significant influence of the geometry of tuned notes. The only factor which did not seem to matter very much was the difference between air cooling and water quenching. These statements will be illustrated by examples selected from the rather extensive set of data from the full set of tests.

Figure 3 shows the damping behaviour of the three types of steel as received. All three show an increasing trend of Q-factor with frequency, or equivalently a decrease of damping with rising frequency. The low-carbon mild steel and the IF steel show rather similar trends, while the BH steel has lower damping at all frequencies. Error bars indicate the total spread across repeated measurements of each value: in most cases the spread is small, indicating good reliability of the results.

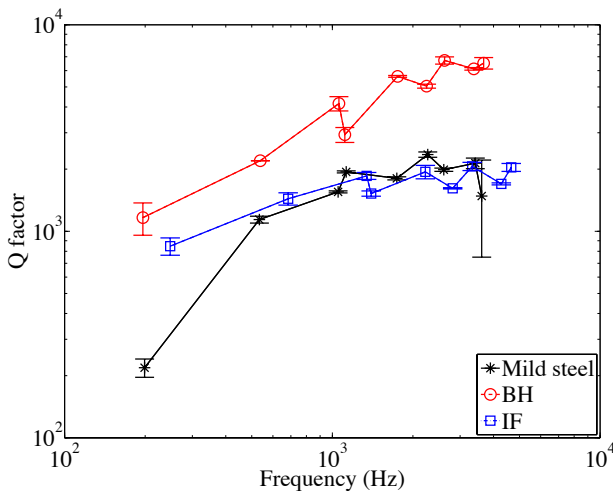


Figure 3. Variation of Q-factor with frequency for the three types of steel in the as-received condition.

Modes 1, 2, 3, 5, 7 and 9 are bending modes, but modes 4, 6 and 8 are torsional modes so that their damping depends on the complex shear modulus as well as the complex Young's modulus [10]. This accounts for the downward "jogs" visible in the figure for these modes, implying that the shear modulus has somewhat higher associated damping than the Young's modulus. From now on, these torsional modes will be omitted, for clarity of interpretation. Error bars are also omitted from now on: all damping measurements have comparable reliability to that shown in Fig. 3, and the trends to be shown later are much stronger effects than this uncertainty.

Figure 4 shows the effect of cold rolling on the mild steel, before any annealing is done. Each mode except the first shows a broadly similar trend: there is a large increase in damping (drop in Q) with the first increment of rolling to give 10% reduction, then a slight further increase with further reduction. The trend to lower damping as frequency increases, seen in Figure 3, manifests itself here by the curves for successive modes appearing approximately in order from bottom to top in the plot. The reason for the first mode behaving differently from the rest is at present a mystery.

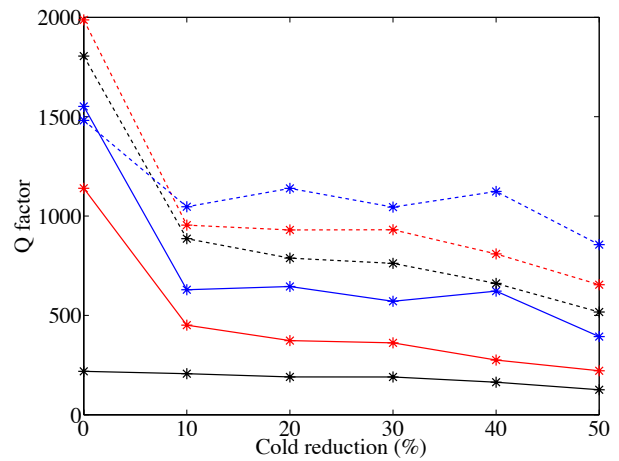


Figure 4. Variation of Q-factor of mild steel with the percentage thickness reduction in cold rolling, before annealing. Curves show the bending modes 1, 2, 3, 5, 7, 9 in order from bottom to top. Values at zero reduction show as-received material.

The next stage involves annealing. For one typical example of the mild steel data, after 20% reduction, Figure 5 shows the variation of damping with annealing temperature. Each mode shows a similar trend: damping is reduced by annealing until the temperature reaches about 350°C, at which point it reaches a plateau. The shape of these curves suggests that there may be (at least) two different mechanisms of damping contributing to the total: one is progressively reduced by annealing, the other is largely unaffected by annealing and is responsible for the plateau.

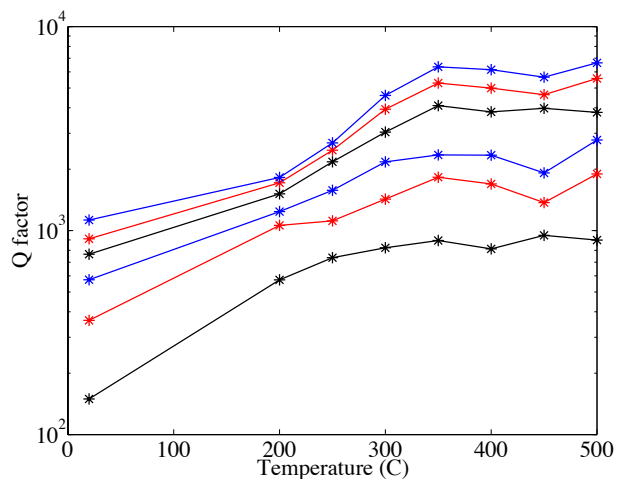


Figure 5. Variation of Q-factor with annealing temperature for mild steel after 20% reduction by cold rolling. Curves show the bending modes 1, 2, 3, 5, 7, 9 in order from bottom to top. Values at 20°C show results before annealing.

Figure 6 repeats these results for one particular mode, and compares them to results for the corresponding mode in strips made from the BH and IF steels. It can be seen that the mild steel tends to lie between the IF and BH results: close to IF with little or no annealing, but rising to a plateau level rather similar to the BH steel, while the IF steel tends to have higher damping until the very high-

est annealing temperatures. Results like this probably give clues about the underlying mechanisms.

Figure 7 shows the same results as Figure 5, but presented in a different way. Q-factor is plotted against frequency for each annealing temperature, on log-log scales. It is immediately clear that the results tend to align along straight lines in this plot, corresponding to power-law variation of Q with frequency f of the form $Q = \beta f^\alpha$ where α, β are constants that can be readily determined by a regression fit to the measurements, indicated by the lines in the plot.

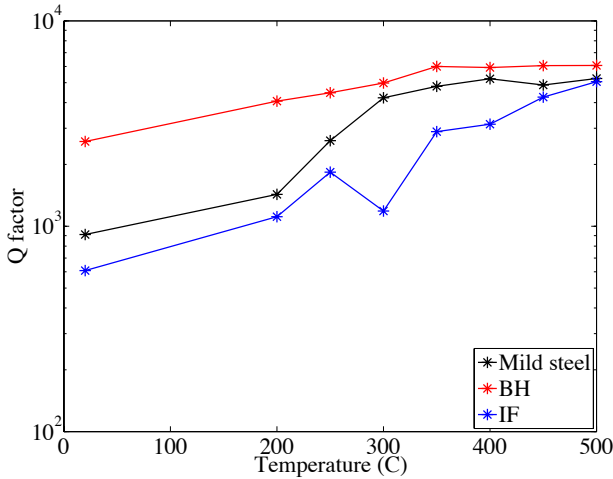


Figure 6. Variation of Q-factor with annealing temperature for bending mode 7, for the three types of steel tested. Values at 20°C show results before annealing.

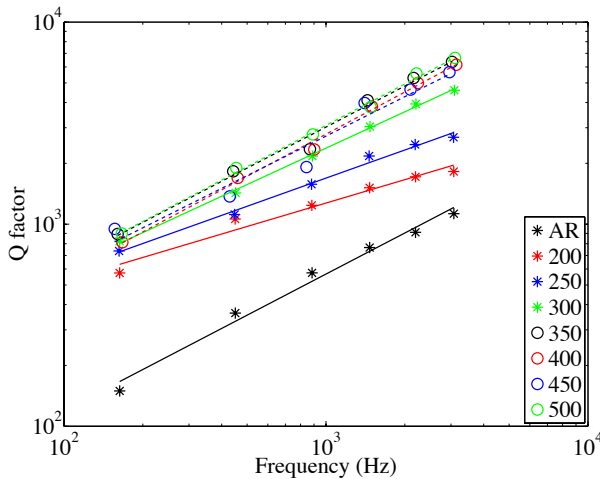


Figure 7. Variation of Q-factor with frequency for mild steel after 20% reduction by cold rolling, for different choices of annealing temperature. Straight lines show linear least-squares fits to the data plotted on log-log axes, to extract power-law variation. “AR” denotes as-received material

The fitted exponent α is plotted as a function of annealing temperature in Figure 8. Three different levels of thickness reduction in the mild steel samples show very similar values, while the BH and IF steels look significantly different. It should be noted that the regression fits cease to be convincing for the IF steel at lower annealing temperatures: the data points for this steel should only be

trusted for 350°C and above. The trends revealed in this plot raise a challenge for modelling the mechanisms at work, a task for future research.

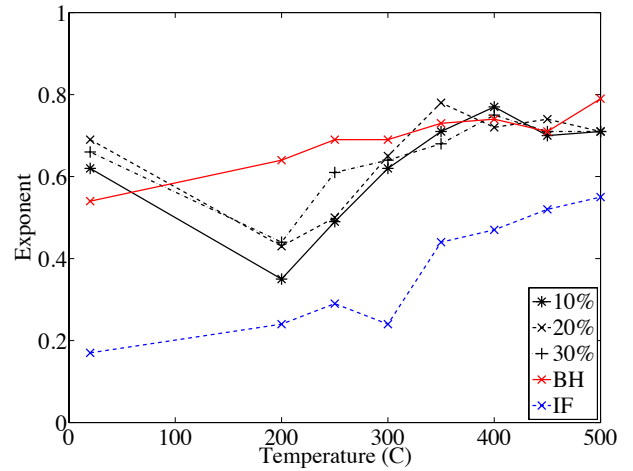


Figure 8. Variation of power-law exponent from the fitting illustrated in Figure 7, for mild steel after three different levels of thickness reduction and for BH and IF steel after 20% reduction. All points give a good regression fit, except for the IF steel at temperatures of 300°C and below: these 4 data points should not be trusted.

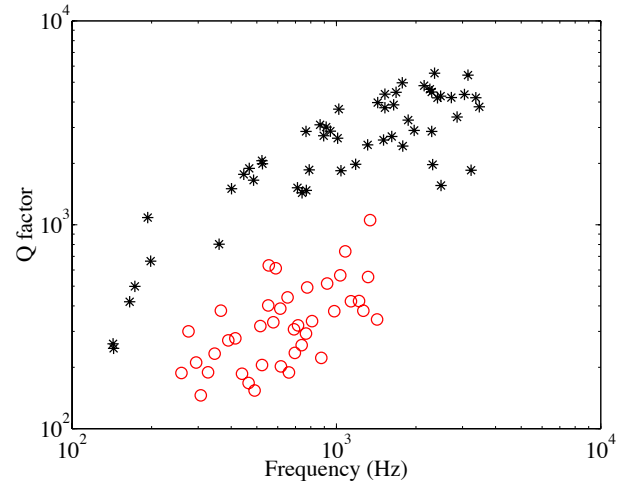


Figure 9. Variation of Q-factor with frequency for small strips extracted from a pan (*), and for confined modes of the notes of the same pan before sample extraction (o).

The final measurements to be shown relate to a real steel pan, rather than to the sheet steel samples discussed so far. A pan was donated by its maker for this project. Q factors were measured from the confined modes of the separate notes of the pan, then small test strips were cut out of the pan and tested in the same manner as the earlier measurements. The results are summarised in Figure 9. Both sets of points show the now-familiar rising trend of Q-factor with frequency, but it is immediately clear that the pan notes have much higher damping than is associated with the underlying material alone. This points to the fact, presumably, that “confined” modes are not quite as confined as one might have imagined: some energy is “leaking out” of the note regions into the rest of the pan,

where it can experience higher damping by mechanisms such as micro-friction around the seam between the head and skirt of the pan.

Evidence of leakage causing inter-note couplings has been shown in earlier holographic studies [1,3]. The present results suggest that further investigation might be carried out, to elucidate how the rate of leakage is influenced by the geometrical details of individual note areas. These details vary greatly: larger note areas tend to be around the outer edge and so are closer to the edge seam. Larger notes also tend to be more crowded together, so that there is less of the spherical dishing around them to effect the mode confinement. Pan-makers probably take advantage of such effects in manipulating tonal quality.

6. CONCLUSIONS

This paper has presented some measurements of the damping behaviour of various steels, including the influence of controlled amounts of cold-working and annealing. The study was motivated by exploring the link between acoustical performance and manufacturing details in Caribbean steel pans. However, it may turn out that the results are important in a wider context because they reveal an unexpectedly rich structure in the damping behaviour of steels. In particular, damping shows a strong variation with frequency, in most cases following a power law reasonably accurately.

The exponent of this power law varies with the composition of the steel, and with the level of cold working and heat treatment, but it does not match the prediction of the simple Debye theory that is normally used to explain frequency-dependent damping in metals [7, 11]. That theory predicts a damping peak associated with each metallurgical mechanism, at a frequency determined by the corresponding relaxation time. The only way a rising trend of Q following a simple power law can arise from that theory is in a frequency range above the peaks, when an exponent with value -1 would be expected. That is not the behaviour seen here: these results raise challenges for modelling the underlying mechanisms of damping in the regime relevant here.

In terms of the specific application to steel pans, perhaps the most surprising result is shown in Figure 9, where damping factors of notes in a pan are compared to the material damping from strip samples extracted from those same note regions. The pan notes have significantly higher damping than the underlying material. This must mean that the “confined” modes of the pan notes are less well confined than those in the closely related problem of confined modes in the “musical saw” [12]. There is obviously scope for theoretical investigation of this difference between one-dimensional and two-dimensional curvature as they affect mode confinement.

Although the damping of pan notes as played is clearly influenced by factors in addition to the local material damping of the steel, nevertheless one would expect the damping behaviour documented here to play some sig-

nificant role. It is a commonplace of musical acoustics that relative small physical effects often have unexpectedly strong perceptual significance to instrument makers and musicians. When a pan maker describes a particular drum as “sounding like lead” they may very well be noticing an effect of the kind documented here. To avoid material uncertainty of this kind, pan makers might find it worthwhile to explore the use of bake-hardenable steel to give them a consistent material that has sufficiently good formability but that can be reliably brought to a stable and low-damping state by a simple heating procedure.

Acknowledgments

The authors thank Prof Harry Bhadeshia for helpful discussions and obtaining the BH and IF samples, and Mr. Aubrey Bryan for donating the pan for destructive testing.

7. REFERENCES

- [1] C. Y. Barlow, B. E. Richardson and J. Woodhouse, “The Caribbean steel drum”, *Proc. Inst. Acoustics*, vol. 10 Part 2, pp. 193-198, 1988.
- [2] S. E. Maloney. Acoustics and manufacture of Caribbean steel pans. PhD thesis, University of Cambridge, 2010.
- [3] N. H. Fletcher and T. D. Rossing, *The physics of musical instruments*, Springer-Verlag, 1991
- [4] U. Kronman, *Steelpan tuning: a handbook for steelpan making and tuning*, Stockholm, Sweden, Musikmuseet, 1991.
- [5] J.D. Eshelby, “Dislocations as a Cause of Mechanical Damping in Metals”, *Proc. Royal Soc. London, Series A*, vol. 197, pp. 396-416, 1949.
- [6] C. Zener, “Internal Friction in Solids IV. Relation Between Cold Work and Internal Friction” *Physical Reviews*, vol. 53, pp. 582-586, 1938.
- [7] C. Zener, “Internal friction in solids”, *Proc. Physical Society*, vol. 52, p. 152-166, 1940.
- [8] A.H. Cottrell and B.A. Bilby, “Dislocation Theory of Yielding and Strain Ageing of Iron”, *Proc. Physical Society Section A*, vol. 62, pp. 49-62, 1949.
- [9] J. Woodhouse and R.S. Langley, “Interpreting the input admittance of violins and guitars”, *Acta Acustica united with Acustica*, vol. 98, pp. 611-628, 2012.
- [10] M. E. McIntyre and J. Woodhouse, “The influence of geometry on linear damping”, *Acustica*, vol. 39, pp. 209-224, 1978.
- [11] M.S. Blanter, I.S. Golovin, H. Neuhäuser and H.-R. Sinning, *Internal friction in metallic materials*, Springer-Verlag, 2007.
- [12] J. Woodhouse and J.F.M. Scott, “Vibration of an elastic strip with varying curvature”, *Phil. Trans. Royal Soc. London*, vol. A339, pp. 587-625, 1992.

An objective approach for assessing the tuning properties of historical carillons

Vincent Debut, Miguel Carvalho and Jose Antunes

Applied Dynamics Laboratory

Campus Tecnológico e Nuclear, Instituto Superior Técnico/Universidade Técnica de Lisboa

2686-953 Sacavem, Portugal

mailto:vincentdebut@ctn.ist.utl.pt

ABSTRACT

The carillons of the Mafra National Palace are undergoing a restoration project. Together, the pair of carillons represent the largest surviving 18th century carillons in Europe. To guarantee the historical significance of these outstanding musical instruments, a detailed diagnosis of their current tuning state was achieved and results were analyzed with respect to historical, acoustical and musical concerns. In a first stage, we developed a suitable polyreference modal identification technique to infer the tuning status of bells from their modal parameters and we then systematically performed modal testing experiments on a selection of historical bells of the Mafra carillons. For each carillon bell which plays a separate note of the instrument, tuning charts displaying the frequency relationships between its most important partials were obtained, as well as the modeshapes, decay times and beating frequencies between modal-doublers for every single musical partial of the bell. In a second part, since carillon bells also must be tuned very accurately one relative to the others, the important topic of estimating the reference pitch and musical temperament of the musical instrument was addressed by developing optimization techniques. After presenting the modal identification procedure and optimal strategies devised for this work, the feasibility and interests of this instrumental approach are finally illustrated for one of the Mafra carillons.

1. INTRODUCTION

Beyond its historical significance as a religious and architectural heritage, the Mafra National Palace constitutes also a notable research resource for musicologists. Located in Portugal, near Lisbon, the building is one momentous baroque palace-monastery (see Figure 1) which comprises six pipe organs and two carillons from the 18th century [1]. Neglected for decades, the Mafra Palace is currently undergoing a restoration process which has recently been awarded by the European Union Prize of Cultural Heritage, and for which one challenging aspect still remain the tuning and restoration

of the ensemble of carillons. For reasons of preservation of cultural heritage, the collection of the 102 historical bells must deserve special scientific considerations and their recuperation must comply with scientific criteria of both musical acoustics and musicological concerns. Before attempting any restoration work, a diagnosis of their current state is essential and a clear idea of their actual tuning must be asserted by paying attention to the standards of bell tuning and performance practice at the time and place where the instruments were played.

To that end, an instrumental technique for assessing the tuning properties of historical carillons was developed. The method is rooted in the vibrational analysis of the carillon bells and examines issues related to musicological features by devising suitable optimization strategies. This enables particularly to characterize the sounding properties of each note of the instrument by examining the modal parameters of each separate bell. The important topic of estimating the reference pitch and musical temperament of the complete carillon was then addressed with respect to criteria of musicological concerns, by formulating the problem in terms of optimization techniques. Objective is then to minimize the difference between the identified modal frequencies and the frequencies of a given musical temperament. As a result, an optimal reference pitch is obtained and one can therefore estimate how the instrument was played in the past, by testing different musical temperaments. It must be emphasized that if important papers are devoted to the casting, tuning and modelling of individual bell [2–5], only few works are concerned with the detailed study of the musical features of a complete carillon [6, 7].

In this paper, the measurement and post-processing techniques used for the modal identification of the historical bells are described and then, the optimization problem is presented. Finally, the feasibility and interests of the technique are illustrated for one of the Mafra carillon, built by the renowned founder Willem Witlockx (1669-1733). If some tuning differences among the collection of bells forming the carillon are highlighted by our identification, results revealed that errors in bell casting and tuning are larger than differences between temperaments which demands stringent tuning precision. It thus prevents any conclusion on the targeted musical temperament of the instrument. The technique, which has the specificity to be objective, is instrumental and can benefit the importance of a cultural heritage when similar

problematics are concerned.



Figure 1: The Mafra National Palace and its two carillons.

2. VIBRATIONAL PROPERTIES OF BELLS

When struck, a bell vibrates in a rather complicated way which results in a complex sound. The emitted tone comprises a large number of partials which all combine together to create a resultant pitch. To obtain pure and clear tones, bell founders deliberately tune the modal frequencies of the lower five modes, which typically (but not always) fall into specific musical frequency relationships: $0.5f_0$ (hum note), f_0 (prime), $1.19f_0$ (minor third), $1.5f_0$ (fifth), $2f_0$ (octave).

An other specificity of bell sound stems from their axial symmetry. Normal modes occur in degenerate pairs which become non-degenerate when slight alteration of the symmetry is introduced, and consequently, one partial is constituted by a pair of modes with near-identical frequencies. In general, because of the inaccuracy inherent to the delicate process of bell casting and tuning, perfect symmetry in the bell geometry is rarely achieved, resulting in a characteristic beating pulsation in the sound of bells. The amount of beats between modal pair frequency components is traditionally said to influence the qualities of bell [3, 8]. Although bell founders and researchers have only a vague idea of the degree of beats needed for a good quality bell, the identification of the two modal families undoubtedly constitutes relevant information for assessing their acoustic properties, in particular to quantify what campanologists called the *warble* [8]. For the case of the Mafra carillons, such data are of even more importance since no information on the mode-doublets can be found in previous studies [6]. For these reasons, a specific attention was paid for the experimental modal identification of the collection of bells, and to that end, a robust modal identification techniques capable of discriminating very close modal frequencies was developed.

3. EXPERIMENTAL MODAL IDENTIFICATION

3.1 Experimental test rig

Bells were manually struck on their original support with an impact testing hammer, at several location. For each

bell, a mesh of 32 test locations regularly spaced near the rim was defined and impact excitation was performed on all of the points, resulting in a collection of 96 impulse responses functions. The vibrational radial responses were recorded with three piezoelectric accelerometers, coupled to charge amplifiers, and the acquired time signals were 12s long. The accelerometers were glued on the outer rim of the bell, in the same horizontal plane, at 3 positions defined by angles of 0° , 22.5° and 146.25° . A Siglab/Spectral Dynamics acquisition board (Model 20-42) ensures the analog digital conversion. Care was taken to roughly control the frequency content of each impact excitation, and because of the large differences in size for the carillon bells (diameters from 0.21m to 2.45m), two instrumented impact hammers were employed during the measurements to ensure a proper excitation of the low frequencies.

3.2 Modal identification approach

Modal identification was achieved by implementing a sophisticated MDOF algorithm, called the Eigensystem Realization Algorithm [9]. The technique has been recognized as being very effective for the modal identification of complex systems. In particular, it provides good estimates of the modal parameters for structures presenting repeated eigenfrequencies due to the specificity of being a general polyreference multi-input/multi-output approach. The algorithm works in the time domain and is based on a state space formulation of the system dynamics. In essence, it attempts to identify a linear mathematical model to match the impulse responses of the structure by using a least-square fit. It combines a set of free decay responses in the form of a generalized Hankel matrix and then uses a singular value decomposition to estimate the minimum order of the mathematical model. The last step of the algorithm consists in computing the eigenvalues of the chosen minimum model from which the modal parameters of the system are extracted. Mathematically, the analytical model considered for the impulse response h_{ij} measured at j from an excitation at location i is given by:

$$h_{ij}(t) = \sum_{r=1}^{2N} A_r^{ij} e^{\lambda_r t} \quad (1)$$

where λ_r are the complex eigenvalues and A_r^{ij} are the complex modal amplitude coefficients, N being the order of the original system. The modal frequencies ω_r and modal damping value ζ_r are extracted from the λ_r noting that:

$$\lambda_r = -\zeta_r \omega_r + j\omega_r \sqrt{1 - \zeta_r^2} \quad (j^2 = -1) \quad (2)$$

Finally, the modeshapes of the system stem from the knowledge of the modal amplitude coefficients at each location along the rim, which can be either computed from the ERA realization matrices or by best-fit with respect to some measurements.

In practice, the presence of noise in the input data perturbs the identifications and in general, it manifests through the identification of nonrealistic modes. Consequently, the model order should be systematically

overestimated, of about ten times the true model order, and this makes delicate the estimation of the model size in practice. To overcome such a difficulty, a stability diagram was implemented as part of the ERA algorithm. By tracking the estimates of the modal parameters as a function of the model size, it is a useful tool to assist in the selection of the system modes: indeed, the physical modes tend to stabilize at low model order whereas nonphysical modes do not stabilize at all during the process because of the random nature of noise. In addition, the identified modeshapes of the bell vibrations were also used as an indicator for the selection of the physical modes. Finally, the overall success of the estimation procedure is achieved by comparison of the synthesized impulse response functions and transfer functions with the ones obtained from measurements.

4. PROCESSING OF VIBRATION SIGNALS

The modal parameters of each bell were extracted proceeding as follows:

1. Computation of all the transfer functions $H_{ij}(\omega) = \ddot{X}_i(\omega)/F_j(\omega)$ by Fourier transform of the measured accelerations and force excitation time signals,
2. Computation of the corresponding impulse responses by inverse Fourier transform,
3. Selection of the beginning of the impulse response,
4. Filtering of the useful signals, in the frequency domain, in order to identify at least, the five most prominent vibrational modes,
5. Construction of the stability diagram by applying, in a sequential fashion, the ERA algorithm for increasing model order,
6. Estimation of the minimum order of the realization from the stability diagram,
7. Among the set of modes identified for the chosen order, selection of the physical modes by analysing the identified modeshapes,
8. Reconstruction of all the impulse responses and transfer function using the modes selected,
9. Identification of the modal amplitude coefficients by pseudo-inversion to fit the measured impulse responses,
10. Comparison of the synthesized impulse responses and transfer functions with the ones measured at the sensor locations,
11. Storage of the identified modal parameters.

A software application was developed to facilitate the analysis of the measurement data. However, as for most modal identification problems, one notes that the proposed procedure still involve the user interaction at two stages: for the determination of the minimum order of the model and for the distinction between structural modes and computational modes. To assist the user's

judgment, the convergence of the modal parameters values obtained for successive orders has been analyzed and displayed on the stability diagram and also other modal indicators, namely the MIF and CMI [10], have been implemented. Moreover, modes with unphysical modal parameters values are directly deleted to make clearer the stability diagram and hence facilitating its interpretation.

5. RESULTS FOR THE WITLOCKX CARILLON

The southern tower of the Mafra National Palace contains one of the rare carillon built by W. Witlockx which is still in condition to be played. However, in 1986, the original 46-bell carillon was restored and expanded up to 53 bells by The Royal Eijsbouts. On one hand, this enabled the carillon to be heard after a long period of inactivity. But, on the other hand, from the musicological stand point, it has ruined its historical significance since most of the original bells were re-tuned following the destructive process of material removal. Consequently, today, only 12 bells remain from the original work of Witlockx. Since the original motivation of this work stems from musicological interests, only those 12 historical bells were analyzed by using the previously described modal identification technique. When we turn to investigate the sounding qualities of the entire musical instrument, we make use of the modal frequencies measured by Lehr in 1984 [6], prior to the instrument re-tuning, noting that they correspond to mean frequency values of the degenerate frequencies since Lehr did not account for the modal doublets in the definition of each partial. By doing so, it provides us a manner of studying the musical properties of the carillon as originally built by Witlockx. Obviously, care was previously paid to compare the values measured by Lehr and the ones extracted by our technique in a preliminary work. The obtained results undoubtedly gave confidence in the values collected by Lehr as well as the ones obtained by our experimental procedure.

5.1 Bell internal tuning: tuning of the musical partials of individual bell

As a first step in assessing the overall tuning of the carillon, the proper tuning of each individual bell was analyzed in a systematic manner. For each bell, we identified the modal frequency and modal damping values of the most prominent modes-doublets, as well as their shapes in the horizontal plane of the rim. Despite the numerous works devoted to individual bells [2, 11], it can be said that the determination of all the modal parameters together is rarely found in the literature although they all constitute relevant information to bell founders and researchers. Actually, the clarity and pitch of the bell are directly related to the relationships between the bell partial frequencies, while its decay time varies according to the modal frequency and modal damping value of each partial. Also, the difference between the double modal frequencies enables to characterize the presence of beats in the sound which is known to influence the pitch perception and sound quality of the bell. Finally, the modeshapes can be analyzed as

a tool to detect inaccuracies in bell casting which may be relevant for the case of historical bells.

As already explained, results of our modal identification comprise the 12 historical bells casted by Witlockx pertaining to notes between G_0 - B_1 . As one can notice, some tones are missing in this series because today, some bells are laying on scaffolds, preventing any modal measurement (4 bells in this interval). Figure 2 presents one example of impulse response function and corresponding transfer function stemming from measurements.

To illustrate the well-behaviour of the ERA-based modal identification technique, Figure 3 represents the synthesized impulse responses computed using two different numbers of modes, compared with their respective (band-filtered) measurements. When 26 modes are used, one observes a satisfactory comparison between the two curves, giving confidence in the modal parameter identified. Selecting a lower number of modes for the mathematical model results in larger difference due to the filtering effect. To validate our identification procedure, a detailed analysis of the partials of a laboratory-bell was achieved before performing the in-field work and our identified modal frequencies were compared with the theoretical values obtained by a Finite Element Method analysis. Actually, after a perfect match of the first mode, it was observed that the differences in frequency between FEM and experimental identifications was always less than 2% for the following four relevant partials.

The performance of the technique to identify the near-identical modal-pairs is illustrated in Figure 4 where the identified modes in the horizontal plane of the most prominent partials are displayed for the case of a bell with written note A_1 . Modeshapes are normalised such as their maximum amplitude is equal to unity, and both real (pink) and imaginary (yellow) parts are represented. The dotted lines display the undeformed shape while dots stands for the accelerometer locations. Our identifications result in the classical vibrational patterns of bells and clearly distinguish the two modal families as highlighted by the phase difference. Modeshapes are essentially real.

From the identified modal frequencies, one can then analyze the internal tuning of each bell forming the carillon. Since it is the overall relative tuning between the lower partials which is relevant for a bell to be well-tuned, it must be understood that internal tuning deviation is independent of the partial used as reference. For us, such information is used as a mean to characterize the practical know-how of the bell founder at the time of the carillon construction. However, the choice of the reference partial becomes particularly important when restoring a complete carillon, and such internal tuning charts can be useful to highlight the nature of the corrections to execute to the bell, internally. It gives the frequency changes necessary to correct the bell tuning. As a result of our identification, Figure 5 shows the internal tuning of the first six historical bell cast by Witlockx, with respect to their nominal, expressed in cents. As seen, the bell associated to the played tone B_0 appears particularly well-tuned. Note

that there is no doubt on the relation between the played bell and the emitted note since musical notes are written on the inner part of the bells. Finally, the amount of warble for each partial can also be deduced from the frequency difference between the modal-doublets components as illustrated for the carillon bells in Figure 6. Beats between doublets are very slow, indicating bells with small symmetry imperfections.

Decay rate computed from the identified modal damping values are displayed in Figure 7, as well as the corresponding modal frequencies. Although identified modal damping values are in the range 0.01-0.04% (and almost identical for the two modes of a pair), note that each partial decays with its own rate which undoubtedly contributes to the complex pitch recognition of bell sounds. Finally, Figure 8 displays the decay time of each partial of the carrillon for the 12 historical bells studied in this work.

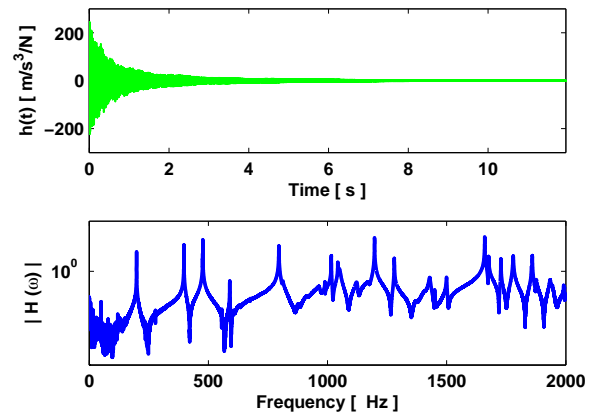


Figure 2: Measured impulse response function (up) and corresponding transfer function (bottom) for the bell with written note A_1 . Excitation and response were both at location 1.

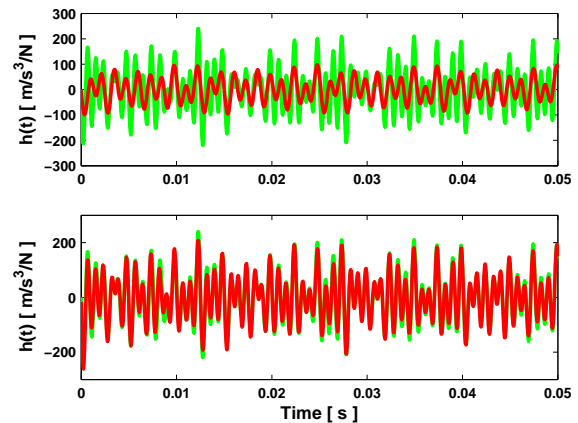


Figure 3: Details of the measured (green) and reconstructed (red) impulse response functions for two different model sizes. Up: 10 modes. Bottom: 26 modes.

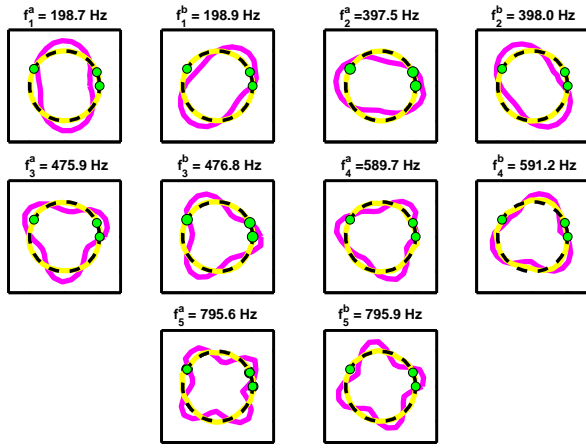


Figure 4: Experimentally identified modeshapes of the most prominent partials for one bell casted by Witlockx (written note A_1). Pink: real part; yellow: imaginary part; —: undeformed shape; green dots: sensor locations.

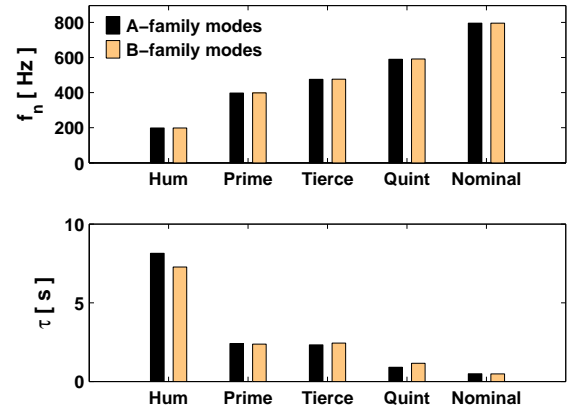


Figure 7: Identified modal frequencies and computed decay time for one bell casted by Witlockx (written note A_1). Modal parameters for the two families of modes are presented.

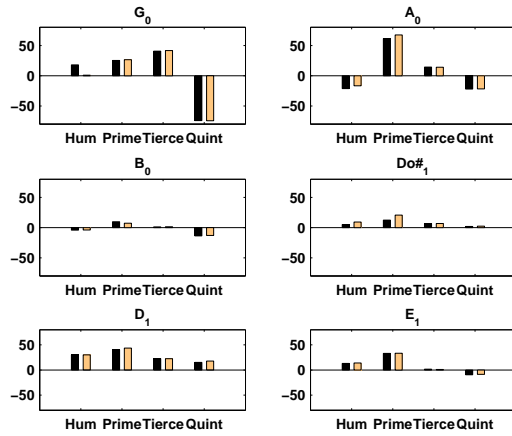


Figure 5: Internal tuning for the first 6 historical bells casted by Witlockx. Reference partial is Nominal and used frequency ratios are such that: $[0.5, 1, 1.2, 1.5]/2$. Values for the two modal families are presented for each partial.

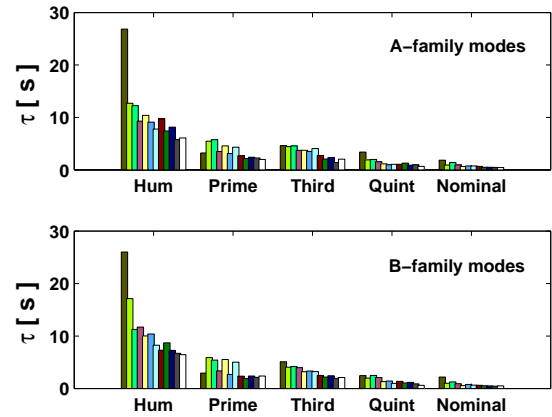


Figure 8: Decay time computed for the 12 historical bells casted by W. Witlockx, ordered in musical bell partial. Values for the two modal families are presented (up and bottom).

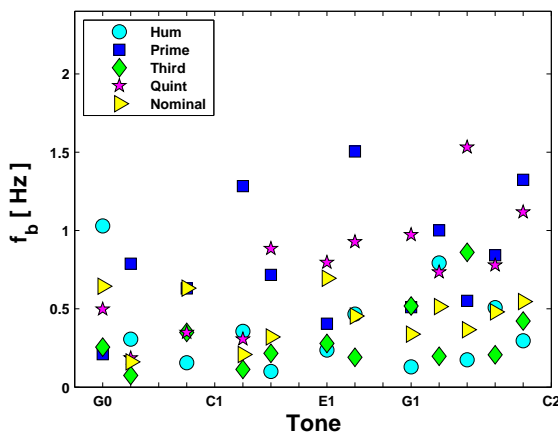


Figure 6: Computed beating frequencies between frequency component of each bell partial, for the 12 historical bells casted by Witlockx.

5.2 External tuning: reference pitch and musical temperament of the carillon

To answer to part of some questions relevant in music history studies, a method is now proposed for inferring, from the collection of bell modal frequencies, the reference pitch and temperament of the musical instrument. Obviously, since a carillon is an instrument composed of a set of tuned bells in order to execute both melodies and chords, its sound qualities not only depend on the internal tuning of its individual bell - important to sound pure and clear as already mentioned - but also on the relative tuning between bells which must follow a given musical scale. From the historical stand point, this aspect represents a question of utmost importance for who is interested in how music was performed in a place and at a time, in the past.

To that end, we devised several related optimization strategies. The basic idea is to minimize an error

between the identified modal frequencies of the collection of bells and the theoretical frequencies of each assumed musical scale. By testing sequentially various plausible temperaments, various reference pitches can be identified and an error, which quantifies the overall detuning of the instrument, can be obtained. Finally, the tuning properties of the instrument are deduced from the reference pitch/temperament combination which results in the smallest error. Mathematically, the optimization problem consists in minimizing a norm of the error \mathcal{E} defined by:

$$\mathcal{E} = \{\mathcal{F}_{mes}\} - F_{ref} \{\mathcal{S}_t\} \quad (3)$$

where $\{\mathcal{F}_{mes}\}$ is a vector of the identified modal frequencies of the set of bells, $\{\mathcal{S}_t\}$ is a vector containing the musical intervals of a given temperament and F_{ref} is the reference pitch of the carillon which is adjusted by best-fit during the optimization process. The identification is hence linear and one useful aspect of Eq.(3) is the possibility to specify in the vector $\{\mathcal{F}_{mes}\}$ which bell and which musical partial of the bell are accounted for the definition of the carillon tuning. Actually, computations can be perverted by several less-than-perfect bells and there is only few information available on how carillons were tuned in the 18th. Furthermore, it is strongly plausible that some bell founders, depending on their experience, did favor the tuning of few partials only. The proposed technique can thus moderately discard the bell founder practice, since it enables different optimizations to be performed, with respect to one given partial only or by including all the partials of the bells in the optimization procedure. In all the cases, the best fit is provided by pseudo-inversion of formulation (3).

Several historical musical temperaments were considered in this study. Based on historical and performance considerations, it appears that either the meantone or the just temperaments are strong candidates for the case of the Mafra carillons. One specificity of these two tuning systems is highlighted by the presence of different wolf intervals associated to fifths with a different interval than all of the others. Distinctions between these two temperaments then reside on the widths of the fifths but also on the size of other intervals, mostly the major and minor thirds. Furthermore, it exists various ways of defining the meantone temperament according to how the syntonic comma is divided. It hence imposes between meantone temperaments slight differences in the width of the fifths which are used to generate the musical scale. The equal temperament, which is the standard tuning scale for today music, has also been considered. Table 1 presents the values of the intervals between each of the 12 notes as defined for the musical temperaments considered in this study.

Tuning deviations computed from the equal, just and 1/4 comma meantone temperaments are shown in Figure 9, for each bell musical partial and over the entire musical range of the instrument. Optimizations were performed accounting simultaneously for the five most important partials and by considering bells producing the notes from G_0 to E_3 (shown between vertical red lines in Figure 9),

for which bell partials appear quite accurately in tune. By analysing the dispersion of the tuning deviation along the musical scale, Figure 9 confirms the use of either the meantone or just temperament by the founder, as expected. To make clearer such a statement, Figure 10 displays the standard deviations of the tuning deviations computed accounting for all the individual bell tuning deviations. As a strong result, it appears from Figure 10 that no preferential tuning of the carillon with respect to the tested target temperaments was achieved. Actually, errors in tuning are larger than differences between temperaments. Interestingly, Figure 10 also confirms one actual difficulty of bell tuning. As can be seen, if the partials Hum and Nominal are in general satisfactorily well-tuned over all the musical range of the instrument, the tuning of the Quints are globally less precise. Such aspect is certainly related to the difficulty of tuning this partial as attested by the tuning functions of individual partials given in the literature [12]. They particularly highlight the impossibility to enlarge this interval by traditional tuning techniques. Consequently, bell founders usually casted their Quints always somewhat to high.

Table 2 presents the computed reference pitch - corresponding to the A=440 Hz for today music - as a function of both the assumed temperament and the partial considered for the optimization procedure. Interestingly, one can observe that all the computations result in near-identical values for the basic pitch, around 396 Hz, except for the just temperament which gives a slight lower value. It can therefore be stated that the Witlockx carillon is approximately tuned 180 cents lower than the actual reference pitch (440 Hz) currently used in classical music. This value has now to be confronted to the standard pitch used in the 18th century in the peninsula iberica, knowing that access to such information is difficult in general. Further work has to be done in that direction by musicologists.

The data shown in Figure 9 can also be splitted into separate plots to display the tuning deviations of each bell partial. A useful aspect of these plots resides in the possibility of examining how the carillon was tuned by the founder during his work. By doing so, one can also have an idea of the practice and skills developed by the bell founder. From the results in Figure 11, one can see that Witlockx apparently did favor the tuning of the bell partials Hum and Nominal for tuning his carillon. Of even most importance, a close examination of the tuning deviations over all the musical range of the carillon reveals that Witlockx only achieved an accurate tuning of one single partial, the Hum tone. Actually, if both Hum and Nominal partials present small deviations compared to the optimal computed pitch for notes in the range G_0 to E_3 , large tuning dispersion are highlighted for the Nominal above the note E_3 . It might be an indication that Witlockx did define the Hum tone as the reference bell partial for tuning its carillon. An other plausible interpretation is that Witlockx only attempted a good tuning of the Hum tone for these notes being somehow aware of the decreasing sensivity of the ear toward very high frequencies. Finally, it would

	Equal	Just	Meantones				
			1/4	1/3	1/5	1/6	2/9
C	1	1	1	1	1	1	1
C#	1.0595	1.0413	1.044907	1.0374	1.0494	1.0525	1.0474
D	1.1225	1.1251	1.118034	1.1157	1.1194	1.1203	1.1188
E \flat	1.1892	1.1996	1.196279	1.2	1.1941	1.1926	1.1951
E	1.2599	1.2498	1.25	1.2448	1.2531	1.2552	1.2517
F	1.3348	1.3325	1.337481	1.3389	1.3367	1.3361	1.3370
F#	1.4142	1.4061	1.397542	1.3889	1.4027	1.4063	1.4005
G	1.4983	1.5	1.495349	1.4938	1.4963	1.4969	1.4959
G#	1.5874	1.5619	1.5625	1.5496	1.5703	1.5755	1.5668
A	1.6818	1.6663	1.671851	1.6667	1.6750	1.6770	1.6736
B \flat	1.7818	1.7994	1.788854	1.7925	1.7866	1.7852	1.7876
B	1.8877	1.8747	1.869186	1.8595	1.8750	1.8789	1.8724
C	2	2	2	2	2	2	2

Table 1: Definition of the musical scale considered in this work.

	Equal	Just	Meantone				
			1/4 comma	1/3 comma	1/5 comma	1/6 comma	2/9 comma
Hum	395.7	393.3	395.2	394.9	395.4	395.5	395.3
Prime	398.3	395.8	397.7	397.4	397.9	398.0	397.8
Third	401.5	395.6	398.6	397.1	399.5	400.1	399.1
Quint	392.1	389.3	392.4	392.5	392.3	392.3	392.3
Nominal	397.8	395.3	397.2	397.5	397.4	397.5	397.3
All	396.9	393.9	396.2	395.8	396.5	396.6	396.4

Table 2: Southern tower, Witlockx carillon. Computed reference pitch for each tested musical scales. Optimization was performed by considering either a single partial (Hum, Prime, Third, Quint or Nominal) or all of them (All).

be interested to relate the discontinuity in tuning errors observed above the note E₃ for the higher musical partials, to a possible change in the profiles of the corresponding bells.

6. CONCLUSIONS

This work aimed at analysing the tuning of the Mafra carillons, which form the largest surviving 18th century carillons in Europe. We developed suitable multi-reference identification techniques to extract the bell vibration modes from experiments, and devised optimal strategies to infer the tuning status of complete carillons with respect to historical tempered scales. The modal identification technique succeeded in distinguishing the modal pairs defining a musical partial, and therefore enabled a detailed analysis of the bell sounding properties by examination of its modal properties. Implementation of optimization techniques was successful to estimate, in an objective manner, the reference pitch and temperament of one of the Mafra carillons. Overall, the described procedure can be useful to answer to questions of both acoustical and musicological concerns.

Acknowledgments

The authors thank the director of the Mafra National Palace M. Pereira for his invaluable collaboration, as well as J. Soeiro de Carvalho and J.N. Cordeniz for helpful discussions on the musicological and historical aspects of the carillons. We also acknowledge support from the Fundação para a Ciência e Tecnologia (FCT) under grant research project PTDC/EAT-MMU/104255/2008.

7. REFERENCES

- [1] L. Gama, *Os Carrilhões de Mafra: Palácio Nacional de Mafra*. Lisboa: Direcção-Geral dos Edifícios e Monumentos Nacionais, 1989.
- [2] R. Perrin, T. Charnley, and J. Depont, "Normal modes of the modern english church bell," *Journal of Sound and Vibration*, vol. 90, pp. 29–49, 1983.
- [3] T. D. Rossing and R. Perrin, "Vibrations of bells," *Applied Acoustics*, vol. 20, pp. 41–70, 1987.
- [4] A. Lehr, *The Designing of Swinging Bells and Carillon Bells in the Past and Present*. Astén: Athanasius Kircher Foundation, 1987.
- [5] P. Roozen-Kroon, "Structural optimization of bells," Ph.D. dissertation, Technical University of Eindhoven, Eindhoven, 1992.
- [6] A. Lehr, *De twee klokkenspelen op het nationaal paleis te Mafra*. Astén: Athanasius Kircher Foundation, 1984.
- [7] X. Boutillon and B. David, "Assessing tuning and damping of historical carillon bells and their changes through restoration," *Applied Acoustics*, pp. 901–910, 2002.
- [8] A. Lehr, "The removal of warbles or beats in the sound of a bell," *Acta Acustica*, vol. 86, pp. 550–556, 1997.
- [9] J. Juang, *Applied System Identification*. New Jersey: PTR Prentice-Hall, Inc., 1994.
- [10] R. S. Pappa, K. Elliott, and A. Schenk, "A consistent-mode indicator for the eigensystem realization algorithm," NASA, Tech. Rep., 1992.
- [11] T. D. Rossing, *Acoustics of Bells*. Stroudsburg, Pennsylvania: Van Nostrand Reinhold, 1984.
- [12] E. W. v. Heuven, *Acoustical measurements on church-bells and carillons*. Gebroeders Van Cleef, 1949.

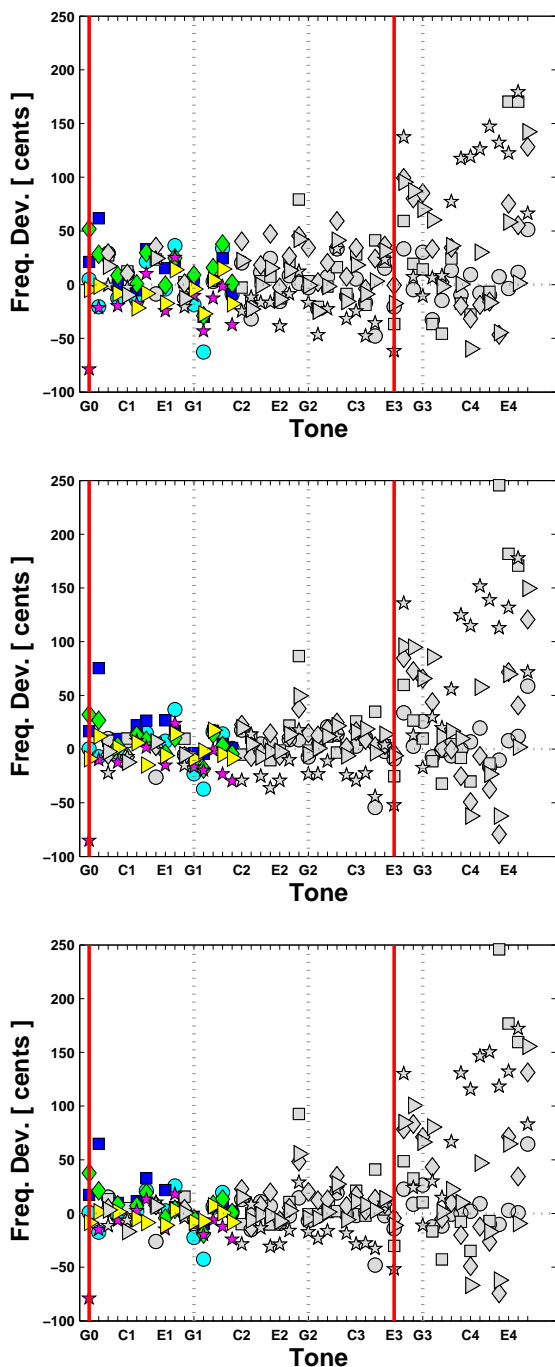


Figure 9: Deviation from historical musical temperaments. Equal (up), just (middle) and 1/4-comma meantone (bottom) temperaments. Reference pitch computed are respectively 396.9 Hz, 393.9 Hz and 396.3 Hz. Legend is the same as in Figure 6. Grey marks stand for bells measured by Lehr [6].

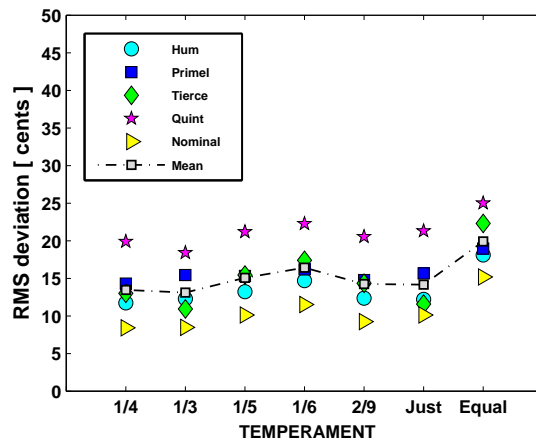


Figure 10: Standard deviations of the tuning errors of the Witlockx carillon with respect to the tested historical temperaments.

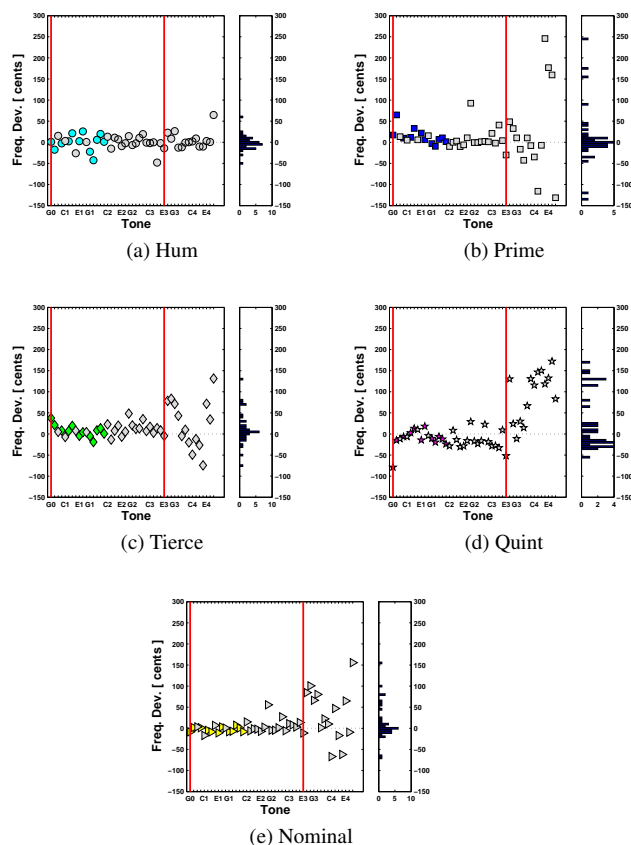


Figure 11: Deviation of individual partials from the quarter-comma meantone temperament. Reference pitch computed is 396 Hz. Grey marks stand for bell measured by Lehr [6].

Nonlinear vibrations of steelpan: analysis of mode coupling in view of modal sound synthesis.

Mélodie Monteil

UME, ENSTA-ParisTech

melodie.monteil@ensta-paristech.fr

Cyril Touzé

UME, ENSTA-ParisTech

cyril.touze@ensta-paristech.fr

Olivier Thomas

ENSAM-LILLE

olivier.thomas@ensam.eu

ABSTRACT

Steelpan are musical percussions made from steel barrels. During the manufacturing, the metal is stretched and bended, to produce a set of thin shells that are the different notes of the instrument. In normal playing, each note is struck, and the sound reveals some nonlinear characteristics which give its peculiar tone to the instrument. In this paper, an experimental approach is first presented in order to show the complex dynamics existing in steelpan's vibrations. Then two models, based on typical modal interactions, are proposed to quantify these nonlinearities. Finally, one of them is observed in free oscillations simulations, in order to compare the internal resonance model to the steelpan vibrations behaviour in normal playing. The aim is to identify the important modes participating in the vibrations in view of building reduced-order models for modal sound synthesis.

1. INTRODUCTION

Steelpan belong to a musical instruments family coming from the island of Trinidad and Tobago. They are usually played in steelbands (see Fig. 1(a)), that are orchestras composed of steelpan covering a range of several octaves.

A steelpan is a tuned percussion, built from cylindrical steel barrels subjected to several stages of metal forming that stretch and bend the structure. The top of the barrel is pressed, hammered, punched and burnt in order to obtain a sort of main thin bowl within which convex substructures are formed. Each convex dome corresponds to a musical note, which natural frequency is precisely tuned according to harmonic relationships (f , $2f$, 3 (or 4) f , ...). Usually, this instrument is played by striking each note with a stick covered by a piece of rubber, as shown on the Fig. 1(b). When a note is stroke with a stick, vibration amplitudes are such that geometric nonlinearities cannot be neglected, and are recognized as a key feature for explaining the peculiar tone of the steelpan [1]. This nonlinearity combined with harmonic relationships, due to steelpan tuning, can activate internal resonances creating strong energy transfers between eigenmodes.

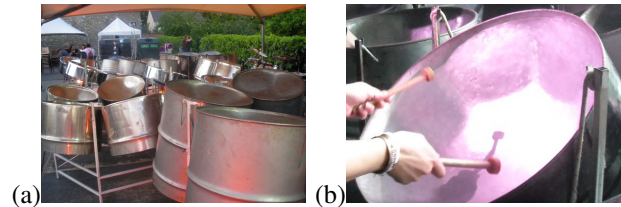


Figure 1. Pictures of steelpan : (a) Steelpan of a typical steelband. (b) Steelpan in normal playing.

Vibrations modelling have been proposed in a series of papers by Achong et al. [2]. In these works, the steelpan is considered as a nonlinear system of oscillators, and energy transfers between normal modes of vibration are highlighted. Rossing et al. [3, 4] have performed modal analyses by holographic techniques to observe modal interactions between harmonically tuned notes. More recently, numerical modal analyses with the finite element methods have been proposed [5]. The steelpan sound radiation has been recently addressed experimentally by Copeland et al. [6]. Finally, some metallurgical issues during the steelpan making have been considered in [7].

In this contribution, we propose a refined analysis of modal couplings and energy exchanges occurring in nonlinear vibrations of steelpan. A modal analysis first reveals the appearance of pairs of modes from the second harmonic, for each note of the pan. These pairs of modes, having nearly equal eigenfrequencies, are interpreted as a consequence of the localization of the vibration into the notes area. Secondly, experiments in forced vibrations allow to reveal the complex nature of the energy exchanges between the modes, that are excited for very small levels of vibration amplitudes. Simple original models including a 1:2:2 and a 1:2:4 resonance are then fitted to the experiments, showing that: (i) mode pairing substantially complexifies the dynamical behaviours and favours the appearance of unstable regimes, (ii) few details of the resonance curves are missed by those 3-dofs models, indicating that even at very small amplitudes of vibrations, a complex dynamics involving more than 3 modes is at hand.

Those findings are used to derive oscillator models for sound synthesis. The first results with three dofs show that the main features (energy transfers and envelope modulation) are recovered.

2. STEELPAN TUNING AND LINEAR ANALYSIS

In steelpan making, a fine tuning of the notes already formed is performed at the end of the building process. Most of the time, the steelpan maker tunes the first three harmonics of each note [8]. He begins with the fundamental frequency (the pitch) by modelling the center of the note. Then, he focuses on the overtones. He adjusts the frequency of two upper harmonics (partials) by modelling the periphery of the note area, as it is shown of Fig. 2(a). This procedure can be easily understood by considering that each harmonic of a given note is associated to a particular vibration mode shape, as it will be explained in the following.

Modal analysis is usually used to characterize the linear behaviour of a structure by identifying eigenfrequencies, mode shapes and modal damping coefficients. A home-made non-contact coil/magnet exciter is used to excite the steelpan at a given point. The equivalent point force is estimated by recording the current intensity in the coil [9]. The steelpan vibratory response, in velocity, is measured with a laser vibrometer.

The steelpan, shown on the figure 2(b), is a right barrel of a double second (middle-high frequency steelpan). It is composed of 19 precisely tuned notes, distributed on three concentric circles, the lower notes being on the outer circle. Previous studies have shown that the vibration remains confined around the solicitation area [3]. Hence, the scan is more particularly focused on G3 (of fundamental frequency f_1) and its harmonically tuned neighbours G4 ($2f_1$) and G5 ($4f_1$).

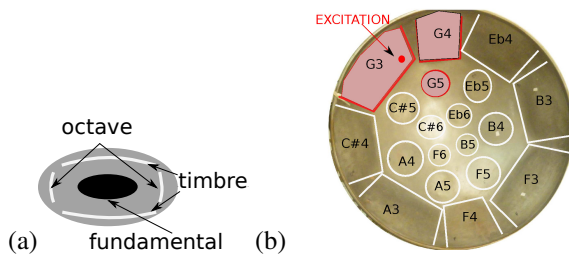


Figure 2. (a) Steelpan tuning. (b) Modal analysis of the steelpan used for the experiment (right barrel of a double second) excited on G3.

Fig. 3 shows the transfer function, measured on the excitation point, in the frequency range [0, 1700] Hz, and the associated mode shapes of the structure. One can see that the first three modes are perfectly tuned like f_1 , $2f_1$ and $4f_1$, while the fourth and the fifth departs a little from the perfect harmonic relationship, and are slightly shifted from the exact $6f_1$ and $8f_1$ relation.

More precisely, around f_1 , the modal shape of the structure is focused on the excited note only. Around $2f_1$, a double peak is clearly visible indicating that the mode is degenerate with two mode shapes around the same frequency. The first one has for eigenfrequency $f_2 = 390$ Hz and is composed of the second vibration pattern of the G3 note together with the fundamental vibration mode of the G4

note. The second one has its eigenfrequency at $f_3 = 397.8$ Hz and its mode shape is similar except the fact that the pattern on the G4 note is out of phase. Finally the measurement reveals also that at $4f_1$, two degenerate modes are also at hand, with eigenfrequencies $f_4 = 789.5$ Hz and $f_5 = 799.3$ Hz, and companion mode shapes.

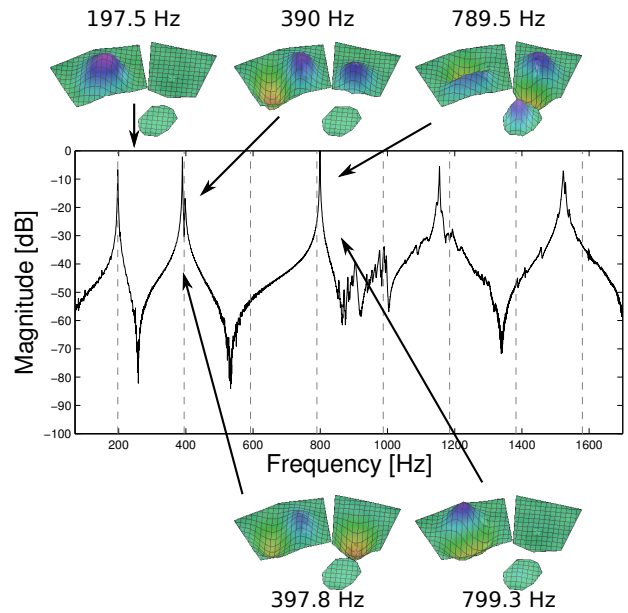


Figure 3. Modal analysis of the steelpan excited on the note G3. FRF measurement and associated mode shapes.

This modal analysis reveals that in steelpan vibrations, modes appear by parts from the second partial. This is a consequence of the fact that vibrations are strongly localized into notes areas, and is a classical feature in mode localization that has been observed for instance in simple beam systems [10].

Forced vibrations at higher force amplitudes will now be detailed to depict how energy is transferred between these modes.

3. FORCED OSCILLATIONS

The previous linear analysis shows that the first frequencies of the steelpan note are organized as a 1:2:2:4:4 sequence. In a nonlinear regime, and in the case of thin curved structures, harmonic relationships create strong modal interactions via internal resonances [11]. In order to investigate the nonlinear behaviour of the steelpan, the same home-made coil/magnet exciter is used, but higher force amplitudes are applied. The coil/magnet exciter has been thoroughly analyzed for calibration in [12], where it has been shown that for the vibration amplitudes encountered, harmonic distortion is less than 1%, ensuring a clean and reproducible harmonic excitation. An external sinusoidal current is applied with an external excitation frequency around the first eigenfrequency ($f_{dr} \simeq f_1$). The vibratory response of one point of the excited note is measured. The first three harmonics (at f_1 , $2f_1$ and $4f_1$) of the response

are precisely measured. Then, two analytical models of internal resonance relationships are proposed to fit the experiment and identify some modal interactions in steelpan vibrations. More details can be found in [13, 14].

3.1 Experiments

In Fig. 4 the nonlinear response of the steelpan excited around 197.5 Hz with $I = 5A$ is shown. The harmonics 1, 2 and 4 of the recorded displacement are shown, they are denoted respectively by w_{H1} , w_{H2} and w_{H4} . A strong coupling between these harmonics, oscillating at f_1 , $f_2 \simeq 2f_1$ and $f_3 \simeq 4f_1$, is revealed. Markers are inserted into the figures to precisely locate, in frequency, the different eigenfrequencies of the system.

First, a 1:2:2 internal resonance is observed through the complex behaviour of the measured amplitude at twice the excitation frequency. The energy transferred gives a large amplitude of vibration to the second (390 Hz) and the third eigenmode (789.5 Hz). Secondly, coupling with the fourth harmonic is also observed where w_{H4} seems to be slaved to w_{H2} due to the similar global shapes of the two corresponding amplitudes. One can assume that a 1:2:2:4 resonance is here activated.

A strong peak of amplitude exists around $f_{dr} \simeq 192$ Hz. At this frequency, the first mode, directly excited, reaches 0.06mm, the second harmonic 0.08mm and the fourth one 0.003mm. Considering the thickness of a typical steelpan initial barrel around 1 mm, one can conclude that geometric nonlinear effects are exhibited for vibration amplitudes of 1/16 times the thickness.

Then, a quasiperiodic regim, around 193 Hz, reveals a complex dynamics of the system. This implies that there is a difference between the forward and the backward excitation. A jump phenomenon is observed at 192.6 Hz.

Compared to the theory [14, 15], the shapes of the solutions are slightly different than a 1:2 or a 1:2:2 internal resonance, probably because the cubic nonlinearity is not anymore negligible. An other remark on the effects of cubic nonlinearities can be the shift of the curves to the left part of the graph, compared to the linear initial frequencies. The maximum of each harmonic response is observed around $f_{dr} \simeq 192$ Hz.

3.2 3-dofs internal resonance models fitting

The complicated dynamics exhibited by the forced vibrations is now modeled by two simple systems involving three internally resonant modes, so as to highlight the most salient features of the dynamics of the steelpan. The two models involve either a 1:2:2, or a 1:2:4 eigenfrequency relationship. For these 3-dofs models, analytical solutions are accessible via multiple scales analysis [14]. Model fitting to experimental measurements will thus shed new light on the identification of nonlinear coupling coefficients as well as energy transfers.

Considering the steelpan as a curved thin structure, the transverse displacement w can be discretised by expanding

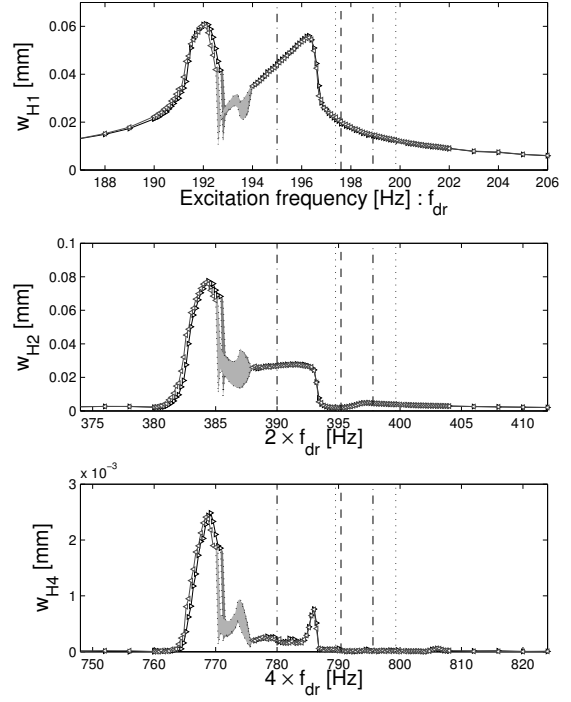


Figure 4. Frequency response curves $I = 5A$, excitation frequency f_{dr} in the vicinity of the first mode. Perlins: harmonics 1, 2 and 4 of the measured displacement w . Forward (black line: \triangleright) and backward (grey line: \triangleleft) frequency sweeps. Linear eigenfrequency markers f_1 (---); $2f_1$ (- · -) and $4f_1$ (· · ·).

w on the linear modes basis $\{\Phi_k\}$ as

$$w(\mathbf{x}, t) = \sum_{k=1}^N \Phi(\mathbf{x}) q_k(t), \quad (1)$$

where q_k is the modal coordinate and Φ_k is the mode shape amplitude value at the spatial point \mathbf{x} .

Now two 3-dofs simple models are proposed, considering the most important coupling found in the modal analysis, and thus displaying respectively a 1:2:2 and the 1:2:4 internal resonance. The 1:2:2 internal resonance considers three modes which eigenfrequencies are tuned such that $\omega_3 \simeq \omega_2 \simeq 2\omega_1$. The associated normal form of the amplitude of the three corresponding modes reads :

$$\begin{cases} \ddot{q}_1 + \omega_1^2 q_1 = \varepsilon [-2\mu_1 \dot{q}_1 - \alpha_1 q_1 q_2 - \alpha_2 q_1 q_3 + F_1 \cos \Omega t], \\ \ddot{q}_2 + \omega_2^2 q_2 = \varepsilon [-2\mu_2 \dot{q}_2 - \alpha_3 q_1^2], \\ \ddot{q}_3 + \omega_3^2 q_3 = \varepsilon [-2\mu_3 \dot{q}_3 - \alpha_4 q_1^2], \end{cases} \quad (2)$$

where $\omega_k = 2\pi f_k$ denotes the angular frequency of the mode k and μ_k its damping coefficient. $\Omega = 2\pi f_{dr}$ is the external frequency and F_1 the amplitude of the external force. In first approximation, only quadratic nonlinear coupling terms α_i are kept. Only four of them are present. They correspond to the resonant terms. All other possible nonlinear quadratic terms have no importance for the global dynamics and can be cancelled. Nonlinear coefficients, damping terms and external forcing are assumed to

be small as compared to the linear oscillatory part, and thus are scaled by $\epsilon \ll 1$. According to the multiple scales development, the solutions of the dynamical system Eqs. (2) are obtained as:

$$\begin{cases} q_1 = a_1 \cos(\Omega t + \varphi_1), \\ q_2 = a_2 \cos(2\Omega t + \varphi_2), \\ q_3 = a_3 \cos(2\Omega t + \varphi_3), \end{cases} \quad (3)$$

where a_k and φ_k are amplitudes and phases of the solution q_k , respectively. In term of transverse displacement, the equation (1) leads to $w(\mathbf{x}, t) = w_1 \cos(\Omega t + \varphi_1) + w_2 \cos(2\Omega t + \gamma_2)$, where $w_1 = \Phi_1 a_1$ and where w_2 and γ_2 are combinations of $\Phi_2, \Phi_3, a_2, a_3, \varphi_2$ and φ_3 .

The 1:2:4 internal resonance, is a 3-dofs nonlinear dynamical system constructed with the following eigenfrequency relationship: $\omega_3 \simeq 2\omega_2 \simeq 4\omega_1$. It reads:

$$\begin{cases} \ddot{q}_1 + \omega_1^2 q_1 = \varepsilon [-2\mu_1 \dot{q}_1 - \alpha_5 q_1 q_2 + F_1 \cos \Omega t], \\ \ddot{q}_2 + \omega_2^2 q_2 = \varepsilon [-2\mu_2 \dot{q}_2 - \alpha_6 q_1^2 - \alpha_7 q_2 q_3], \\ \ddot{q}_3 + \omega_3^2 q_3 = \varepsilon [-2\mu_3 \dot{q}_3 - \alpha_8 q_2^2], \end{cases} \quad (4)$$

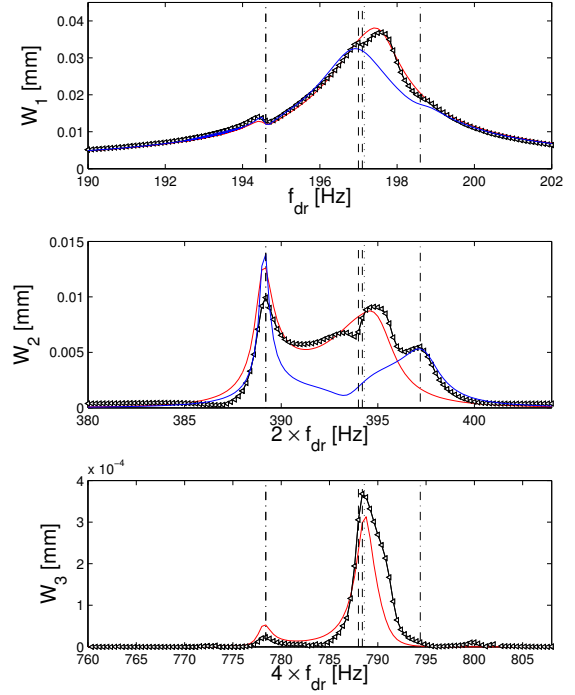
where $\alpha_{5,6,7,8}$ are the four new nonlinear coupling coefficients associated to the 1:2:4 internal resonance system. Solutions of Eqs. (4) are:

$$\begin{cases} q_1 = a_1 \cos(\Omega t + \varphi_1), \\ q_2 = a_2 \cos(2\Omega t + \varphi_2), \\ q_3 = a_3 \cos(4\Omega t + \varphi_3). \end{cases} \quad (5)$$

In that case, two upper-harmonics, oscillating at 2Ω and 4Ω , simultaneously appear. The transverse displacement is $w(\mathbf{x}, t) = w_1 \cos(\Omega t + \varphi_1) + w_2 \cos(2\Omega t + \varphi_2) + w_3 \cos(4\Omega t + \varphi_3)$, where $w_1 = \Phi_1 a_1$, $w_2 = \Phi_2 a_2$ and $w_3 = \Phi_3 a_3$.

Fig. 5 presents the 1:2:2 model fitting (in red color), first compared to the experiment, then compared to a 1:2:4 internal resonance model fitting. Experimental fitting, presented in Fig. 5, determines the value of each nonlinear coefficient while the linear parameters (ω_k, μ_k) are identified by modal analysis. The amplitude of the external current is $I = 2A$. The two 3-dofs models allow to locally identify some modal interactions in steelpan vibrations. Each of them locally fits parts of the experimental curves. Markers of linear frequencies indicate the mode activated by nonlinear coupling. This fitting shows that simple models can be used to enhance the comprehension of the complicated dynamics experimentally observed. More complicated scenarios invoking also the presence of the modes at $6f_1$ and $8f_1$ may also be activated in certain vibratory regimes.

The steelpan is harmonically tuned (both between notes and between harmonics in a note). Therefore, the thinness of the curved structure and the amplitude of sollicitation of the note lead to nonlinear behaviour with modal interactions and energy transfers, consequently. Internal resonance models displaying a 1:2:2 and a 1:2:4 frequency relationships allow to recover the main features of the FRFs,



	Mode 1	Mode 2	Mode3
$F_k = \omega_k / 2\pi$ [Hz]	197.1	389.2	788.6
$\xi_k = \mu_k / \omega_k$ [%]	0.0045	0.001	0.001
$\Phi_k(\mathbf{x})$	0.295	0.339	0.06

Figure 5. Frequency response curves and experimental fitting ($I = 2A$) (black line). Nonlinear coupling coefficients identified: 1:2:2 internal resonance model (blue line) - $\alpha_1 = 0.42$, $\alpha_2 = 0.46$, $\alpha_3 = 0.46$, $\alpha_4 = 0.39$ and 1:2:4 internal resonance model (red line) - $\alpha_5 = 0.24$, $\alpha_6 = 0.34$, $\alpha_7 = 0.88$, $\alpha_8 = 0.53$.

while non-modelled effect appears to be easily interpreted. The most complete model for that case should be a 1:2:2:4:4 one, unfortunately analytical solutions for that problem are not tractable.

4. NONLINEAR MODEL FOR SOUND SYNTHESIS

4.1 Free oscillations measurement

In Fig. 6, two time-frequency representations of free oscillations of the G3 note in normal playing are shown. A weak stroke (see Fig. 6(a)) reveals that most of energy is stored within the first two harmonics. Comparatively, a strong stroke (see Fig. 6(b)) has much more energy on the higher frequencies. It can be noticed that, on these two spectrograms, there is no energy at $F = 3f_1$. For the low frequency range, only the tuned harmonics are excited when the steelpan is played.

Fig. 7 represents the measurements of the Fig. 6 filtered on the harmonics 1, 2 and 4 (denoted H_1, H_2, H_4), showing the evolution with the time of the amplitude of each

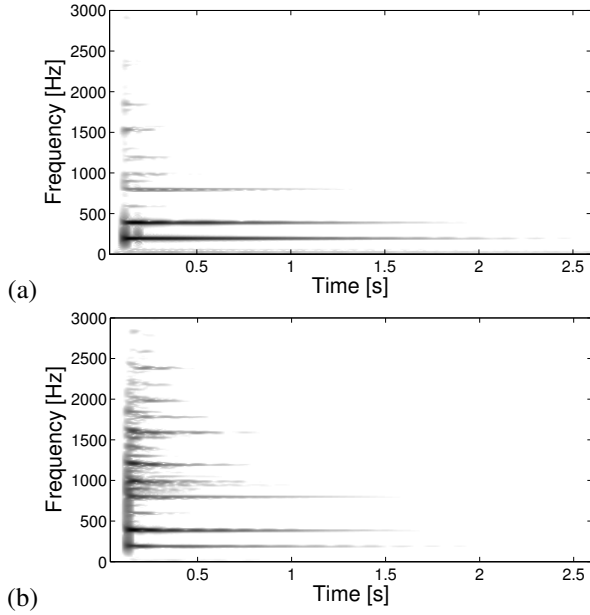


Figure 6. Experimental spectrograms of two free oscillation measurement with two stick-impact intensities. (a) Weak stroke. (b) Strong stroke.

harmonic. First, on the Fig. 7(a), it can be observed that the second harmonic reaches its maximum later than the first one does. Then, the amplitude of the maximum is a little bit higher than the first one. The fourth harmonic presents a very small amplitude. For a strong stroke (see Fig. 7(b)), the amplitude of the second harmonic is much larger than the first one. It can be deduced that energy is transferred from the first mode to the second one, and the intensity of the transfer depends on the intensity of the excitation. The fourth one has much more energy than in the first case. Its amplitude is very close to the first harmonic. Thus, it is shown that initial conditions are very important in the steelpan response.

4.2 Free oscillations simulation

The 1:2:4 adjusted model is now analyzed in free oscillations behaviour. The ordinary differential equation solver (ODE45) of Matlab is used. The linear parameters and the nonlinear coupling coefficients values are deduced from the experimental fitting in forced oscillations (see Fig. 5). An initial condition in displacement is given so as to mimic a delta dirac temporal excitation.

Fig. 8 shows the result of two simulations. The first one (Fig. 8(a)) is for a weak initial condition ($w_0 = 0.15$ mm) and the second one (Fig. 8(b)) is for a larger one ($w_0 = 0.65$ mm). The acceleration of the three transverse displacements (w_1 , w_2 and w_3) are represented versus time. Compared to the filtered measurement (Fig. 7), the global time evolution of a steelpan sound is recovered for the two initial conditions. The right amplitude ratios between the modes are observed. Also, one can remark that the period of the oscillations is qualitatively the same in the experi-

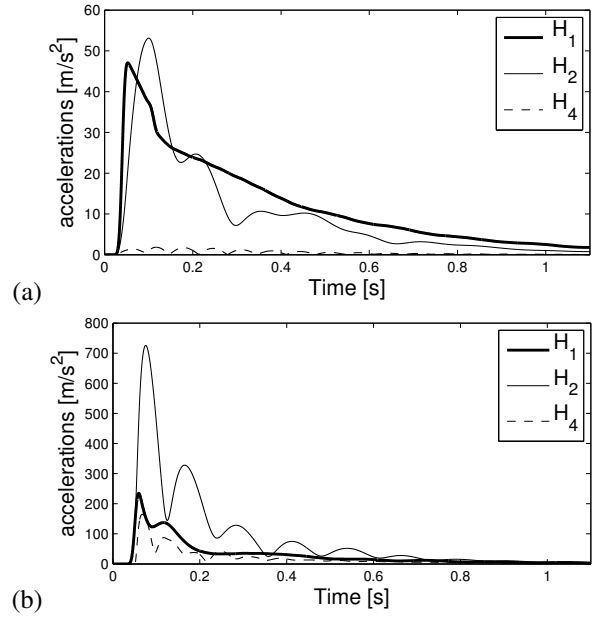


Figure 7. Same measurements than Fig. 6, filtered on harmonics 1, 2 and 4. Evolution of each harmonics with time.

ments and the simulations.

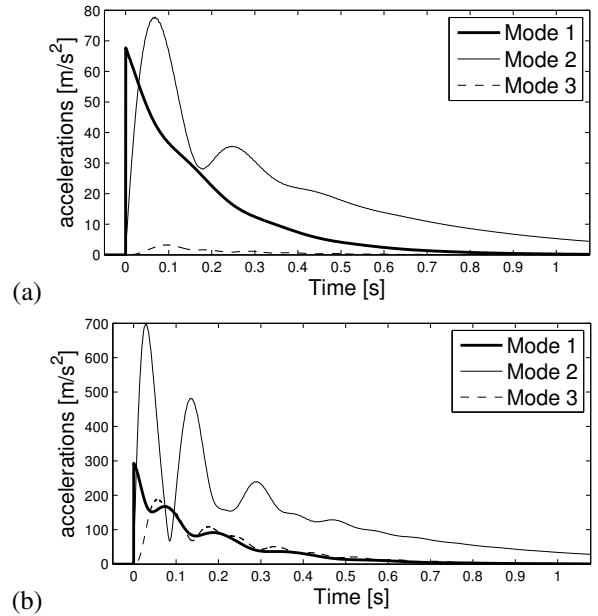


Figure 8. 1:2:4 internal resonance model in free oscillations simulation. Initial conditions in displacement: (a) $w_0 = 0.15$ mm, (b) $w_0 = 0.65$ mm.

These time simulations confirm that with regard to the global dynamics carried by the envelope modulation of the main harmonics, the 3-dofs model allows to recover the most salient nonlinear features, and are thus identified as a key component in steelpan's sound and vibrations.

5. CONCLUSION

Nonlinear vibrations of a steelpan presenting a 1:2:4 modal tuning (f , $2f$, $4f$) have been investigated thanks to a refined modal analysis, measurements of frequency-response curves in forced oscillations and time domain simulations in free, impacted vibrations, using 3-dofs models identified from the forced response.

The main outcomes of the present study are the following

- The modal analysis has clearly evidenced the fact that modes appear by pairs from the second harmonic of each note, a feature that is only scarcely mentioned in the literature. This pairing is interpreted as a consequence of the strong localization of the vibrations into the notes area. Indeed, simpler systems consisting of beams with stiffness imperfections reveals the same degeneracy which is thus generic for structural systems having at least two minima of stiffness [10].
- A consequence of this mode pairing is that the dynamical equations are complexified. Analytically, one can also show that the appearance of 1:2:2 (instead of simpler 1:2) or *e.g.* 1:2:2:4:4 (instead of simpler 1:2:4) internal resonances in a nonlinear systems favour instabilities [14].
- The forced response analysis reveals the complex dynamics of steelpan with appearance of energy exchange and quasiperiodic regimes for very small amplitudes of vibration.
- 3-dofs models that are analytically solvable have been fitted to measured resonance curves, showing that the main features are recovered by considering 1:2:2 and 1:2:4 resonances. However small details in the resonance curves are not fitted, advocating for the fact that the precise dynamics is more complex with energy exchange between pairs of modes and thus solutions comprising *e.g.* 1:2:2:4:4 resonances and even for some amplitudes 1:2:2:4:4:6 and 1:2:2:4:4:8.
- The temporal simulations allow recovering the main features of the energy exchange, once again confirming that the most important part of the dynamics is carried by the identified resonances. However, small details are still missing to the ear for a sound synthesis approach, stating clearly that the higher modes (at 6f and 8f) have to be included in a simple modal model for sound synthesis.

At the conference, sound synthesis will be presented in order to more precisely quantify the importance of the higher modes for the sound production, as well as the impact model.

6. REFERENCES

- [1] A. Achong, "Mode locking on the non-linear notes of the steelpan," *Journal of Sound and Vibration*, vol. 266, pp. 193–197, 2003.
- [2] —, "The steelpan as a system of non-linear mode-localized oscillators, I: Theory, simulations, experiments and bifurcations," *Journal of Sound and Vibration*, vol. 197(4), pp. 471–487, 1996.
- [3] T. D. Rossing, U. J. Hansen, and D. Hampton, "Vibrational mode shapes in Caribbean steelpan. I. Tenor and double second," *Journal of the Acoustical Society of America*, vol. 108(2), pp. 803–812, 2000.
- [4] T. D. Rossing and U. J. Hansen, "Vibrational mode shapes in Caribbean steelpan. II. Cello and bass," *Applied acoustics*, vol. 65, pp. 1233–1247, 2004.
- [5] D. Gay, "Finite Element Modelling of Steelpan Acoustics," in *Acoustic2008*, 2008.
- [6] F. Muddeen and B. Copeland, "Sound radiation from Caribbean steelpan using nearfield acoustical holography," *Journal of the Acoustical Society of America*, vol. 131(2), pp. 1558–1565, 2012.
- [7] L. Murr, E. Esquivel, S. Lawrie, M. Lopez, S. Lair, K. Soto, S. Gaytan, D. Bujanda, R. Kerns, P. Guerrero, and J. Flores, "Fabrication of aluminum, Caribbean-style, musical pan: Metallurgical and acoustical characterization," *Materials Characterization*, vol. 57, pp. 232–243, 2006.
- [8] U. Kronman, *Steel pan tuning: a handbook for steel pan making and tuning*. Musikmuseet, 1992.
- [9] O. Thomas, C. Touzé, and A. Chaigne, "Non-linear vibrations of free-edge thin spherical shells: modal interaction rules and 1:1:2 internal resonance," *Internal Journal of Solids and Structures*, vol. 42, pp. 3339–3373, 2005.
- [10] A. Luongo, "Mode Localization in Dynamics and Buckling of linear Imperfect Continuous Structures," *Nonlinear Dynamics*, vol. 25, pp. 133–156, 2001.
- [11] A. Nayfeh, *Nonlinear interactions*, J. Wiley and sons, Eds., 2000.
- [12] O. Thomas, C. Touzé, and A. Chaigne, "Asymmetric non-linear forced vibrations of free-edge circular plates. Part 2: experiments," *Journal of Sound and Vibration*, vol. 265, pp. 1075–1101, 2003.
- [13] M. Monteil, C. Touzé, and O. Thomas, "Complicated dynamics exhibited by thin shells displaying numerous internal resonances : application to the steelpan ," in *19th International Congress on Sound and Vibrations (ICSV)*, Vilnius, Lithuania, 8-12 juillet 2012.
- [14] M. Monteil, C. Touzé, O. Thomas, and J. Frelat, "Vibrations non linéaires de steelpan : couplages modaux via la résonance interne 1:2:2," in *20ème congrès français de mécanique (CFM)*, Besançon, France, 28 août - 2 septembre 2011.
- [15] A. Nayfeh and D. Mook, *Nonlinear Oscillations*, I. John Willey & Sons, Ed. Wiley Classics Library, 1979.

TIME-RESOLVED INTERFEROMETRY AND PHASE VOCODER ANALYSIS OF A CARIBBEAN STEELPAN

Andrew Morrison
Joliet Junior College
amorrison@jjc.edu

Daniel Zietlow
Rollins College
daniel.zietlow@cfl.rr.com

Thomas Moore
Rollins College
tmoore@rollins.edu

ABSTRACT

The Caribbean steelpan is one of the most recently developed tuned percussion instruments and has been the subject of much scientific study in recent years. Electronic speckle pattern interferometry (ESPI) is a useful method for characterizing the operating deflection shapes (ODS) and modes of vibration of musical instruments, and previous studies have used time averaged ESPI to characterize the ODS of resonances of the notes on steelpan. Using ESPI in conjunction with a high speed camera, capable of capturing images at rates of several thousand frames per second, allows for time-resolved examinations of transient motion of the note strike and the subsequent decay. High speed ESPI movies of note strikes of a low tenor steelpan were acquired while simultaneously recording the sound of the strike. The comparison of the time-resolved interferometry data with the analysis of the sound recordings allows for insights into the evolution of coupling between note areas.

1. INTRODUCTION

The Caribbean steelpan has developed from its humble roots as a folk instrument to become the only acoustic tuned musical instrument invented in the 20th century. The rapid evolution of the instrument was driven initially by skilled tuners in Trinidad who experimented with different techniques for constructing the steelpan. The construction and tuning of each pan is still done primarily by hand, as no manufacturing process can produce a pan of sufficient quality for learning to play or performance. The tuning process involves many steps, and is more fully described elsewhere in the literature.

The basic construction process starting from a new oil barrel involves sinking the bottom surface, cutting the pan skirt to length, scribing the note layout, rough tuning, firing the pan, and final tuning. Each pan tuner is free to use whatever note layout desired, but most tuners use a layout of notes familiar to the majority of players around the world.

In recent years there have been many studies related to the acoustics of the steelpan family. Kronman and Jansson de-

scribed characteristic sound spectra of steelpan strikes [1]. Rossing et. al have described the vibration of various members of the steelpan family, including both mode shapes of notes of the steelpan and the vibration of the whole pan [2]. Copeland et al. have characterized radiation patterns of the steel pan using a sound intensity probe and nearfield acoustical methods. [3–5] There have also been efforts to understand the material properties of steelpans. Murr et al. described the metallurgy of a steelpan constructed using a hydroforming process [6].

A main factor that contributes to the unique timbre of the instrument is that the individual notes of the instrument are embedded into the same piece of steel. Therefore, the striking of a single note causes the entire surface of the instrument to begin vibrating. Energy from the struck note is transferred to sympathetically vibrating notes, and notes with resonant frequencies close to the frequencies of the struck note will tend to vibrate with the largest amplitudes of all the sympathetic vibrations. One of the most common note layouts used by steelpan tuners is commonly referred to as the “fourths and fifths” layout, where each note is neighbored by its musical fourth and musical fifth. For steelpan drums having more than one ring of notes, the inner ring notes are usually the octave of the outer ring notes they neighbor. This layout puts notes of the two most consonant intervals (the octave and the perfect fifth) next to each other. The notes in this layout form a “circle of fifths,” which is a concept familiar to many musicians. The vibrations of these notes are strongly-coupled, although an understanding of the coupling between notes has yet to be explained. The present work is a first attempt to characterize the vibration of strongly-coupled notes.

2. EXPERIMENTAL METHOD

The steelpan used in this investigation was a low tenor steelpan tuned by Bertrand Kellman. This steelpan is tuned with the now standard fourths and fifths layout of the notes, shown in Figure 1. The approximate dimensions of the steelpan are 58.5 cm in diameter, a 22.5 cm bowl depth, and a 18 cm long skirt. A more complete description of this pan has been reported previously [3].

In ordinary playing conditions, the steelpan is suspended on hooks attached to a stand by nylon cords threaded through holes on opposite sides of the instrument. The tenor steelpan hangs from the stand inclined slightly towards the performer. A performer traditionally uses a wooden shaft between 7-8.5 inches in length wrapped with a strip of natural rubber to form a mallet for striking the steelpan. The rub-

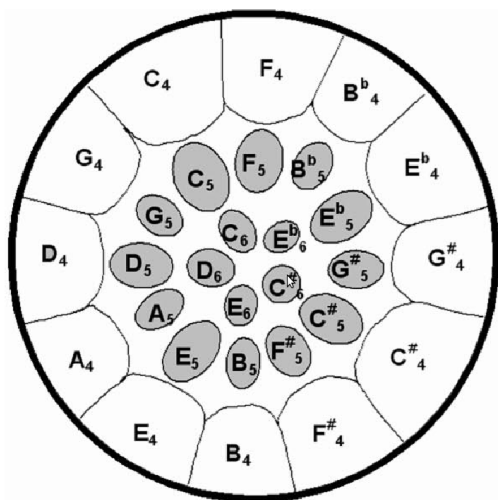


Figure 1: Layout of the Kellman tenor steelpan.

ber is often 1/32" thick in 1" wide strips wrapped several times around the shaft to form a mallet head approximately double the diameter of the shaft. Performers wrap their own mallets, and although there are traditional methods of wrapping the mallets, each performer is free to wrap the mallet to their own liking.

In our experiments, the instrument was removed from the stand and securely mounted near the locations of the mounting holes on an optical table. In our experiments the drum was struck with a small hand-held steel ball-tipped driver at levels well below even the softest amplitudes used in performance. The elasticity of the rubber-tipped mallets likely has an interaction during the strike with the pan that we wanted to avoid. Our goal was to understand the motion of the pan itself and the sound produced after a strike without concern for an interaction with the mallet.

A recording of each drum strike was made with a sensitive microphone or pair of matched microphones triggered to coincide with the high-speed photography as described below. The audio data was captured at a sample rate of 44100 Hz with a program written in LabView. The audio sample was then converted to an audio file for analysis in the SNDAN software package [7].

The suite of audio analysis tools called SNDAN includes a tool called pvan which is a phase vocoder harmonic spectrum analysis tool. Phase vocoder analysis allows for characterization of the time varying nature of the harmonic amplitudes of audio signals with harmonic content [7, 8]. The analysis is most useful if there is a large degree of harmonicity in the sound. A skilled steelpan tuner works hard to make sure the first three resonances of each of the notes is harmonically tuned. The use of rubber tipped mallets tends to damp the high frequency resonances above these lowest tuned harmonics, leaving mostly only harmonic content in the radiated sound. Thus, using the phase vocoder analysis is a powerful and appropriate tool for studying the time dependence of the first three resonances of a steelpan note.

To investigate the vibration of coupled notes a low tenor steelpan was located in the image plane of an electronic

speckle pattern interferometer (ESPI) located inside a hemi-anechoic chamber. The ESPI is a two-beam optical interferometer which has been more completely described elsewhere [9]. The ESPI was equipped with a high speed CCD camera which was set to capture images at a rate of 15,000 frames per second. Frames are captured by the camera, stored on a computer, and then image subtraction is performed using a program written in LabView. The image was framed to capture the most strongly-coupled notes. In particular, the coupling between notes separated by an octave and notes separated by a perfect fifth were investigated. Individual notes were struck with a hand-held steel ball-tipped driver from the rear of the pan so that the striking instrument was not in the image frame. The sensitivity of the ESPI system is such that the vibration amplitude must be relatively small. The strikes on the pan, while clearly audible inside the anechoic chamber were significantly lower than what would be used in typical playing conditions.

3. RESULTS

3.1 Phase vocoder analysis

Kronman and Jansson made sonogram measurements which indicated that the higher harmonics in the spectrum do not reach a maximum until 25-30 ms after the strike [1]. Their work indicated that the higher the frequency of the harmonic, the longer the delay after the strike before reaching maximum amplitude.

Using the phase vocoder analysis routine in the SNDAN package, amplitudes of the first three harmonics of the sound samples were obtained. Figure 2 shows amplitude plots of the first three harmonics of the steelpan struck on the C_4 note. A pair of closely matched microphones was used: one microphone was located directly above the C_4 note as shown in 2a; the other microphone was located directly above the C_5 note as shown in 2b. The position of the microphones blocked the view of the camera, so no ESPI data was recorded.

The amplitude of the fundamental frequency rises quickly after the strike occurs before starting to decay. The amplitude of the second harmonic reaches a maximum approximately 50 ms after the strike occurs. The amplitude of the third harmonic reaches its maximum value approximately 75 ms after the strike occurs. Although the rise in amplitude of the second and third harmonic was 2-3 times slower than that reported previously, the order in which the harmonic reach their maximum amplitude are consistent with those reported by Kronman and Jansson [1].

The measurements shown in Figure 2a indicate that a greater portion of the acoustic energy in the fundamental frequency of the sound is due to the vibration of the struck note area: C_4 . Figure 2b shows that a greater portion of the acoustic energy in the second harmonic of the sound is due to the vibration of the C_5 note area.

All other audio recordings were made with a single microphone located 1 meter from the center of the pan and out of the view of the camera. Recordings were made simultaneously with ESPI imaging of note strikes.

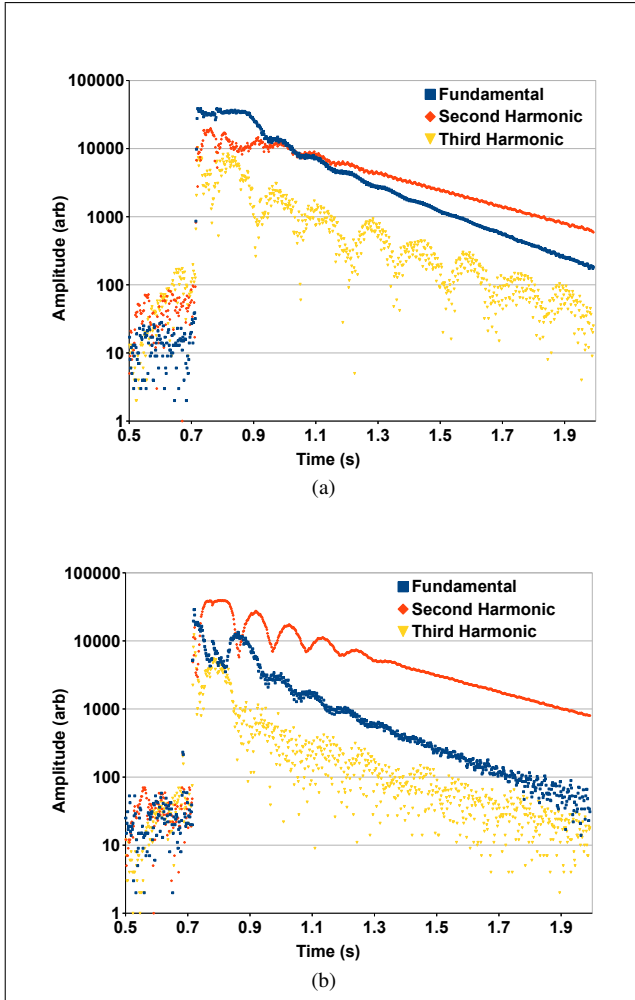


Figure 2: Amplitude measurements of first three harmonics after striking C_4 note: (2a) Microphone above strike note. (2b) Microphone above C_5 note.

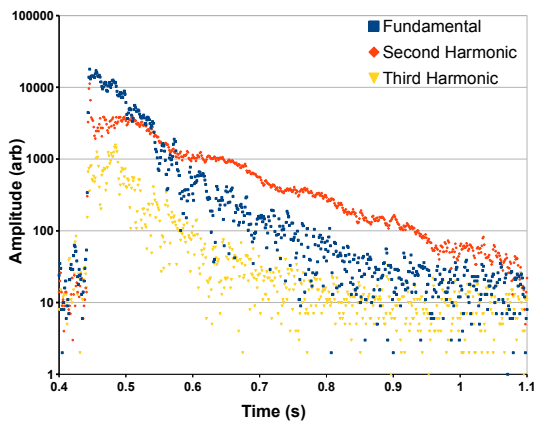


Figure 3: Amplitude measurements of the first three harmonics of the G_4^\sharp note.

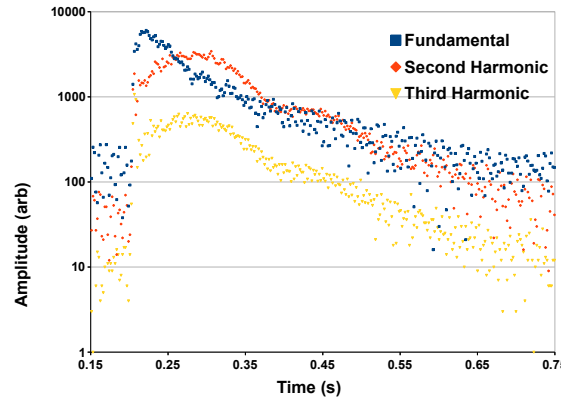


Figure 4: Amplitude measurements of the first three harmonics of the C_4 note.

A plot of amplitude measurements for the first three harmonics of the G_4^\sharp note is shown in Figure 3. It was observed that the second harmonic in this case reached a maximum 25 ms after the strike. However, the third harmonic reached a maximum only 15 ms after the strike, contrary to the behavior observed by Kronman and Jansson.

Analysis of an audio recording of the strike of the C_4 note is shown in Figure 4. The amplitude of the second harmonic and third harmonic was observed to rise in sync with each other in this case. Both harmonics reached a maximum amplitude approximately 80 ms after the onset of the strike. After measuring many of the amplitude-time profiles of the notes on this steelpan it was impossible to determine any clear pattern as to which conditions would lead to a predictable order of rise of the amplitude of the second and third harmonics. Given the nonlinear behavior of the steelpan, it is not a surprise that no clear pattern has yet emerged.

3.2 High speed interferograms

To understand how the sound generated by the steelpan evolves in time, we used ESPI with a high speed camera to capture frames immediately after the strike. The ESPI frame capture was done at the same time that the audio recordings were made.

The initial propagation of the traveling wave due to the strike on the note is shown in Figure 5. In this figure, the camera was zoomed in to image only a small part of the steelpan near the location of the strike. The frame rate of the camera was set to record just over 15,000 frames per second. Within 1/1000th of a second, the initial wavefront has reached the rim of the pan as seen on the left side of the interferograms. After this point, reflections off of the rim make tracking the initial wave impossible.

The high frame capture rate means that a large number of images are stored for analysis not in real time, but after the strike has occurred. The processed images are then combined into a movie that can be played back for interpretation. Only a small subset of the images analyzed are able to be included here. The frames shown in Figure 5, for example, represent, on average, less than 1% of the images

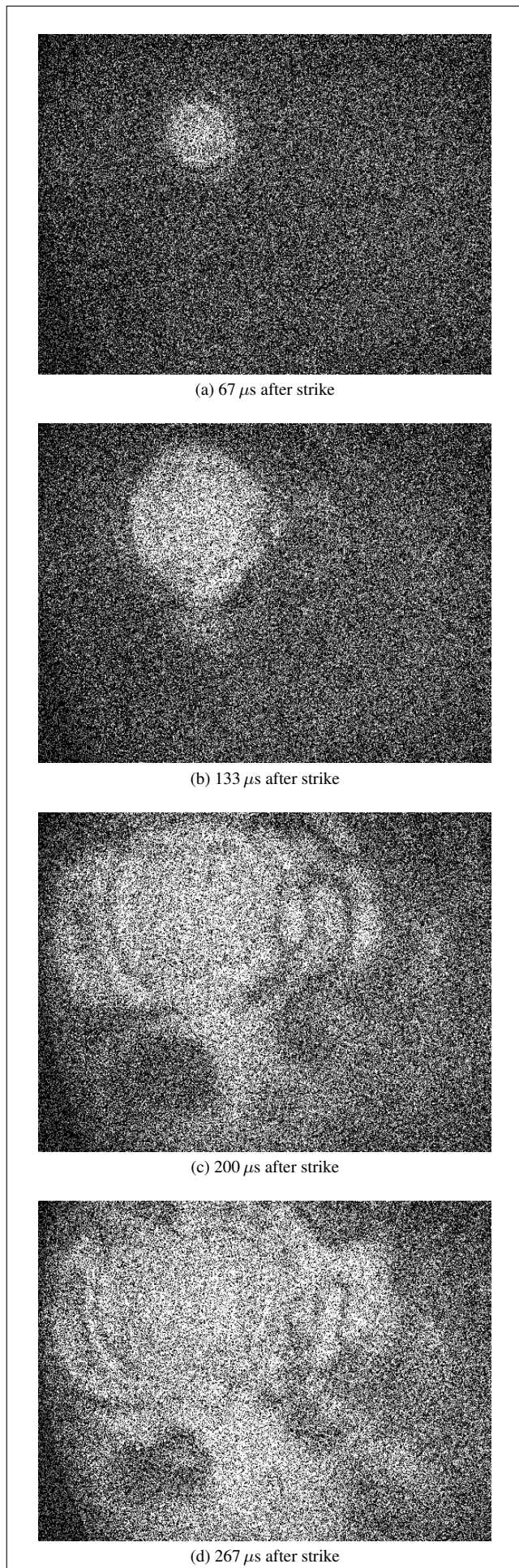


Figure 5: First four frames after the strike of the pan showing initial propagation of the wave in the region near a single note next to the rim.

needed to show the motion of the pan as the harmonics reach their maximum amplitude.

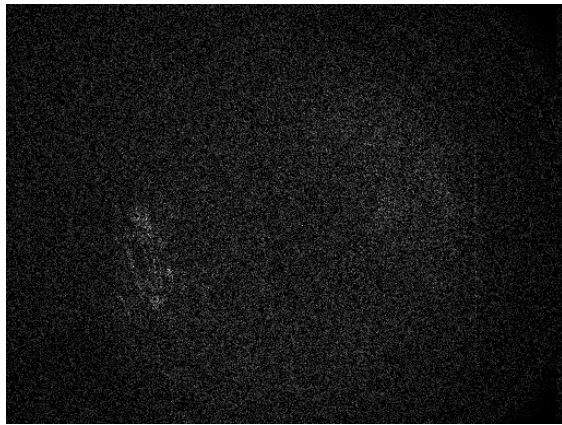
High speed ESPI data were acquired for the strikes of the G_4^\sharp and C_4 notes discussed in Section 3.1 and shown in Figures 3 and 4. Four selected frames after striking the G_4^\sharp are shown in Figure 6. In these images, the entire steelpan is visible. For reference, the steelpan has been rotated such that the strike note G_4^\sharp is located at the 8 o'clock position in the figure.

The location of the strike note is shown in Figure 6a. After approximately 1 ms as shown in Figure 6b a few notes (C_5^\sharp and E_5^\flat) are vibrating sympathetically at low amplitudes, and the strike note is vibrating at a larger, albeit still relatively small, amplitude. After approximately 6 ms as shown in Figure 6c the strike note has a large vibration amplitude, the G_5^\sharp note is growing in amplitude, and several other notes are vibrating sympathetically at low amplitudes. Figure 6d shows the steelpan 20 ms after the strike when the G_5^\sharp note has reached a large vibration amplitude. There is a good correlation between the ESPI images in Figure 6a and the phase vocoder analysis for this strike as shown in Figure 3.

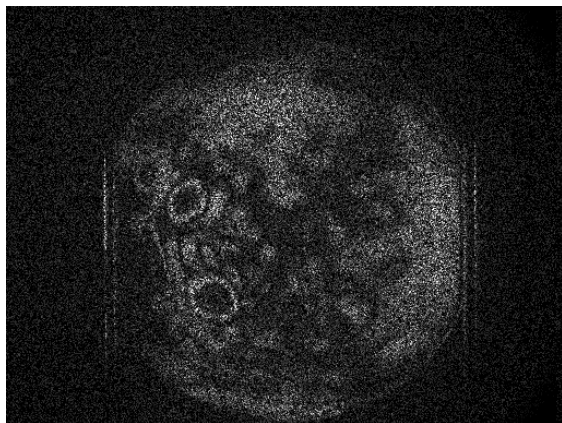
A detailed view of the part of the pan near the strike note is shown in Figure 7. The image shows the notes which are in motion near the C_4 note approximately 48 ms after the strike occurs. At this time the amplitude of the second harmonic is increasing, and is approximately equal to the amplitude of the fundamental. In the ESPI image, the C_4 note (located at the 2 o'clock position) shows a strong second harmonic component, as indicated by the nodal line running parallel to the rim. The C_5 note (located slightly to the right of center) is clearly vibrating at its fundamental frequency as evidenced by the single anti-nodal region. The C_6 note (located slightly to the left of center) is also vibrating at its fundamental frequency. Comparing the sequence of ESPI images with the phase vocoder analysis as shown in Figure 4 shows clear correlations between the vibration of notes near the C_4 and the amplitude of the harmonics in the audio recording.

4. CONCLUSIONS

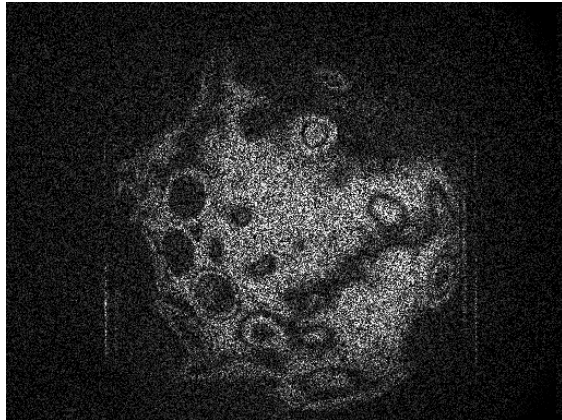
The Caribbean steelpan is a highly nonlinear instrument. Understanding how the pan surface vibrates after striking a note and being able to correlate the vibrations to the spectrum produced by the strike has been notoriously difficult. In this article, we have described the use of time-resolved electronic speckle pattern interferometry and phase vocoder analysis to study the Caribbean steelpan. We determined that delays for the onset of higher harmonics may be 2-3 times longer than previously reported in the literature. We showed that the general behavior of higher harmonics having later onset previously reported is not true for all instruments or notes on an instrument. By making audio recordings of note strikes while simultaneously using ESPI equipped with a high speed camera we are able to correlate the motion of the steelpan surface with the phase vocoder analysis of the amplitude versus time measurements of the first three harmonics.



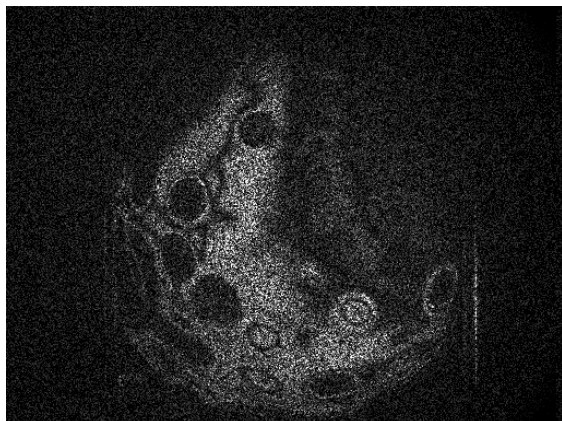
(a) 67 μ s after the strike



(b) 1 ms after the strike



(c) 6 ms after the strike



(d) 20 ms after the strike

Figure 6: Four processed frames showing the full pan after the strike of the G_4^\sharp note.

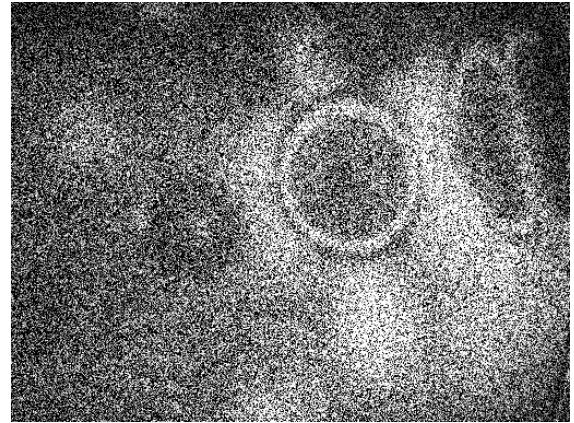


Figure 7: Detailed view of the region around the C_4 note. Image was captured 48 ms after the strike of the C_4 note. In this image, the strike note, the C_5 note, and the C_6 note are shown to be vibrating.

Acknowledgments

This work was supported in part by the Joliet Junior College Faculty Development fund and NSF grant #PHY-0964783.

5. REFERENCES

- [1] U. Kronman and E. Jansson, "Acoustics of the Steelpan; Tone Generation and TuningAn Introductory Study of Methods for Measurements," *Speech, Music and Hearing Quarterly Progress and Status Report*, vol. 29, no. 4, pp. 59–73, 1988. [Online]. Available: http://www.speech.kth.se/prod/publications/files/qpst/1988/1988_29_4.059-073.pdf
- [2] T. D. Rossing, U. J. Hansen, and D. S. Hampton, "Vibrational mode shapes in Caribbean steelpans. I. Tenor and double second," *J. Acoust. Soc. Am.*, vol. 108, no. 2, pp. 803–812, Aug. 2000.
- [3] B. Copeland, A. Morrison, T. D. Rossing, and Andrew Morrison, "Sound radiation from {Caribbean} steelpans," *J. Acoust. Soc. Am.*, vol. 117, no. 1, pp. 375–383, Jan. 2005.
- [4] B. Copeland, Andrew Morrison, and T. D. Rossing, "Sound radiation from {Caribbean} steelpans (abstract only)," *J. Acoust. Soc. Am.*, vol. 110, p. 2673, 2001.
- [5] F. Muddeen and B. Copeland, "Sound radiation from Caribbean steelpans using nearfield acoustical holography," *The Journal of the Acoustical Society of America*, vol. 131, no. 2, pp. 1558–65, Feb. 2012. [Online]. Available: <http://link.aip.org/link/?JASMAN/131/1558/1>
- [6] L. E. Murr and E. al., "Metallurgical and acoustical comparisons for a brass pan with a {Caribbean} steel pan standard," *J. Materials Science*, vol. 39, pp. 4139–4155, 2004.
- [7] J. W. Beauchamp, "Unix workstation software for analysis, graphics, modification, and synthesis of musical

sound,” in *Audio Engineering Society Convention 94*, 1993.

- [8] —, “Analysis and synthesis of musical instrument sounds,” in *Analysis, Synthesis, and Perception of Musical Sounds*. Springer, 2007, pp. 1–89.
- [9] T. R. Moore, “A simple design for an electronic speckle pattern interferometer,” *American Journal of Physics*, vol. 72, pp. 1380–1384, 2004.

Numerical Experiments with Non-linear Double Membrane Drums

Alberto Torin

Acoustics and Audio Group
University of Edinburgh
A.Torin@sms.ed.ac.uk

Stefan Bilbao

Acoustics and Audio Group
University of Edinburgh
sbilbao@staffmail.ed.ac.uk

ABSTRACT

Drums with two membranes are very common; snare drums, tom-toms and bass drums can be found in Western music, but there are examples in Eastern music, as well (the Indian mridangam, the Japanese taiko, etc.) These instruments can have considerable physical dimensions; bass drum heads can sometimes reach a radius of half a meter. Given the size, the low tension at which the membranes are generally tuned and the amplitude of vibrations, it is unlikely that a linear model could capture the most salient features of the sound of these instruments.

Pitch glide effects and an increase of high-frequency energy has been observed at high excitation amplitudes for the bass drum [1] and more recently for tom-toms [2]. Similar phenomena have been observed, for example, in strings [3] and plates [4], and are often related to the presence of non-linearities in the system.

In this paper we present a finite difference time domain model of double membrane drums (i.e., tom-toms and bass drums) with air coupling and with non-linear terms (due to von Kármán) in the equations of motion for the two membranes. Some of the computational difficulties stemming from this particular choice will be discussed. Simulation results and sound examples will be presented.

1. INTRODUCTION

Percussion instruments, and drums in particular, have received growing attention in recent years. Several experimental and numerical studies have been performed on kettledrums [5,6] and snare drums [7,8], but research on bass drums remains quite scarce [1].

Numerical simulations of musical instruments based on physical models are becoming an attractive approach both in acoustical studies and sound synthesis. Among the various techniques adopted, finite difference time domain methods appear to be a versatile tool to tackle these kinds of problems rigorously, especially in the presence of several coupled components, possibly with non-linear interactions [9]. The main disadvantage of such methods, namely, their heavy computational complexity, is becoming less urgent due to the increasing availability of parallel hardware, such as graphical processing units (GPGPUs). In fact, for

algorithms whose update for the various components can be performed with a high degree of independence, as is often the case for FDTD methods over regular grids, the speed-up that one can obtain with respect to a code written for a CPU can be significant (even tens of times) [10].

Despite these improvements, however, some challenges are still open. The solution of a linear system, for example, remains a difficult task even in the case of FDTD methods, where linear systems to be solved are sparse. Simple methods, like Gaussian elimination, because of their essentially serial nature, are not suited for implementation on a GPU. One of the goals of the ongoing NESS Project [11] at the University of Edinburgh is to devise numerical algorithms for sound synthesis that can eventually be parallelized on GPGPUs. The present paper may also be considered, then, as a first step in the study of parallel implementations of large sparse linear systems. The non-linear drum model described below becomes an ideal test case for these kinds of experiments.

The physical description of the system, which is mainly based on approaches recently applied to the snare drum [9] and to the timpani drum [12], is given in Section 2, and includes the coupling of the drum with the surrounding air field. The main novelty of the present model is the introduction of a von Kármán non-linearity in the membrane equation [13]. In Section 3 a finite difference numerical scheme will be presented which is stable and which dissipates energy strictly. The core of the algorithm is the solution, in the run time loop, of a sparse linear system which must be constructed anew at each time step. Given the size of the system, this results in a computationally very demanding task. In Section 4 we will discuss a possible method to accelerate the computation based on purpose-designed iterative methods. Finally, simulation results will be presented in Section 5.

2. MODEL SYSTEM

The drum model under consideration can be schematically divided into several interacting components: two membranes, a rigid shell connecting them, the acoustic field in which the system is embedded. (See Fig. 1.) Typical bass drums have membranes with radius between 40 cm and 50 cm, with height between 35 cm and 45 cm, while tom-toms are generally smaller [14].

The approach adopted here and the geometry of the system are similar to those used recently in [9], therefore some of the details will be omitted. A finite computational space \mathcal{V} is needed in which the virtual drum can be embedded.

This enclosure, or “box”, is not intended to simulate a real 3D room, but rather to emulate the behaviour of an infinite space by applying absorbing conditions at the boundaries $\partial\mathcal{V}$ (see Sec. 2.3.)

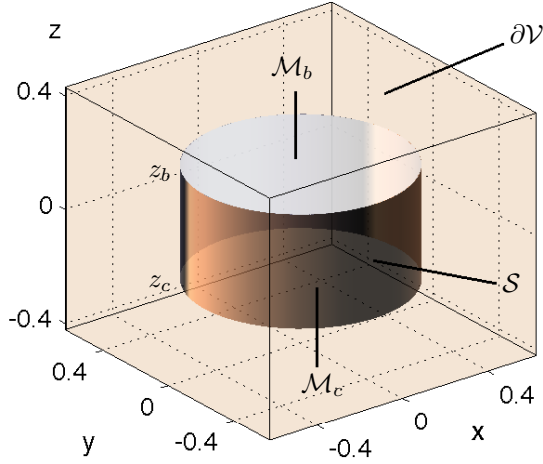


Figure 1. Geometry of the model. Batter membrane \mathcal{M}_b at position z_b , carry membrane \mathcal{M}_c at position z_c , rigid shell \mathcal{S} , computational region \mathcal{V} with boundaries denoted by $\partial\mathcal{V}$.

2.1 Membranes

The most important components of bass drums and tomtoms are two heads, which are responsible for the sound production. The upper one, called the batter head, is usually struck with a mallet, while the lower one, called the carry membrane, is set into motion by the air inside the cavity. They are defined over two circular regions \mathcal{M}_b and \mathcal{M}_c of radius R and at positions $z = z_b$ and $z = z_c$, respectively. As is usual in the literature, we will only consider the displacement $w(x, y, t)$ of the membrane normal to the surface, at position (x, y) and time t .

As pointed out in [1], typical blows on the batter head can cause displacements of several millimeters, at least one order of magnitude bigger, therefore, than the thickness of the membrane itself. It is reasonable to expect non-linear effects to become important in these conditions, similarly to what happens for strings [15].

We can write, now, the equations of motion for the two membranes:

$$\begin{aligned} \frac{\partial^2 w^{(i)}}{\partial t^2} = & c_i^2 \Delta_{2D} w^{(i)} - \kappa_i^2 \Delta_{2D} \Delta_{2D} w^{(i)} + \sigma_i \Delta_{2D} \frac{\partial w^{(i)}}{\partial t} \\ & \frac{1}{\rho_i H_i} (f_i^+ + f_i^-) + \delta_{i,b} \frac{1}{\rho_b H_b} \delta(x_b - x_0, y_b - y_0) f_{exc} \\ & + \frac{1}{\rho_i H_i} \mathcal{L}(w^{(i)}, \Phi^{(i)}). \end{aligned} \quad (1)$$

Here, the index $i = b, c$ refers to the batter or carry membranes respectively. The first line groups the terms for a linear stiff membrane with viscoelastic loss [16]. The differential operators Δ_{2D} and $\Delta_{2D} \Delta_{2D}$ are the Laplacian

and biharmonic operators, with

$$\Delta_{2D} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \quad (2)$$

The membrane speeds c_i and stiffness parameters κ_i are given by:

$$c_i = \sqrt{(T_i / \rho_i H_i)} \quad \kappa_i = \sqrt{E_i H_i^2 / 12 \rho_i (1 - \nu_i^2)}, \quad (3)$$

where (dropping the subscripts) T is the membrane tension per unit of length, in kg/s^2 , ρ is the membrane density, in kg/m^3 , H is the membrane thickness, in m, E is Young's modulus, in $\text{kg/s}^2\text{m}$, and ν is the dimensionless Poisson's ratio. σ is the coefficient governing viscoelastic losses, in m^2/s . All these parameters can in principle be distinct for the two membranes.

The two terms in the second line of (1) take into account external forces acting on the membranes. f^+ and f^- represent the pressure due to the presence of air acting above and below the surface of the membrane, while f_{exc} is a pointwise excitation force, acting on the batter membrane only.

The last term in (1) introduces in the system geometric non-linear effects, that are widely discussed in relation to plate vibration [13]. The action of the non-linear operator \mathcal{L} on two test functions ξ and χ is given by:

$$\mathcal{L}(\xi, \chi) = \frac{\partial^2 \xi}{\partial x^2} \frac{\partial^2 \chi}{\partial y^2} + \frac{\partial^2 \chi}{\partial x^2} \frac{\partial^2 \xi}{\partial y^2} - 2 \frac{\partial^2 \xi}{\partial x \partial y} \frac{\partial^2 \chi}{\partial x \partial y}. \quad (4)$$

$\Phi(x, y, t)$ is the so-called Airy's stress function, and must satisfy the following constraint:

$$\Delta_{2D} \Delta_{2D} \Phi^{(i)} = - \frac{E_i H_i}{2} \mathcal{L}(w^{(i)}, w^{(i)}). \quad (5)$$

(1) and (5) form, then, a set of two coupled equations for each membrane.

Two boundary conditions for both $w^{(i)}$ and $\Phi^{(i)}$ must be supplied; the membranes are assumed to be fixed at the boundary but free to rotate, a condition generally referred to as “simply supported” in the plate literature, while for the Airy's functions a free condition is used [17]:

$$w^{(i)} = 0 = \Delta_{2D} w^{(i)}, \quad \Phi^{(i)} = 0 = \mathbf{n}_{i,ext} \cdot \vec{\nabla}_{2D} \Phi^{(i)}, \quad (6)$$

where $\mathbf{n}_{i,ext}$ is the unit vector normal to the boundary and $\vec{\nabla}_{2D}$ is the gradient.

2.2 Air

Wave propagation in air is modelled by means of a 3D wave equation:

$$\frac{\partial^2 \Psi}{\partial t^2} = c_a^2 \Delta_{3D} \Psi, \quad (7)$$

where c_a is the speed of sound in air (here, 340 m/s), and the 3D Laplacian operator is defined as:

$$\Delta_{3D} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (8)$$

Ψ is a velocity potential related to the pressure p and particle velocity \mathbf{v} by:

$$p = \rho_a \frac{\partial \Psi}{\partial t} \quad \mathbf{v} = -\vec{\nabla}_{3D} \Psi, \quad (9)$$

where ρ_a is the density of air (1.21 kg/m³) and $\vec{\nabla}_{3D}$ is the 3D gradient.

2.3 Absorbing Boundary Conditions

As mentioned before, it is necessary to house the virtual drum in a finite computational region \mathcal{V} . However, since the purpose is that of simulating (ideally!) an infinite space, we must impose absorbing boundary conditions over the walls $\partial\mathcal{V}$ of this box.

One possible way of achieving this is by means of perfectly matched layer, a standard approach drawn from electromagnetism [18]. In the present case, however, in order to reduce the computational complexity, we chose to adopt Engquist Majda equations [19], a family of absorbing conditions with increasing order of accuracy. In the present work, we will use a first order approximation defined as:

$$\left(\frac{\partial}{\partial t} + c_a \mathbf{n} \cdot \vec{\nabla}_{3D} \right) \Psi = 0, \quad (10)$$

where \mathbf{n} denotes the unit vector normal to the wall and pointing outwards.

2.4 Cavity

The drum cavity is modelled as a perfectly rigid shell \mathcal{S} enclosing the air between the two membranes. It is a cylindrical surface of radius R , between $z = z_c$ and $z = z_b$. Neumann conditions are applied to the velocity potential over the shell surface

$$\mathbf{n}_S \cdot \vec{\nabla}_{3D} \Psi = 0 \quad (11)$$

where \mathbf{n}_S is the unit vector normal to the cylindrical surface. This condition must be imposed both inside and outside the cavity.

2.5 Coupling Conditions

As we anticipated in Sec. 2.1, an explicit expression for f^+ and f^- appearing in (1) must be given. The pressure acting on the membranes must be equal to the pressure of the acoustic field above and below the surfaces. Given the relation between Ψ and p , we can write:

$$f_i^+ = -\rho_a \lim_{z \rightarrow z_i^+} \partial_z \Psi |_{\mathcal{M}_i} \quad f_i^- = \rho_a \lim_{z \rightarrow z_i^-} \partial_z \Psi |_{\mathcal{M}_i} . \quad (12)$$

Another condition must be supplied, namely, that the velocity of the membrane be equal to the velocity of the acoustic particles on either side of the membrane. Mathematically,

$$\partial_t w_i = - \lim_{z \rightarrow z_i^-} \partial_z \Psi |_{\mathcal{M}_i} = - \lim_{z \rightarrow z_i^+} \partial_z \Psi |_{\mathcal{M}_i} . \quad (13)$$

Conditions (12) and (13) hold over the membrane regions \mathcal{M}_b and \mathcal{M}_c .

2.6 Excitation and Output

The excitation mechanism underlying the mallet-membrane interaction has been the subject of several studies. Its inherently non-linear nature is discussed, e.g., in [6]. A finite difference model of such an excitation is indeed possible [20], but is complicated by the interaction with the non-linearity of the membrane. Given the very short duration of the mallet-membrane interaction (on the order of 2-4 ms), and to simplify the implementation, we chose to adopt another approach. Each strike is modelled as a raised cosine impulse in time acting at a single point of the membrane specified by a 2D Dirac function [20]. More realistic models that would allow a finer control of the sound are currently under study.

Output sounds are obtained by sampling the variations of pressure generated at a certain point in the acoustic field by a strike on the batter membrane.

3. FINITE DIFFERENCE SCHEMES

In this section we discuss the numerical implementation of the model presented above using finite difference time domain methods [21].

The system described in the previous section will be approximated over regular Cartesian grids and at discrete time instants, which are integer multiples of a time step k . In sound synthesis simulations, it is customary to choose the sample rate F_s *a priori* for perceptual reasons, and to obtain k consequently as $k = 1/F_s$.

The displacement function $w^{(i)}(x, y, t)$ for each membrane can be approximated by a discrete function $w_{l,m}^{n,(i)}$ defined over a spatial grid of spacing h and time step k , such that:

$$w^{(i)}(x, y, t) \approx w_{l,m}^{n,(i)} \equiv w^{(i)}(lh, mh, nk) \quad (14)$$

for integers l, m and n .

We can now introduce forward, backward and identity shift operators in time, whose action on $w_{l,m}^n$ is given by:

$$e_{t+} w_{l,m}^n = w_{l,m}^{n+1}, \quad e_{t-} w_{l,m}^n = w_{l,m}^{n-1}, \quad 1 w_{l,m}^n = w_{l,m}^n. \quad (15)$$

Analogous relations hold for spatial shift operators; along the x axis, for example, we have:

$$e_{x+} w_{l,m}^n = w_{l+1,m}^n, \quad e_{x-} w_{l,m}^n = w_{l-1,m}^n, \quad 1 w_{l,m}^n = w_{l,m}^n. \quad (16)$$

These operators form the basic units constituting finite difference operators.

The choice of Cartesian vs. polar grids over a circular geometry is dictated by various concerns, discussed in detail in [9], the most important of which is the bandlimited output that is generated by schemes in polar coordinates. Typical polar schemes operating at a sample rate of 44.1 kHz produce outputs bandlimited to approximately 3 kHz.

For computational reasons, it is useful to store 2D grids as column vectors, by taking the values of the array columnwise. (See Fig. 2 for an example.) Difference operators can be written, therefore, as sparse matrices.

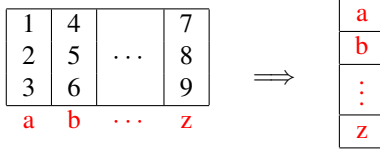


Figure 2. A 2D grid (left) is reshaped into a column vector (right) by placing each column (in red) in a vertical array.

3.1 Membranes

A finite difference scheme for the coupled equations (1) and (5) may be written as:

$$\begin{aligned} \delta_{tt}w^{(i)} &= c_i^2\delta_{2\Delta}w^{(i)} - \kappa_i^2\delta_{2\Delta}\delta_{2\Delta}w^{(i)} + \sigma_i\delta_{t-}\delta_{2\Delta}w^{(i)} \\ &+ \frac{1}{\rho_i H_i} (f_i^+ + f_i^-) + \delta_{i,b} \frac{1}{\rho_b H_b} \delta(l_b - l_0, m_b - m_0) f_{exc} \\ &+ \frac{1}{\rho_i H_i} l(w^{(i)}, \mu_t, \Phi^{(i)}), \end{aligned} \quad (17)$$

$$\delta_{2\Delta}\delta_{2\Delta}\mu_{t-}\Phi^{(i)} = -\frac{EH}{2} l(w^{(i)}, e_{t-}w^{(i)}), \quad (18)$$

where, in order to simplify an already heavy notation, we put $w^{(i)}$ instead of $w_{l,m}^{n,(i)}$. Integers l_0 and m_0 represent the nearest grid point to the continuous excitation point (interpolation may be used, as well.) The explicit actions of the various operators involved in the previous equations can be written in terms of shift operators defined in (15) and (16):

$$\delta_{tt} = \frac{1}{k^2} (e_{t+} + e_{t-} - 2), \quad (19a)$$

$$\delta_{2\Delta} = \frac{1}{h^2} (e_{x+} + e_{x-} + e_{y+} + e_{y-} - 4) \quad (19b)$$

$$\delta_{t-} = \frac{1}{k} (1 - e_{t-}) \quad (19c)$$

$$\mu_{t-} = \frac{1}{2} (1 + e_{t-}) \quad (19d)$$

$$\mu_t = \frac{1}{2} (e_{t+} + e_{t-}). \quad (19e)$$

This particular choice for the discretization of the non-linear terms appearing in (1) and (5) is but one in a family of finite difference schemes for the von Kármán equation [22], and leads to provable energy conservation when the following expression for l is used:

$$\begin{aligned} l(\xi, \chi) &= \delta_{xx}\xi\delta_{yy}\chi + \delta_{yy}\xi\delta_{xx}\chi \\ &- \frac{1}{2} (\delta_{x-,y-}\xi\delta_{x-,y-}\chi + \delta_{x+,y-}\xi\delta_{x+,y-}\chi \\ &+ \delta_{x-,y+}\xi\delta_{x-,y+}\chi + \delta_{x+,y+}\xi\delta_{x+,y+}\chi). \end{aligned} \quad (20)$$

The finite difference version of boundary conditions (6) can be written as follows:

$$w^{(i)} = 0 = \delta_{2\Delta}w^{(i)}, \quad \Phi^{(i)} = 0 = \delta_n\Phi^{(i)} \quad (21)$$

where δ_n represents the non-centred spatial derivative over the boundary.

Coupling conditions f^+ , f^- and f_{exc} in (17) are the discrete counterpart of those in (1), and their explicit expression will be given in Sec. 3.3.

3.2 Acoustic field

The acoustic field $\Psi(x, y, z, t)$ can be approximated by a discrete function $\Psi_{l,m,p}^n$ over a 3D Cartesian grid with spacing h_a . The finite difference discretization of (7) that has been adopted is:

$$\delta_{tt}\Psi = c_a^2\delta_{3\Delta}\Psi, \quad (22)$$

where the 3D Laplacian operator is defined as:

$$\delta_{3\Delta} = \frac{1}{h_a^2} (e_{x\pm} + e_{y\pm} + e_{z\pm} - 6), \quad (23)$$

and $e_{j\pm} = e_{j+} + e_{j-}$ for $j = x, y, z$.

The boundary condition (11) over the drum shell can be implemented with a simple staircase approximation [9]. For a grid point Ψ_{in} inside the cavity with a nearest neighbour Ψ_{out} outside, a modified version $\bar{\delta}_{3\Delta}$ of the 3D Laplacian can be written as:

$$\bar{\delta}_{3\Delta}\Psi_{in} = \delta_{3\Delta}\Psi_{in} + \frac{1}{h_a^2} (-\Psi_{out} + \Psi_{in}). \quad (24)$$

A similar equation holds for Ψ_{out} , with the subscripts exchanged, and it can be easily extended when the nearest neighbours outside or inside are two.

At the boundary of the computational box, Engquist Majda condition (10) can be discretized with centred operators. At $p = 1$, e.g., we can write:

$$\delta_t.\Psi_{l,m,1} - c_a\delta_z.\Psi_{l,m,1} = 0. \quad (25)$$

This equation can be inserted into (22) to obtain an explicit update form for the acoustic field over the boundary [9].

3.3 Coupling Conditions and Excitation

The coupling mechanism described in Sec. 2.5 can be implemented in an easier way if the vertical positions z_i of both membranes are chosen half way between two neighbouring sets of points of the acoustic field with indices p_i^- and p_i^+ , so that $p_i^- = z_i - h_a/2$ and $p_i^+ = z_i + h_a/2$. This may require to adjust the value of h_a accordingly. (See Fig. 3.)

Since the various grids adopted for the membranes and for the velocity potential have different spacings, interpolation is also necessary. Considering the coupling between the batter membrane and air first, we introduce two interpolants \mathcal{J} and \mathcal{I} operating from the membrane to a 2D subset of the acoustic field grid, and vice versa. When written in matrix form, they must satisfy the following relation for energy conservation reasons [9]:

$$\mathcal{I} = \frac{h_a^2}{h^2} \mathcal{J}^T, \quad (26)$$

where the superscript T denotes matrix transposition. Now, we can write (12) as:

$$f^+ = -\rho_a \mathcal{I} \delta_t. \Psi^+, \quad f^- = \rho_a \mathcal{I} \delta_t. \Psi^-, \quad (27)$$

where Ψ^+ and Ψ^- represent the 2D slices of the grid at position p^+ and p^- , respectively. Equations (13) become:

$$\mathcal{J} \delta_t. w = -\delta_{z-} \Psi^+ = -\delta_{z+} \Psi^-. \quad (28)$$

Equations (28) can be inserted into (27) to obtain an expression for f^+ and f^- in terms of known values and w^{n+1} , which leads to an implicit update expression for the membrane itself.

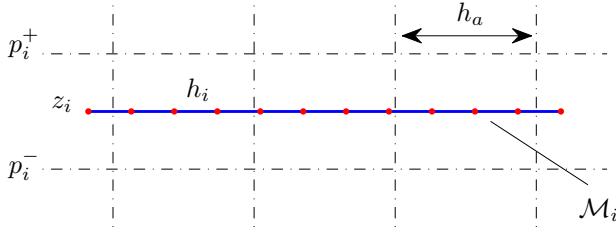


Figure 3. Vertical cross section of the acoustic field grid (in black). The membrane \mathcal{M}_i (in blue) is placed at position z_i , halfway between 3D grid points p_i^+ and p_i^- . Spacings for the 2D and 3D grids are h_i and h_a , respectively.

3.4 Stability and Numerical Energy

Stability conditions for the finite difference schemes written above can be obtained through energy analysis [20]. For the membranes schemes (17) we obtain:

$$h_i^2 \geq c_i^2 k^2 + 4k\sigma_i + \sqrt{(c_i^2 k^2 + 4k\sigma_i)^2 + 16\kappa_i^2 k^2}, \quad (29)$$

while for the acoustic field equation (22) we can write

$$h_a^2 \geq 3c_a^2 k^2. \quad (30)$$

In both cases, spatial grid steps are chosen as close as possible to their minimum value.

It can be shown that a numerical energy \mathfrak{h} exists for the system, which is positive definite and strictly dissipated (that is, $\mathfrak{h}^{n+1} < \mathfrak{h}^n$). If coefficients σ_i are set to zero and reflecting conditions are applied over $\partial\mathcal{V}$, \mathfrak{h} remains constant to machine precision. This is a powerful tool for debugging purposes, as virtually any error in the code has an influence on \mathfrak{h} . See Section 5.4.

4. IMPLEMENTATION DETAILS

The implementation of the model described in the previous sections has been carried out in MATLAB. Operations involving matrices and vectors can be performed in a natural way in this environment. Problems that require the solution of a sparse linear system, for example, could be solved with a direct method without much difficulties. However, with some extra work it is possible to take advantage of the particular characteristics of the system under consideration to devise more efficient ways to perform this operation. In the present case, a naïve implementation of the scheme for the non-linear membranes in (17) and (18) would lead to a code that would barely run, even at low sample rates, on standard machines.

In this section, we will quantify the size of the system and we will describe and compare different implementation strategies for the membrane equation, namely exact

methods and an *ad hoc* iterative method. The energy conservation of the scheme can be preserved to machine accuracy in both cases.

4.1 Algorithm and System Size

A schematic description of a possible implementation of the algorithm outlined in the previous section is given in [12]. In the present case, however, the size of the system can be much larger. Typical physical parameters for the membranes and the air at audio sample rate lead to the grid sizes in Table 1.

Component	Dimensions	Total grid points
2D membranes	265×265	70,225
3D space	$110 \times 110 \times 110$	1,331,000

Table 1. Dimensions of finite difference grids for the various components at $F_s = 44.1$ kHz. Here the 2D grid is for a single bass drum membrane of radius $R = 45$ cm with typical physical parameters [1] in a 1.5 m^3 box.

4.2 Linear System Solution

The update scheme for the membranes given by the coupled equations (17) and (18) is inherently implicit, due to two independent factors: the presence of air coupling, on the one hand, and the non-linearity on the other. Their separate action has been analysed in detail in [9] and [22], respectively, but they have never been adopted simultaneously, so far. The Berger-type non-linearity described in [12], in fact, despite its appearance can be written as an explicit update. Though there is no conceptual difficulty in finding the update scheme, it becomes very challenging from a computational point of view to put into practice the algorithm in the present case, especially because of the large number of grid points involved.

It is possible to “vectorize” the grid functions w and Φ for the two membranes and combine them into a single column vector. When written in matrix-vector form, and after taking into account the air coupling conditions (27) and (28), the update recursion for w^{n+1} and Φ^{n+1} can be written schematically as

$$\begin{bmatrix} \mathbf{1} + \alpha\mathcal{I}\mathcal{J} & -L_1 \\ L_2 & \Omega \end{bmatrix} \begin{bmatrix} w^{n+1} \\ \Phi^{n+1} \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad (31)$$

where $\mathbf{1}$ is the identity matrix, $\alpha\mathcal{I}\mathcal{J}$ is the symmetric matrix coming from air coupling with \mathcal{I} and \mathcal{J} the interpolants defined in (26) and α a dimensionless coupling constant, $\Omega = \delta_{2\Delta}\delta_{2\Delta}$ is the biharmonic operator, L_1 and L_2 are square matrices related to the von Kármán operator (20), and $a = a(w^n, w^{n-1}, \Phi^{n-1})$ and $b = b(\Phi^n)$ are the vectors constructed from values from previous time steps. It is interesting to notice that the diagonal blocks are constant, while L_1 and L_2 must be computed at every time step, since they depend on the action of the \mathcal{L} operator defined in (4) on w^n .

The system (31) has now the form $Sx = q$, and it is possible to solve it straightforwardly using, e.g., Gaussian elimination. However, the results in terms of computation speed

suggest to find a possible alternative (see Section 5.3.) It becomes natural, then, to take advantage of the fact that (under typical conditions for real drums) the coupling term α is small (order 10^{-2}) compared to the identity matrix to create an iterative method for solving the system. We start by writing $S = S_1 + S_2$, where

$$S_1 = \begin{bmatrix} \mathbf{1} & -L_1 \\ L_2 & \Omega \end{bmatrix}, \quad S_2 = \begin{bmatrix} \alpha \mathcal{I} \mathcal{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (32)$$

where $\mathbf{0}$ represents a square matrix with all entries equal to 0 and $\mathbf{1}$ is the identity matrix. Incidentally, S_1 is the update matrix that one obtains from the simulation of a non-linear von Kármán model in a vacuum [22]. With these positions, it is possible to write the iteration

$$S_1 x_{j+1} = -S_2 x_j + q, \quad (33)$$

starting from an initial guess x_0 . This method produces an approximate solution of (31) with the desired precision, if and only if every eigenvalue λ of $S_1^{-1} S_2$ satisfies $|\lambda| < 1$. Furthermore, the rate of convergence of this algorithm depends on the maximum size of $|\lambda|$ [23]. In the present case, the first condition holds and the algorithm requires 8-12 iterations to converge within machine accuracy. We can write (33) explicitly as

$$\begin{bmatrix} \mathbf{1} & -L_1 \\ L_2 & \Omega \end{bmatrix} \begin{bmatrix} w^{n+1} \\ \Phi^{n+1} \end{bmatrix}_{j+1} = \begin{bmatrix} \bar{a} \\ b \end{bmatrix}_j, \quad (34)$$

with $\bar{a}_j = -\alpha \mathcal{I} \mathcal{J} w_j^{n+1} + a$. The iterative method only concerns w , and $w_{j=0}^{n+1} = w^n$ is the natural choice for the initial guess. Now, dropping the superscript $n+1$, it is possible to write the system in terms of Φ_{j+1} only and subsequently update w_{j+1} :

$$(\Omega + L_2 L_1) \Phi_{j+1} = b - L_2 \bar{a}_j \quad (35a)$$

$$w_{j+1} = \bar{a}_j + L_1 \Phi_{j+1}. \quad (35b)$$

For the solution of the linear system (35a) there are two options: an exact method or an iterative method. For sound synthesis purposes, in particular, single precision solutions may be satisfactory. In this case, an iterative solver with a less stringent tolerance may be adopted, which leads to an additional speed-up of the code (see Section 5.3 for details.) One possibility is to use a preconditioned conjugate gradient (*pcg*) algorithm [24], capitalising on the symmetric (apart from low rank corrections) and positive definite structure of the reduced matrix $S_r = (\Omega + L_2 L_1)$. It has already been said that L_1 and L_2 depend on w^n , and they are responsible for the non-linear effects in the system. In fact, the product $L_2 L_1$ in S_r can be seen as a non-linear correction to the biharmonic operator Ω . This contribution is important at high vibration amplitudes of the membrane, but becomes less and less relevant as the energy of the system is dissipated. In the light of these considerations, it is natural to choose Ω as preconditioner, and to feed it to the *pcg* algorithm in the form of a (possibly incomplete) Cholesky factorization [24].

A schematic outline of the algorithm presented above is given in Fig. 4.

```

 $w_{j=0} \leftarrow w^n;$  ▷ initial guess for  $w_{j=0}^{n+1}$  is  $w^n$ 
for  $j = 0 \rightarrow \text{maxiter} - 1$  do
     $\bar{a}_j \leftarrow -\alpha \mathcal{I} \mathcal{J} w_j + a$  ▷ calculate  $\bar{a}_j$ 
     $\text{temp} \leftarrow b - L_1 \bar{a}_j$ 
     $(\Omega + L_2 L_1) \Phi_{j+1} = \text{temp}$  ▷ solve lin sys (35a)
     $w_{j+1} \leftarrow \bar{a}_j + L_1 \Phi_{j+1}$  ▷ update  $w_{j+1}$ 
end for
 $\Phi^{n+1} \leftarrow \Phi_{\text{maxiter}}$  ▷ update  $\Phi^{n+1}$ 
 $w^{n+1} \leftarrow w_{\text{maxiter}}$  ▷ update  $w^{n+1}$ 
    
```

Figure 4. Pseudocode for the iterative method described in Sec. 4.2.

5. RESULTS

5.1 Non-linear Membrane

The behaviour of a non-linear system can be dramatically different from that of a linear one. The principle of superposition no longer holds, energy exchanges between different modes are allowed, and the system can ultimately exhibit chaotic behaviour [4].

One among typical non-linear phenomena is the pitch glide effect; an example of this is provided in Fig. 5, where an unnaturally high excitation is used for illustration purposes.

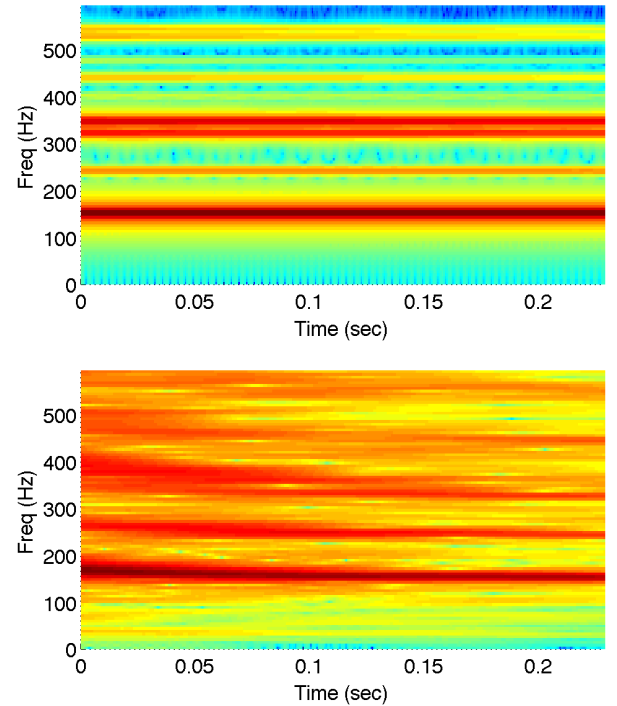


Figure 5. Spectrograms for sound outputs from a virtual tom-tom ($R = 20$ cm) at low (up) and high (down) striking amplitudes. A pitch glide can be seen at high amplitudes of excitation.

5.2 3D Acoustic Field

With an FDTD simulation, one has access to the entire state of the system and, in particular, to the full 3D acous-

tic field. Figure 6 shows, as an example, some snapshots of a cross section of the acoustic field after a short raised cosine excitation (0.3 ms) on the batter membrane. The propagating pressure waves are clearly visible.

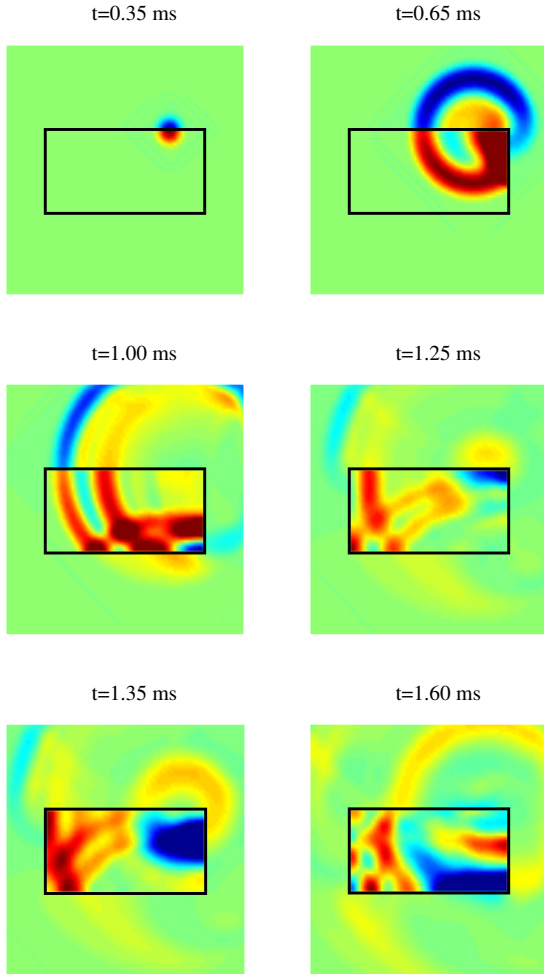


Figure 6. Snapshots of the cross section of the acoustic field after a strike on the batter membrane of the drum, at times as indicated.

5.3 Linear System

In this section we compare different methods for the solution of the linear system described in Section 4.2, the exact solver (31) and the recursive approach (34) based on the reduced system (35). The latter, in particular, can be tackled by means of an exact method or an iterative one. For this second case, one possibility is to use *pcg* solvers, with the Cholesky factorization of Ω as preconditioner and single precision tolerance.

Table 2 shows the computation times at different sample rates for the three methods described. The difference between exact and iterative methods is apparent. This result is even more striking if we think that the solution of the reduced linear system is performed many times (fewer than 10, in this case) inside the recursion.

sample rate	dir	iter+dir	iter+pcg
8 kHz	87 s	15 s	8 s
16 kHz	802 s	206 s	102 s

Table 2. Comparison between different approaches to the linear system solution at different sample rates (dir=direct method applied to (31), iter+dir=iterative method (Fig. 4) with exact solution of (35a), iter+pcg=iterative method (Fig. 4) with iterative solution of (35a) using *pcg*.) The computation time refers to 0.01 s of output and for a drum with $R = 0.20$ cm.

5.4 Energy Conservation

In a test simulation with no losses in the membrane equations, and with perfectly reflecting conditions over the boundary ∂V , we can check that the numerical energy of the system be conserved to double precision, even with an iterative solution of the linear system.

Figure 7 shows numerical energy conservation for the system, with energy variations $(h^n - h^0)/h^0$ of the order of machine accuracy.

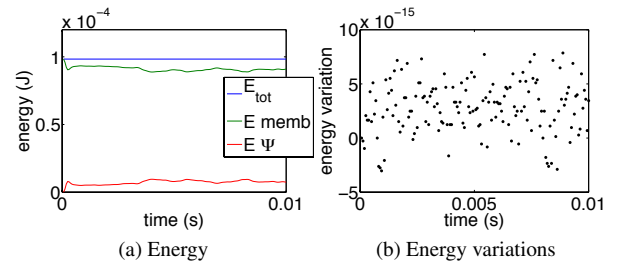


Figure 7. The energy of the system is shown in (a): total energy remains constant, while there is an energy exchange between the membranes and the air. Normalized energy variations in (b) are of the order of machine accuracy.

5.5 Sound examples

Sound examples produced using the model described above can be found at

<http://www2.ph.ed.ac.uk/~s1164558>

6. FINAL REMARKS

In this paper, we have addressed one of the questions raised in [12] by including a fuller non-linear model in the equations for the membranes. This seems to have been so far (at least one of) the missing pieces of a more realistic model of the drum, perhaps with regard to the dramatic attack that can be heard at high striking amplitudes. The Berger model adopted in [12], in its simplicity, can produce pitch glide effects, but is not able to render the build-up of high-frequency energy typical of high excitations.

As mentioned before, however, this must be considered as a preliminary study of this system, and additional research will be performed in several directions. First of all, an experimental validation of the parameters used here

must be conducted, especially concerning the viscosity parameters of the membranes. Then, an explicit modeling of the drum shell has not been included in the present model. Its importance in the case of the snare drum has been studied in [7], and it is reasonable to expect that a similar behaviour may be found for bass drums and tom-toms, as well. Again, this is something that will be the object of experimental study.

From the computational point of view, the core of the algorithm presented is the solution of a very large sparse linear system. One of the major ongoing efforts at Edinburgh is towards finding an efficient method for the solution of this system that could eventually be parallelized. Some of the progress made is outlined in the present paper, but it is clear that more work needs to be done.

Acknowledgments

This work was supported by the European Research Council, under grant StG-2011-279068-NESS.

7. REFERENCES

- [1] H. Fletcher and I. G. Bassett, "Some experiments with the bass drum," *The Journal of the Acoustical Society of America*, vol. 64, p. 1570, 1978.
- [2] S. Dahl, "Spectral changes in the tom-tom related to striking force," *Speech Music and Hearing Quarterly Progress and Status Report*, vol. 38, no. 1, pp. 059–065, 1997.
- [3] T. Tolonen, V. Välimäki, and M. Karjalainen, "Modeling of tension modulation nonlinearity in plucked strings," *Speech and Audio Processing, IEEE Transactions on*, vol. 8, no. 3, pp. 300–310, 2000.
- [4] A. Chaigne, C. Touzé, and O. Thomas, "Nonlinear vibrations and chaos in gongs and cymbals," *Acoustical science and technology*, vol. 26, no. 5, pp. 403–409, 2005.
- [5] R. S. Christian, R. E. Davis, A. Tubis, C. A. Anderson, R. I. Mills, and T. D. Rossing, "Effects of air loading on timpani membrane vibrations," *The Journal of the Acoustical Society of America*, vol. 76, p. 1336, 1984.
- [6] L. Rhaouti, A. Chaigne, and P. Joly, "Time-domain modeling and numerical simulation of a kettledrum," *The Journal of the Acoustical Society of America*, vol. 105, p. 3545, 1999.
- [7] T. D. Rossing, I. Bork, H. Zhao, and D. O. Fystrom, "Acoustics of snare drums," *The Journal of the Acoustical Society of America*, vol. 92, p. 84, 1992.
- [8] F. Avanzini and R. Marogna, "A modular physically based approach to the sound synthesis of membrane percussion instruments," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 4, pp. 891–902, 2010.
- [9] S. Bilbao, "Time domain simulation and sound synthesis for the snare drum," *The Journal of the Acoustical Society of America*, vol. 131, p. 914, 2012.
- [10] L. Savioja, D. Manocha, and M. Lin, "Use of GPUs in room acoustic modeling and auralization," in *Proc. Int. Symposium on Room Acoustics*, 2010.
- [11] Next Generation Sound Synthesis (NESS) Project. [Online]. Available: <http://www.ness-music.eu>
- [12] S. Bilbao and C. Webb, "Timpani drum synthesis in 3D on GPGPUs," in *Proc. of the 15th Int. Conference on Digital Audio Effects (DAFx-12)*, 2012.
- [13] A. H. Nayfeh and D. T. Mook, *Nonlinear oscillations*. New York: John Wiley and Sons, 1979.
- [14] N. Fletcher and T. Rossing, *The physics of musical instruments*. Springer Verlag, 1998.
- [15] B. Bank, "Physics-based sound synthesis of stringed instruments including geometric nonlinearities," Ph.D. dissertation, Budapest University of Technology and Economics, 2006.
- [16] A. Chaigne and C. Lambourg, "Time-domain simulation of damped impacted plates. I. Theory and experiments," *The Journal of the Acoustical Society of America*, vol. 109, p. 1422, 2001.
- [17] O. Thomas and S. Bilbao, "Geometrically nonlinear flexural vibrations of plates: In-plane boundary conditions and some symmetry properties," *Journal of Sound and Vibration*, vol. 315, no. 3, pp. 569–590, 2008.
- [18] J.-P. Berenger, "Three-dimensional perfectly matched layer for the absorption of electromagnetic waves," *Journal of computational physics*, vol. 127, no. 2, pp. 363–379, 1996.
- [19] B. Engquist and A. Majda, "Absorbing boundary conditions for numerical simulation of waves," *Proceedings of the National Academy of Sciences*, vol. 74, no. 5, pp. 1765–1766, 1977.
- [20] S. Bilbao, *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*. Wiley Publishing, 2009.
- [21] B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time dependent problems and difference methods*. Wiley New York, 1995.
- [22] S. Bilbao, "A family of conservative finite difference schemes for the dynamical von Karman plate equations," *Numerical Methods for Partial Differential Equations*, vol. 24, no. 1, pp. 193–216, 2007.
- [23] G. Strang, *Linear Algebra and Its Applications*. London: Academic Press, 1976.
- [24] Y. Saad, *Iterative methods for sparse linear systems*. Philadelphia: SIAM, 2003.

EXPERIMENTAL STUDY OF COUPLED DRUMHEAD VIBRATIONS USING ELECTRONIC SPECKLE-PATTERN INTERFEROMETRY

Randy Worland

Department of Physics

University of Puget Sound

Tacoma, WA 98416 USA

worland@pugetsound.edu

ABSTRACT

The coupled vibrations of a two-headed musical snare drum were investigated experimentally using electronic speckle-pattern interferometry. A dual interferometer system was used to record images of both vibrating heads simultaneously. Operating deflection shapes and frequencies are reported for the first several coupled modes, along with their relative amplitudes, orientations, and phase relations.

Previously reported results for coupled drumhead modes are verified and extended to include the effects of doubly degenerate mode pairs that are split due to non-uniform tension in the drumheads. The (1,1) mode shapes are found to create four sets of coupled vibrations, with clear angular orientations and phase relations between the two heads in each case. The higher frequency (2,1) mode shapes are less strongly coupled, but do exhibit three of the four possible coupled pairs with this particular drum and tuning.

1. INTRODUCTION

Literature on the behavior of a single musical drumhead includes experimental, theoretical, and numerical work dating back to Rayleigh's discussion of circular membrane vibrations [1-4]. The ideal circular membrane with fixed perimeter produces normal mode shapes designated by integers (m,n) , where m represents the number of nodal diameters and n the number of nodal circles (see Figure 1). Each mode containing nodal diameters ($m > 0$) is doubly degenerate, with two complementary angular orientations corresponding to the same frequency. Thus, the (1,1) mode pair differ in orientation by 90° , while the (2,1) pair differ by 45° . The frequencies produced by real drumheads are not identical to those calculated in the ideal case for several reasons, including air loading and membrane stiffness.

Furthermore, non-uniformities in membrane tension may disturb the symmetry of the nodal patterns in a real drumhead [5]. In such cases the degeneracy is lifted and orthogonal patterns will occur at slightly different frequencies for each non-zero value of m .

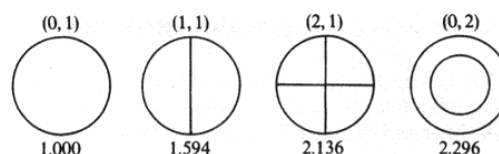


Figure 1. The first four normal mode shapes of an ideal membrane, with relative frequencies indicated below. Integers (m,n) designate the number of nodal diameters and circles respectively. Only one of each degenerate pair is shown for the (1,1) and (2,1) modes.

Many musical drums, such as tom toms, bass drums, and snare drums include two heads, on opposite sides of a cylindrical shell (see Figure 2). These heads are often referred to as the “batter” head (the one that is struck) and the “resonant” head, or in the case of a snare drum, the “snare side” head. The two heads used on these drums are often of different thicknesses and tuned to different pitches. Vibrations of the two heads are coupled acoustically by the air inside the shell, and mechanically by the shell itself. As a result, both heads vibrate when the batter head is struck. Lower frequency modes are known to couple more strongly than the higher frequency modes [2-4].



Figure 2. Tom toms (left) and snare drum (right) with two heads. The snare drum shown is similar to the one used in this study.

The current paper presents optical imaging of coupled drumhead vibrations using electronic speckle-pattern interferometry (ESPI). Techniques for obtaining and interpreting the images are described, along with the method for determining the relative phases of the vibrations. In particular, the effects due to non-uniform tension are examined in both the (1,1) and (2,1) mode shapes. Results are interpreted and compared with previous studies of coupled drumheads [6-8].

2. MODELS

2.1 One-Dimensional Model

The simplest model of coupled oscillators is that of two mass and spring systems connected by a third spring as shown in Figure 3. This one-dimensional system is described in many mechanics texts [9, 10] and has been applied to the case of drumhead vibrations by Rossing *et al.* [6]. The main results of the model are summarized below.

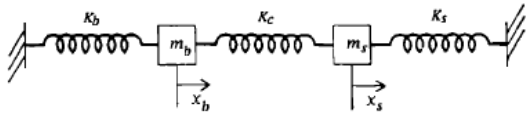


Figure 3. Two mass and spring oscillators coupled by a spring of stiffness K_c [6]. When applied to the (0,1) modes of a two-headed drum, the subscripts b and s represent the batter and snare side heads respectively.

In the absence of coupling, the natural frequency of each independent oscillator (f_b and f_s) depends only on its mass and spring constant. When the coupling spring K_c is added, the combined system has two normal modes consisting of in-phase motion (at a lower frequency), and out-of-phase motion (at a higher frequency). The degree to which the frequencies of the coupled system differ from those of the uncoupled oscillators varies with the strength of the coupling. In general, stronger coupling alters the frequencies by a greater amount, while weak coupling produces relatively small changes in the frequencies. As the degree of coupling approaches zero, the vibrations of the masses m_b and m_s become increasingly independent, at frequencies approaching f_b and f_s .

As described by Rossing, *et al.* [6], each head of the drum may be represented by an oscillator, with the coupling spring accounting for the combined action of the enclosed air mass and the shell. This one-dimensional approach provides a good qualitative understanding of the basic coupling mechanism, particularly as applied to the fundamental mode (0,1) of the drum. Figure 4 shows side views of the resulting modes for both the in-phase (0,1)₋ and out-of-phase (0,1)₊ drumhead motion.

The relative amplitude (x_b / x_s) of the two oscillators also depends on the coupling strength. With strong coupling and similar masses the amplitudes will be comparable, and the amplitude ratio will be on the order of one for both in-phase and out-of-phase motions. In weak coupling the amplitude ratios will differ significantly from unity and will exhibit an inverse relationship between the

in-phase and out-of-phase modes [9]. Thus, with weakly coupled oscillators m_b will exhibit a larger amplitude than m_s at one eigenfrequency of the system, while m_s will display the larger relative amplitude at the other normal mode frequency. These amplitude results will be applied to the experimental drumhead data of Section 4, as a qualitative measure of coupling strength.

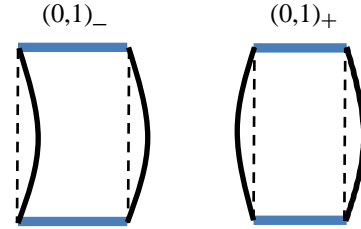


Figure 4. Side view of a drum with coupled heads vibrating in the (0,1) mode shape. The in-phase motion (left) occurs at the lower frequency and is designated (0,1)₋ while the out-of-phase motion (right) occurs at the higher frequency and is designated (0,1)₊.

2.2 Beyond the One-Dimensional Model

Modes beyond the (0,1) pattern can also couple, although typically the lower frequency modes show the strongest coupling. Rossing *et al.* [6] describe the coupling of (1,1) modes and report some evidence of (0,2) mode coupling in experiments in which care was taken to ensure uniform tension in both heads of a snare drum.

The degree of coupling of higher modes is likely to depend on a variety of factors, including drum geometry (diameter and shell depth) and membrane tensions (both absolute and relative). Non-uniform tension in the heads opens up a new set of possibilities, as degenerate modes in each head are split, and mode shapes may no longer be identical in the two heads.

For example, with slightly non-uniform tension the (1,1) mode of a single head splits into an orthogonal pair of patterns, with slightly different resonant frequencies [5]. Thus, the coupled system may have as many as four deflection shapes stemming from the ideal (1,1) mode. The orientations, frequencies, relative amplitudes, and relative phases for these cases are examined experimentally in Section 4, along with those stemming from the (2,1) mode.

3. OPTICAL MEASUREMENT TECHNIQUES

Electronic speckle-pattern interferometry (ESPI) was used to obtain images of the vibrating drumheads. The ESPI setup used here is based on the system described by Moore and Zietlow [11]. Light from a helium-neon laser is split into two coherent beams, with one beam illuminating the vibrating object while the other beam serves as a reference beam of nearly identical path length. The beams are recombined and the resulting speckle-pattern interference images are recorded by a CCD camera. The interferometer is sensitive to out-of-plane motion of the vibrating object on the order of the light wavelength, such that small amplitude steady state vibrational patterns may

be observed. This ESPI technique provides operating deflection shapes (which may be closely related to the normal modes of the system) and the corresponding frequencies. The images show nodal lines as white, and contours of equal amplitude motion as fringes of dark and light gray (see Figures 6-8). The amplitude of the vibrations can be determined from the number and spacing of these fringes.

3.1 The Dual ESPI System

In order to capture images from both vibrating heads of a drum simultaneously, two identical systems of the type described by Moore and Zietlow were set up on an optical table, facing opposite sides of the drum. In the lab, the two images of the vibrating heads were observed in real time on adjacent computer monitors. Because the images were recorded from opposite directions, a mirror was used to reflect the image from one monitor so that the relative orientations of the two patterns could be visually compared as the driving frequency was varied. Similarly, the digital images recorded from one of the two systems were also flipped (rotated 180° about a vertical axis) for comparison purposes (see Figures 7 and 8).

The drum being studied was positioned vertically in the center of the optical table and excited acoustically with a small speaker (e.g. 2-8" diameter) driven by an audio frequency sine wave from a function generator. The speaker was typically placed about 50 cm from one of the heads. Although the speaker location could be varied to enhance the clarity of the images and to optimize the absolute number of fringes, the *relative* orientations and amplitudes of the coupled patterns were not altered by the choice of speaker location.

Coupled vibrations were characterized by deflection shapes of both heads showing well-defined patterns with maximum response at precisely the same frequency, and with identical (or very similar) shapes and angular orientations. Sharp resonances of the lower frequency patterns matching these criteria were unambiguous, and were reproducible to within about 0.1 Hz. Higher frequency patterns, as well as some broader (more highly damped) resonances, were more difficult to interpret. Many *uncoupled* vibrational cases were also observed, primarily at higher frequencies, with only one head showing a distinct pattern at a particular frequency.

This coupling test relies on both heads producing vibrations in a similar range of amplitudes. As the coupling becomes weaker, the amplitude ratio increases such that *simultaneous* resolution of fringes from both heads is not possible. Thus, for our experiment the coupling must be strong enough to produce an amplitude ratio no greater than about 10:1 between the two heads.

A further verification of the coupled vibrations was performed using a frequency analyzer to locate resonant frequencies when one or the other head of the drum was struck with a drum stick. In this test, coupled vibrations correspond to frequencies that appear in the spectrum regardless of which head is struck. Uncoupled patterns correspond to frequencies that only appear when one or the other head is struck. For the experimental results

reported in Section 4 below, the spectrum analysis tests (not shown here) were consistent with the ESPI imaging results.

3.2 Phase Determination

The one-dimensional model described in Section 2.1 distinguishes clearly between coupled heads oscillating *in* or *out* of phase with each other. The more complex two-dimensional mode shapes also exhibit in-phase or out-of-phase motion when coupled, as shown in Figure 5 for the (1,1) doublet [6]. However, the side view does not reveal the angular orientations of the coupled modes. The basic dual ESPI setup produces images that do reveal these orientations, but does not provide the necessary phase information without some additional refinement.

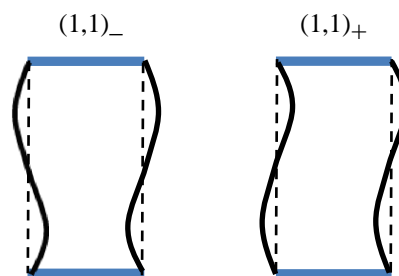


Figure 5. Side view of a drum with coupled heads vibrating in the (1,1) mode. The out-of-phase motion (left) occurs at the lower frequency and is designated (1,1)₋. The in-phase motion (right) occurs at the higher frequency and is designated (1,1)₊.

To determine the relative phases of the vibrating heads, a mirror in each object beam was attached to a mechanical oscillator driven by the same function generator used to drive the speaker. Thus, the mirrors move sinusoidally at the frequency of the drumhead vibrations when activated. This has the effect of altering the object beam's path length at the frequency of the drumhead motion such that the white fringes in the images no longer represent nodal lines. Rather, they represent regions of the drumhead moving in phase with the mirror, and with the same amplitude. The mirror amplitudes can be adjusted separately to optimize the visibility of the white fringes.

In general the drumheads do not move strictly in or out of phase with the speaker cone due to the separation distance between the speaker and the drum. However, once a resonance was found (without the moving mirror), the speaker could be repositioned as needed with the mirror motion turned on until the signature white fringe was observed. This indicated that a portion of the drumhead was moving in phase with the mirror.

In this way, white lines that normally represent nodes are displaced on the image. The direction of this displacement can be interpreted such that it is possible to tell if corresponding portions of the two heads are moving with the same, or opposite, phase. The method produces unambiguous results in cases where the two vibrational shapes are aligned with each other. In the coupled mode data shown below, a high degree of alignment was always observed.

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Snare Drum

The drum used in this experiment is a common 14" x 5" snare drum with a metal shell, with the snares removed. The snare drum was chosen for its relatively large diameter to shell depth ratio (14:5), and for comparison with the results of Reference 6. The batter head used was a single ply "Coated Ambassador" (10-mil Mylar) manufactured by Remo, Inc.; the snare side head was a single ply clear Remo "Ambassador Snare" (3-mil Mylar). Both are typical heads used with snare drums.

The heads were not tuned to the same pitch, as is common with snare drums. Furthermore the individual heads were not tuned carefully for uniform tension. The lack of "ideal" tensioning lifted the degeneracy of modes containing nodal diameters, revealing some interesting results described in Sections 4.3 and 4.4 below.

4.2 The (0,1) Modes

The (0,1) modes of two-headed drums are known to couple strongly [2-4]. The fundamental (0,1)₋ was observed at 142 Hz, with the expected set of nearly circular fringes shown for both heads (see Figure 6). As the driving frequency was varied, both heads showed their maximum response, as judged by a greater number of visible fringes, at this frequency. The higher frequency resonance (0,1)₊ was broad and a bit harder to obtain. The fringe pattern in this case shows very little motion toward the perimeter of the drum, and the observed fringes do not reveal a high degree of circular symmetry. The relative phases for both of the modes shown in Figure 6 are consistent with the previously reported work. The relative amplitudes, as judged by the number of fringes on each head, are similar, as would be expected with strong coupling.

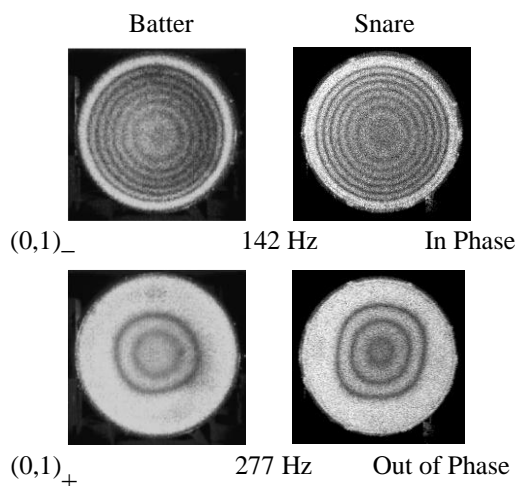


Figure 6. ESPI images showing the (0,1) coupled mode shapes of the batter head (left) and the snare side head (right) for both in-phase motion (top) and out-of-phase motion (bottom). Nodal regions appear white in these images.

4.3 The (1,1) Modes

Although the (1,1) mode of an ideal membrane is doubly degenerate, in practice the degeneracy tends to be lifted by asymmetries in the membrane tension. In the presence of relatively small asymmetries it is still possible for the patterns to rotate in response to the driving mechanism allowing the shapes on the two heads to align with each other. In this way each head may contribute an orthogonal pair of (1,1) shapes for coupling with partner shapes on the opposite head. Based on this model as many as four coupled (1,1) mode shapes may be observed, as opposed to the coupled doublets reported in the case of a uniformly tensioned snare drum [6].

The ESPI results for the coupled (1,1) vibrations are shown in Figure 7. Four resonant frequencies are observed that display patterns very similar to the (1,1) mode shape. The alignment of each coupled pair is striking, even though with non-uniform tension the nodal diameters are not perfectly straight. Each orthogonal alignment contributes both an in-phase and an out-of-phase coupled pair. The out-of-phase motion produces the *lower* frequency set of resonances in this case while the in-phase motion occurs at the *higher* frequencies.

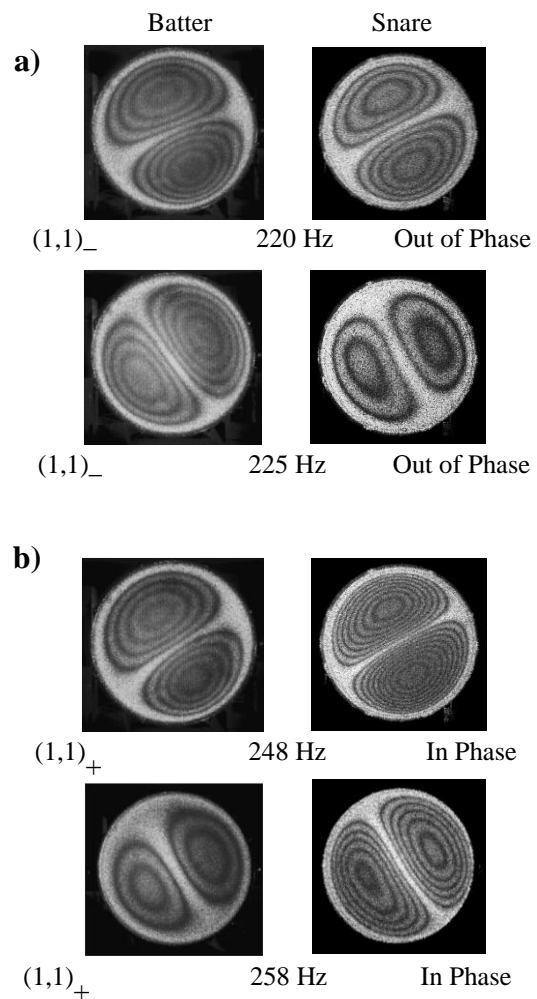


Figure 7. ESPI images showing four sets of coupled vibrations with (1,1) mode shapes. Fig. 7a shows the lower frequency, out-of-phase cases; Fig. 7b shows the higher frequency, in-phase cases.

These frequency vs. phase relations are opposite to those of the (0,1) shape, where the in-phase motion produces the lower frequency. This reversal is consistent with the results of Rossing *et al.* [6], who attributed the lower frequency of the (1,1) in-phase motion to the ability of the enclosed air to slosh back and forth parallel to the heads in modes containing nodal diameters. The key distinction is that modes with $m = 0$ do not conserve the enclosed volume, while modes containing nodal diameters ($m > 0$) do.

The in-phase and out-of-phase motions for a given pattern display the same angular orientations. For example, the top pair in Figure 7a has the same orientation as the top pair in Figure 7b (and similarly for the lower pair of each figure). These common orientations, along with clear phase relations and maximum response at precisely the same frequencies all help to confirm the interpretation of four coupled pairs derived from (1,1) modes.

Although non-uniformity in the tension of the heads is responsible for the production of orthogonal pairs of (1,1) shapes in each head, it appears that for the degree of asymmetry in the current drum at least one pattern is able to rotate freely enough to accommodate the coupling with the other head.

The relative amplitudes shown in Figure 7 are also consistent with the coupled oscillator model, which predicts an inverse relationship in amplitudes for the in-phase and out-of-phase modes [9]. This inverse relationship will be particularly noticeable with decreasing coupling strength. The out-of-phase motion (Fig. 7a) shows slightly less amplitude on the snare side head than on the batter head for both the 220 Hz and 225 Hz resonances. The in-phase cases (Fig. 7b) show the opposite effect; slightly greater amplitude on the snare side head for both the 248 Hz and 258 Hz resonances. The difference here is fairly subtle, as these vibrations still exhibit a significant degree of coupling. The amplitude effect is more striking in the loosely coupled (2,1) patterns described below.

4.4 The (2,1) Modes

The ESPI results for the (2,1) mode shapes are shown in Figure 8. Again, with non-uniform tension each head may contribute an orthogonal pair of (2,1) patterns, creating as many as four coupled patterns. However, in this particular case only three of these coupled patterns are observed. As expected, the higher frequency (2,1) modes are less strongly coupled than the (1,1) patterns described above. The lower frequency cases (Fig. 8a) are again the result of out-of-phase motion, as air can slosh increasingly freely from side to side as more diameters are added to the pattern.

The weakness of the coupling takes these (2,1) patterns close to the limit of detectability with this experimental setup. In Figure 8a the batter head motion is much larger than that of the snare side (at both 320 Hz and 333 Hz) as shown by the large number of batter head fringes, with barely discernible fringes on the snare side head. Nonetheless, the reasonably good alignment of the patterns and their maximum response at the same driving frequency indicate that there is some degree of coupling present. Although requiring some additional care due to

the weakness of the coupling, the phase relationships were obtained unambiguously.

The higher frequency (in-phase) motion revealed only one coupled pair with the (2,1) shape, as shown in the top row of Figure 8b at 349 Hz. The amplitude ratio is qualitatively reversed from that of the top row of Fig. 8a, as expected, with the snare side head now showing the much larger fringe count. However, the (2,1) pair orthogonal to that of the top row is incomplete. The snare side head shows the expected pattern at 392 Hz, but the batter head does not yield a corresponding (2,1) image. Thus, the (2,1) modes are observed to be weakly coupled for three of the four possible cases only.

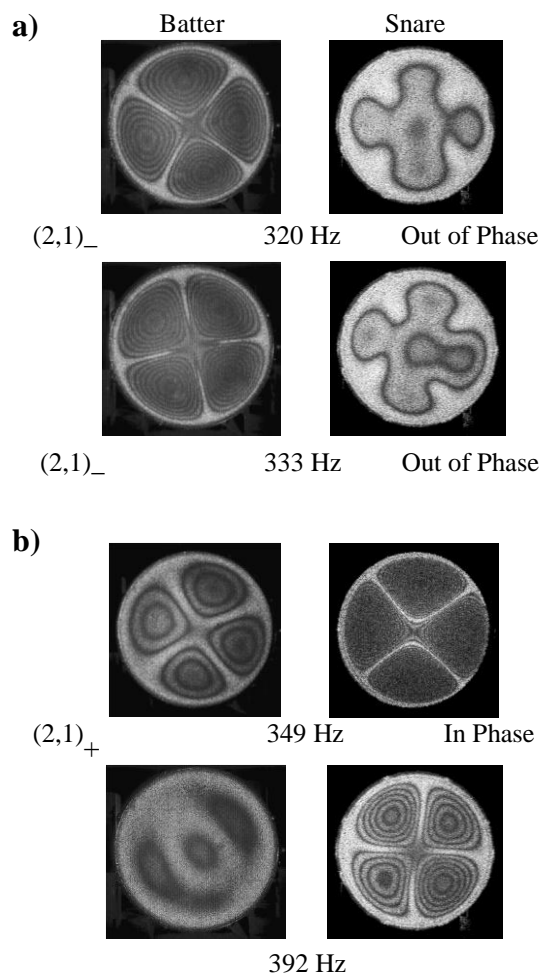


Figure 8. ESPI images showing the (2,1) vibrational mode shapes. Fig. 8a shows the lower frequency, out-of-phase cases, which are very loosely coupled. In Fig. 8b, only the top row shows weakly coupled vibrations. The lower row shows an orthogonal partner on the snare side head, but no matching (2,1) pattern on the batter head.

5. CONCLUSIONS

Reliable techniques were developed for simultaneous real-time ESPI imaging of both sides of a two-headed drum. Information obtained with this system includes the resonant frequency, operational deflection shapes, relative orientations, relative amplitudes, and relative phases of the coupled oscillations.

Previously reported results were confirmed regarding the coupling of both the (0,1) and (1,1) modes in a snare drum, even with heads that were not uniformly tensioned. In addition, weak coupling of the (2,1) modes was observed. Most of the modes with frequencies higher than the (2,1) appeared to be independent, and were either uncoupled, or loosely coupled below the current detection limit, as expected.

With non-uniform membrane tension deliberately introduced, the effects of lifted degeneracy in the (1,1) and (2,1) shapes were observed. The (1,1) case clearly showed four coupled patterns with matching angular orientations and the expected phase relations. The (2,1) shapes showed three of the four possible coupled combinations. Amplitude information obtained from the ESPI images was consistent with very weak coupling for the (2,1) shape.

Future investigations may include a systematic study of the factors influencing the degree of higher frequency mode coupling. Parameters of interest include the diameter and shell depth of the drum, as well as the tension and thickness of each head. Varying the degree of non-uniformity in the applied tension of the heads may also yield additional insight into the practical behavior of coupled vibrations in two-headed drums.

Acknowledgements

The author would like to thank Thomas Moore of Rollins College for many helpful discussions regarding the ESPI techniques, particularly those involving the phase determination.

6. REFERENCES

- [1] J.W.S. Rayleigh, *The Theory of Sound*, 2nd ed., MacMillan, London, 1894 (reprinted by Dover, 1945), Ch. 9.
- [2] T.D. Rossing, *Science of Percussion Instruments*. World Scientific, 2000.
- [3] N.H. Fletcher and T.D. Rossing, *The Physics of Musical Instruments*, 2nd ed. Springer, 1998.
- [4] T.D. Rossing, ed., *Springer Handbook of Acoustics*. Springer, 2007, pp. 642-648.
- [5] R. Worland, "Normal modes of a musical drumhead under non-uniform tension," *J. Acoust. Soc. Am.*, vol. 127, no. 1, pp. 525-533, 2010.
- [6] T.D. Rossing, I. Bork, H. Zhao, and D. Fystrom, "Acoustics of snare drums," *J. Acoust. Soc. Am.*, vol. 92, no. 1, pp. 84-94, 1992.
- [7] P.G.M. Richardson, E.R. Toulson, and D.J.E. Nunn, "Analysis and manipulation of modal ratios of cylindrical drums," *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp. 907-913, 2012.
- [8] H. Suzuki and Y. Hwang, "Coupling between two membranes of a Japanese drum," *Acoust. Sci. & Tech.*, vol. 29, no. 3, pp. 215-220, 2008.
- [9] K.R. Symon, *Mechanics*, 3rd ed. Addison-Wesley, 1971, pp. 191-197.
- [10] J.B. Marion, *Classical Dynamics*, 2nd ed. Academic Press, 1970, pp.408-418.
- [11] T.R. Moore and S.A. Zietlow, "Interferometric studies of a piano soundboard," *J. Acoust. Soc. Am.*, vol. 119, no. 3, pp. 1783-1793, 2006.

Physical modeling

Acoustics of pianos: physical modeling, simulations and experiments

Antoine Chaigne

UME, Ensta ParisTech, France

antoine.chaigne@ensta.fr

Juliette Chabassier

Poems Team, Inria, France

juliette.chabassier@inria.fr

Nicolas Burban

Ecole polytechnique, France

nicolas.burban@polytechnique.edu

ABSTRACT

The outlines of a recently developed model of a grand piano are summarized. Using dedicated numerical methods, the main vibratory and acoustic variables of each constitutive part of the instrument (strings, bridge, soundboard, sound pressure) are simulated in the time-domain. The obtained waveforms are analyzed and compared with experimental data derived from measurements on a Steinway D grand piano. This comparison yields valuable insight into the physics of the instrument. It shows, in particular, that a nonlinear string model is necessary to account for the observed richness of piano spectra. The model is able to reproduce important features of piano sounds, such as the presence of soundboard modes in the transients, precursors and phantom partials. However, one important limitation of the model, in its present state, is that it does not account for the change of polarization observed on piano strings. Experimental observations of this phenomenon are discussed and a preliminary model for explaining the possible role of the zig-zag end condition in string polarization change is presented.

1. INTRODUCTION

A piano model has been recently developed that couples together the hammer, the strings, the soundboard and the air [1]. One particularity of this model is that it is based on a nonlinear description of string motion. As a consequence, an original model of string-soundboard coupling at the bridge is developed, in order to allow transmission of both transverse and longitudinal forces to the soundboard. This implies to elaborate dedicated numerical schemes in order to ensure stability, and to satisfy the strong requirements of musical sound synthesis in terms of dispersion [2] [3]. Since both the model and the associated numerical methods are extensively described elsewhere, only a brief summary is given in Section 2. In Section 3, some selected results of simulations are analyzed both in the time and frequency domains, where the influence of string amplitude in the resulting sounds is highlighted. Comparison with experimental waveforms and spectra shows that the model is able to reproduce the precursors, the phantom partials and the presence of soundboard modes observed in recorded

piano tones. However, in its present state, the model is not capable to account for the time evolution of the string's polarization plane observed in most pianos. In Section 4, a mixed experimental-numerical preliminary study is reported whose aim is to examine whether the usually observed and so-called “zig-zag” end conditions can be responsible for such changes in polarization. These changes are important since they affect the temporal envelope of the tones and are clearly audible.

2. A PIANO MODEL

2.1 Summary of the model

The model allows to calculate the sound of a piano in the time-domain, from the starting time where the hammer hits the strings, initially at rest, with a given velocity. The key mechanism that transmits the player's action from the keyboard to the hammer is ignored. It is considered to include it in future versions. One major feature of the model is that the geometrical nonlinearity the strings, due to large displacements, is taken into account in addition to stiffness. The model allows the transmission of these nonlinearities from strings to soundboard at the bridge. This yields, among other things, the possibility of simulating phantom partials and precursors [4]. The hammer is defined as a dissipative nonlinear spring [5]. The motion of the strings is governed by nonlinear wave propagation equations, accounting for the large displacement that can be observed especially during the attack [6]. This geometrical nonlinearity induces a coupling between the transverse and the longitudinal polarizations of the strings. This coupling has many consequences on the spectral content, and also on the temporal envelopes of the sounds. The string stiffness is modeled by a Timoshenko model, which has better physical and mathematical properties than the Euler-Bernoulli model. The string model is written:

Boundary condition (agraffe side $x = x_a$): (1)

$$u_s(x = x_a, t) = v_s(x = x_a, t) = \frac{\partial \varphi_s}{\partial x}(x = x_a, t) = 0,$$

Initial conditions: (2)

$$\begin{cases} u_s(x, t = 0) = v_s(x, t = 0) = \varphi_s(x, t = 0) = 0, \\ \frac{\partial u_s}{\partial t}(x, t = 0) = \frac{\partial v_s}{\partial t}(x, t = 0) = \frac{\partial \varphi_s}{\partial t}(x, t = 0) = 0. \end{cases}$$

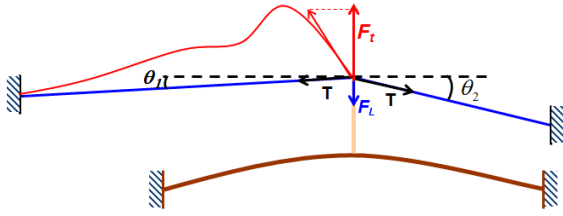


Figure 1. String-soundboard coupling. The force transmitted from string to soundboard has two components. F_t is due to the transverse motion of the string, while F_L is due to its longitudinal motion.

Nonlinear stiff string motion:

$$\left\{ \begin{array}{l} \rho A \frac{\partial^2 u_s}{\partial t^2} - \frac{\partial}{\partial x} \left[EA \frac{\partial u_s}{\partial x} - \frac{(EA - T_0) \frac{\partial u_s}{\partial x}}{\sqrt{\left(\frac{\partial u_s}{\partial x}\right)^2 + \left(1 + \frac{\partial v_s}{\partial x}\right)^2}} \right] \\ + AG\kappa \frac{\partial}{\partial x} \left(\varphi_s - \frac{\partial u_s}{\partial x} \right) = S, \\ \rho A \frac{\partial^2 v_s}{\partial t^2} - \frac{\partial}{\partial x} \left[EA \frac{\partial v_s}{\partial x} - \frac{(EA - T_0) \left(1 + \frac{\partial v_s}{\partial x}\right)}{\sqrt{\left(\frac{\partial u_s}{\partial x}\right)^2 + \left(1 + \frac{\partial v_s}{\partial x}\right)^2}} \right] \\ = 0, \\ \rho I \frac{\partial^2 \varphi_s}{\partial t^2} - EI \frac{\partial^2 \varphi_s}{\partial x^2} + AG\kappa \left(\varphi_s - \frac{\partial u_s}{\partial x} \right) = 0 \end{array} \right. \quad (3)$$

where u_s is the vertical transverse displacement of the string, v_s is the longitudinal displacement, and φ_s is the angle of the cross-sections with the plane normal to the string. The source term S represents the action of the hammer. For clarity, the additional fluid and viscoelastic-like damping terms are not written in Equation (3). The string-soundboard coupling at the bridge is modeled in such a way that both the transverse and longitudinal components of the strings are transmitted to the soundboard. This is obtained by considering that the string is slightly bent due to both the bridge height and soundboard curvature (see Figure 1). As a result, the vibration spectrum of the soundboard (velocity or acceleration) has a full richness, comparable to real tones. The soundboard is modeled as a flat orthotropic Reissner-Mindlin plate of variable thickness [1]. The ribs and the bridge are modeled as local heterogeneities in terms of thickness and elasticity (see Figure 2). The parameters of the soundboard (size, materials and thickness profile) are adjusted with great accuracy, in order to allow comparison with existing soundboards. The number and location of ribs can be adjusted so that comparisons can be made between the sounds produced for various configurations of the soundboard. The 3D sound field around

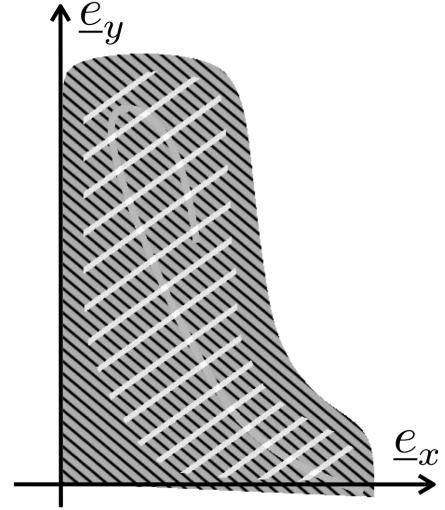


Figure 2. Soundboard model. The black lines indicate the direction of the fibers. The white lines represent the ribs. The grey curve is the bridge.

the instrument is computed in a virtual box bounded by absorbing boundary conditions (or Perfectly Matched Layers), thus simulating a free space as in an anechoic chamber [7]. The rim is taken into account as a rigid obstacle for the sound waves, but its vibrations are ignored. The coupling between the soundboard and the sound field is governed by the continuity of the normal velocity on the soundboard surface, as done for previous modeling of timpani [8] and guitar [9].

2.2 Summary of the numerical methods

Specific methods are used for the discretization of each subsystem and of the coupling terms. To ensure long-term stability, the numerical schemes are based on the formulation of a discrete energy, consistent with the continuous energy of the considered system, which is either constant or decreasing with time. In addition, the discrete formulation of the coupling terms is conservative and fulfils the reciprocity principle. For the strings, two different schemes were adopted: one for the linear part, and another one, specifically developed [2], for the nonlinear part. This new scheme is applicable to a special class of equations called “Hamiltonian system of wave equations”. For the linear part, two different implicit θ -schemes are used: one, conditionally stable with reduced dispersion, for the transverse wave, and another one, unconditionally stable, for both the longitudinal and shear waves. Although the numerical dispersion is higher in this latter case, it has only little consequences since most partials of these waves are beyond the audio range. The stability condition applicable to the linear part yields a condition for the time step, which is selected here equal to $\Delta t = 10^{-6}$ s. For the hammer-string coupling, a nonlinear three time steps formulation is used. For the soundboard, a modal decomposition is made once for all (for given geometry and material properties) followed by a semi-analytic time resolution of the obtained

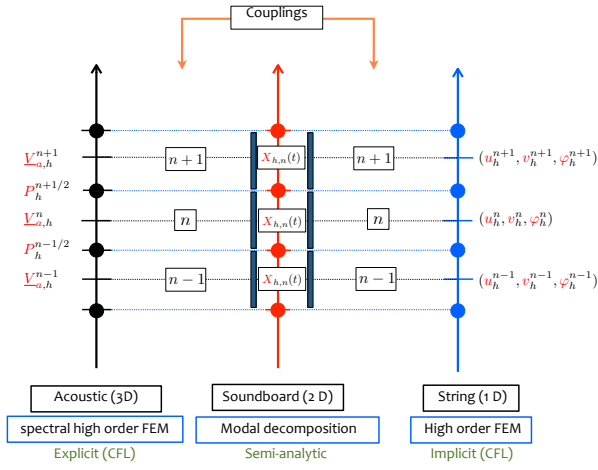


Figure 3. The discretization scheme of the piano model is summarized. The three discrete components of the strings' motion (u_h, v_h, φ_h) and the acoustic velocity $V_{a,h}$ are calculated on the time grid $n\Delta t$. The soundboard modal displacements $X_{h,p}$ and the sound pressure P_h are calculated on the time grid $[n + 1/2]\Delta t$ so as to ensure conservation of the discrete energy.

set of second-order differential equations, where damping is added mode by mode, as done in the past for the guitar [9]. This amounts to assuming that the modal damping matrix is diagonal. The values of the damping terms are extracted from the existing literature and from our own measurements. In practice, 2400 modes have to be calculated for a Steinway D soundboard in the range 0 to 10 kHz. The string-soundboard coupling equations at the bridge are obtained by considering the continuity of the vertical velocity and nullity of the horizontal velocity at the bridge for both subsystems. This allows to couple the discrete string scheme with the semi-analytic soundboard model. In practice, the strings and soundboard unknowns are evaluated on interleaved grids: $\{n\Delta t\}$ for the strings, and $\{n + 1/2\Delta t\}$ for the soundboard (see Figure 3). Finally, for the acoustic propagation, higher-order finite-elements are used. The acoustic space is artificially bounded by Perfectly Matched Layers [7]. Again, the acoustic velocity and sound pressure are evaluated on interleaved grids (see Figure 3). Figure 4 shows an example of computation for the note C2, at time $t = 3.21$ ms after the hammer blow. The upper figure shows the transverse string displacement, while the longitudinal displacement is represented by a color scale within the string. The lower figure shows together the displacement of the soundboard and the pressure field in two planes perpendicular to the soundboard and crossing at the attachment point of the C2-strings on the bridge.

3. ANALYSIS OF SIMULATIONS

3.1 Frequency domain: phantoms and soundboard modes

The effects of transverse-longitudinal coupling in string motion can be seen in Figure 5 which shows the spectrum

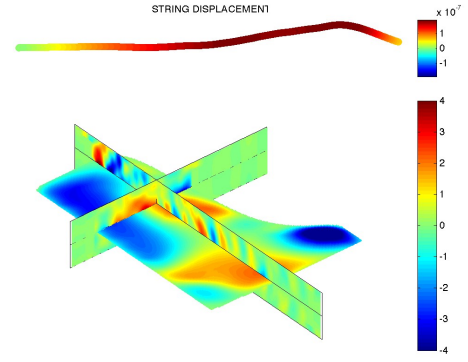


Figure 4. Example of computation for the note C2, at time $t=3.21$ ms after the attack. The figure at the top shows the transverse shape of the string displacement, while the longitudinal component is represented by a color scale within the string. The figure at the bottom shows the displacement of the soundboard and the pressure field in two planes perpendicular to the soundboard crossing at the attachment point of string C2 on the bridge. The lower scale refers to the sound pressure (in Pa).

of the longitudinal component of string F3 (fundamental $f_1 = 175$ Hz), for soft (piano) and strong (forte) hammer blow. The soft attack corresponds to an initial hammer velocity $V_H=0.5$ m/s, leading to a maximum amplitude of the string's displacement equal to 0.34 times its diameter. For the strong attack, we have $V_H=3$ m/s, and a maximum amplitude to diameter ratio equal to 2.1. In both cases, the dominant frequency is the fundamental longitudinal frequency at 2.64 kHz. The coupling is attested by the presence of transverse components in the spectra. As expected, the density of phantom partials is higher for the strong attack. These partials can be seen as additional components between the transverse components, especially in the range 2 to 4 kHz. Accurate frequency analysis shows that the frequencies of these partials are combinations of transverse and longitudinal frequencies due to quadratic and cubic nonlinearities, and are governed by simple arithmetic rules [1]. However, the question of existence of such partials is not trivial and requires a thorough stability analysis [10]. Similar phenomena are observed in other percussive instruments subjected to geometrical nonlinearities, such as gongs and cymbals [11]. Most of these frequencies could not be seen in the case of string without stiffness. They become visible because of the inharmonicity due to stiffness. As stated above, the relative magnitude of both these longitudinal and phantom frequencies in soundboard vibrations and sound pressure critically depend on the coupling conditions at the bridge. Examples of simulated sound pressure spectra are shown in Figure 6 for three different hammer impact velocities corresponding to piano, mezzo-forte and forte playing. Added components below 1 kHz are present, with identical frequencies and similar relative magnitude with regard to the string's partials in all three cases. Accurate spectral analysis shows that these components correspond to soundboard modes.

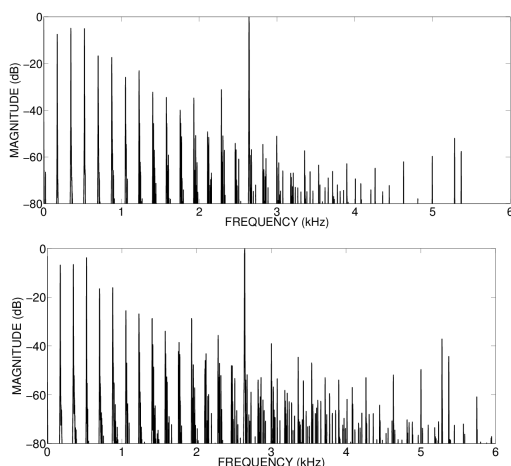


Figure 5. Spectrum of the simulated longitudinal component of the F3 string's displacement for soft (top) and strong (bottom) attack. The longitudinal frequency (at 2.64 kHz) is dominant, and phantom partials can be seen even for a soft attack. The pressure scaling at 0 dB is obtained by taking the magnitude of the strongest partial as reference.

The total bandwidth (within a dynamic range of 100 dB) increases with the initial hammer velocity, from 5 kHz, for the “piano” touch, to 7 kHz for the “forte” touch. This effect is due to the nonlinearity of the hammer felt. Between 1 and 5 kHz, the nonlinearity of the string is responsible for additional frequencies situated between the string's partials and visible as black regions in Figure 6. These frequencies are the so-called “phantom” partials due to the presence of quadratic and cubic terms in the Taylor expansion of both the transverse and longitudinal string force. Zooming on these zones allow accurate measurements of these phantom frequencies, who correspond to combinations (sums and differences) of the eigenfrequencies of the string [1]. In the “forte” case, for this string D \sharp 1, the amplitude of the phantoms is particularly high around 1.2, 1.7, 2.3, 2.8, 3.3, 3.7 and 3.9 kHz. The comparison of Figure 6 with the spectral analysis of sound pressure recorded in the vicinity of a Steinway D grand piano shows similar aspects and some differences. The presented spectra are obtained with three successive attacks, from “piano” to “forte”. In this case the hammer velocity was not measured, but the experimentally observed hammer forces are comparable to the simulations. Again, soundboard modes are visible in the low-frequency range (below 800 Hz). The total bandwidth also increases from bottom (soft impact) to the top (strong impact). However, the bandwidth is reduced, compared to the simulations, growing from 3 to 5.5 kHz (for a dynamic range equal to 100 dB). This discrepancy might be due to an underestimation of the internal damping in strings and soundboard, and/or underestimation of the radiated sound power. In the low-frequency range, it is hard to isolate the vibrations of the soundboard from the rest of the instrument. Therefore, it is perfectly conceivable that a part of the soundboard energy is being transmitted to other elements, such as the rim. Phantom partials are also observed

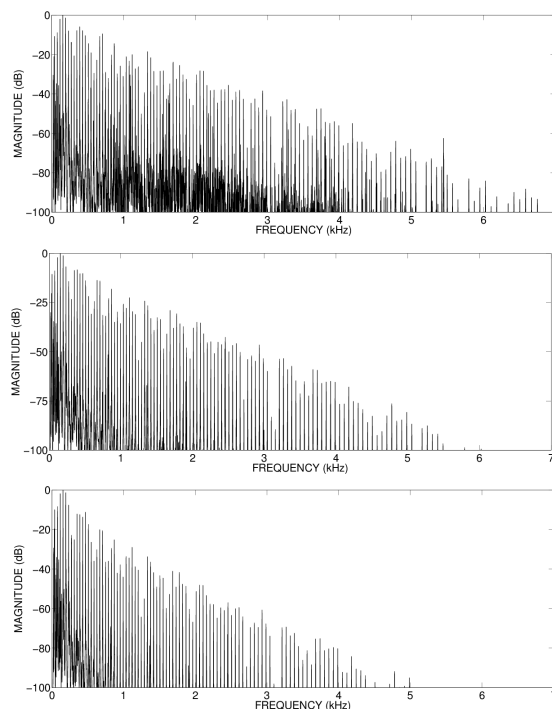


Figure 6. Spectra of simulated sound pressure in the piano near field. Note D \sharp 1 (single string) at three different levels: forte ($V_H=3$ m/s ; top), mezzo-forte ($V_H=1.5$ m/s ; middle) and piano ($V_H = 0.5$ m/s ; bottom). The pressure scaling at 0 dB is obtained by taking the magnitude of the strongest partial as reference. The spectra are averaged during the first 1.6 s of the sound.

in the three cases, with increasing relative amplitudes from “piano” to “forte” sounds. Maxima of the phantoms amplitudes are observed around 1.2, 1.7, 2.0, 2.3, 2.8, 3.3 and 3.9 kHz, which is very similar to the situation observed in the simulations. One main difference here is that the level of the phantoms are significant, even for a soft impact. This might be either due to differences in bridge transmission and/or to higher string amplitude than in the model. Further work is needed here for explaining these results.

3.2 Time-domain: precursors

The effect of string amplitude is also visible in the time-domain on the pressure waveforms (see Figure 8). The small precursor observed for a soft impact is due to the physical dispersion consecutive to string's stiffness. The amplitude of the precursor increases with hammer velocity, and its spectral content changes: frequency analysis shows a spectrum comparable (mutatis mutandis for D \sharp 1) to the longitudinal component in Figure 5 with a noticeable longitudinal fundamental frequency at 544 Hz. The consequence of decreasing bandwidth is visible on the waveforms which become smoother as the initial hammer velocity decreases. Such observations can be made for both measured and simulated pressure waveforms. Comparison between measurements and simulations show that the positions and relative magnitudes of the main pulses are fairly well reproduced, although the details of the wave-

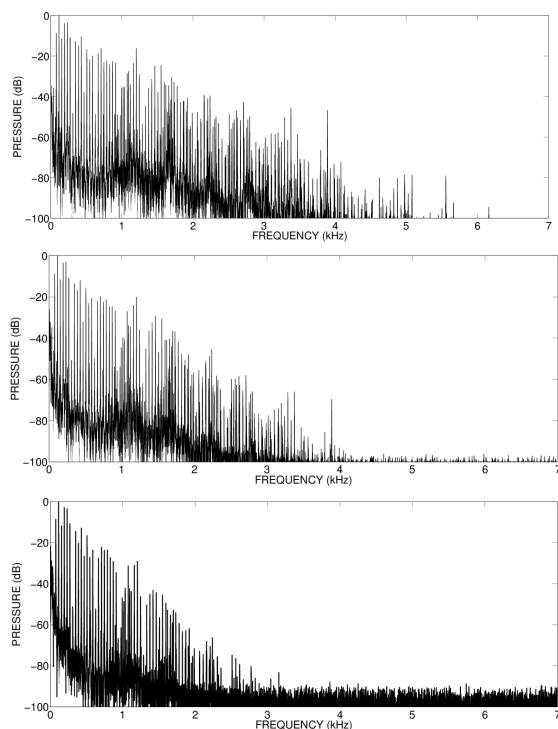


Figure 7. Spectra of measured sound pressure in the vicinity of a Steinway D grand piano. Note D \sharp 1 (single string) at three different levels: forte (top), mezzo-forte (middle) and piano (bottom). The pressure scaling at 0 dB is obtained by taking the magnitude of the strongest partial as reference. The spectra are averaged during the first 1.6 s of the sound.

forms differ. This might be due to phase shifts, since the exact location of the microphone is not perfectly known. More experiments are necessary to check the pertinence of this assumption.

4. BOUNDARY CONDITIONS AT THE BRIDGE

4.1 Experiments

One important limitation of the above presented model is due to the fact that it does not allow any exchange of energy between the two transverse components of the string. As shown in Figure 10, such exchange is currently observed on pianos, even for single strings. Assuming perfect constant circular cross-section and homogeneity, then such exchange can only be driven by the boundary conditions. In this study, we made the *a priori* hypothesis that energy exchange can be due to two essential mechanisms: a rocking motion of the bridge, and particular geometry of the string-bridge contact. In order to isolate both mechanisms, a special monochord has been designed on purpose: it consists of a single string passing over a piece of wood representing a portion of “bridge” and glued on an arched flexible beam that plays the role of the “soundboard”. Two photodetectors HOA1877 situated close to the string end deliver signals proportional to the two transverse components of the string’s displacement. The strings are excited by the blow of mallets with variable hardness. A first preliminary series

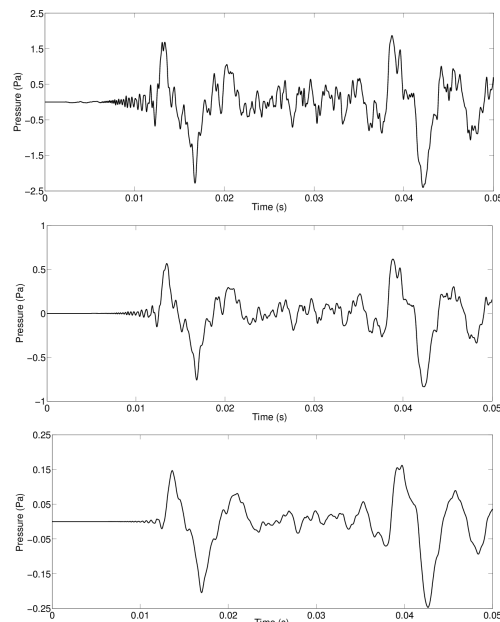


Figure 8. Simulation D \sharp 1. Pressure waveforms during the first 50 ms of the sound. Same hammer velocities as in Figure 6. (Top) forte ; (middle) mezzo-forte ; (bottom) piano.

of experiments is conducted with a “V-shaped” boundary condition, where the “V” is situated in the symmetry plane of the “bridge” (see Figure 11). This prevent us against any change of polarization due to rocking motion of the bridge. As expected, the result shows that a vertical blow induces a vertical motion of the string and that no horizontal component is observable during the 5 s decay (see Figure 12). Another series of experiments is conducted with a so-called “zig-zag” end conditions observed on many pianos. It consists in two needles embedded in the bridge apart from one another in the direction of the string, forming an angle α with the vertical plane in opposite directions (see Figures 13 and 15). Here again, the holes for the needles are drilled in the symmetry plane of the bridge, so as to prevent any influence of rocking motion. In this case, the presence of a horizontal component is observed in the string motion after some time. The energy exchange varies with the angle α . For $\alpha = 90$ degrees, the end of the string is blocked. For $\alpha = 0$, the situation is particular since the string is not fixed at all. In both cases, an initial vertical blow does not induce any horizontal component. For intermediate values, typically for α within the range 20 to 60 degrees, then an horizontal component clearly appears (see Figure 14).

4.2 Modeling and simulation

A simple model of the zig-zag end condition is built as an attempt to explain the observed change of polarization of the string. As shown in Figure 15, the string is allowed to move along the needle. The model also allows a possible loss of contact between the string and the needle. In the absence of friction, the reaction of the needle is perpendicular to it. The model is tested first without friction

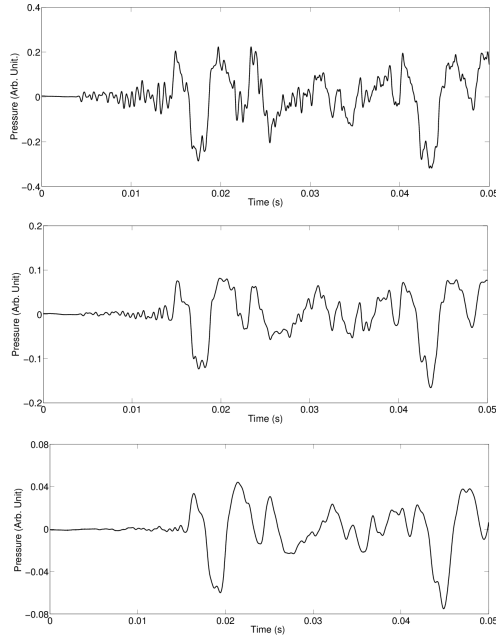


Figure 9. Measurements D#1. Steinway D. Recorder pressure waveforms during the first 50 ms of the sound. Similar hammer forces as in Figure 8. (Top) forte ; (middle) mezzo-forte ; (bottom) piano.

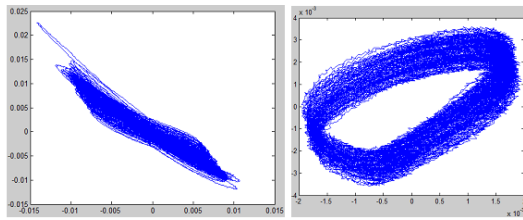


Figure 10. Observed change of polarization on a piano string. The photodetectors form an angle of 45° with the vertical plane. The left figure shows the vertical motion of the string just after the blow, while the right figure shows its horizontal motion after several seconds.

and secondly with a friction force that limits the possibility of slipping along the needle. To simplify the model, the speaking length of the string is represented by a mass-spring system composed of two identical masses M and N , and three springs of identical stiffness (see Figure 16). These lumped parameters are adjusted so as the eigenfrequencies of this 2-dof system are of the order of magnitude of the first partials of the continuous string. The masses are allowed to move in the x , y and z directions, although the z -component (longitudinal motion) can be neglected compared to the two others (see Figure 15). The first bend of the zig-zag end condition is modeled at point Z . A last spring with same stiffness as in the lumped string is simulated between Z and the point P . This last point represents the second bend of the zig-zag and is assumed to be fixed. The model is solved numerically using finite differences for the differential equations of the oscillators, and a predictor-corrector method for the boundary conditions. A typical result can be seen in Figure 17 showing the mo-

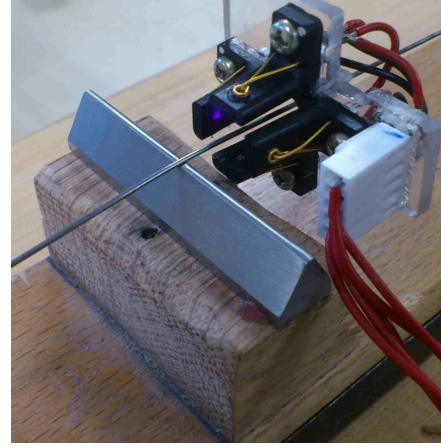


Figure 11. Experimental set-up showing the string, the photodetectors, the V-shaped boundary condition, the bridge and the beam.

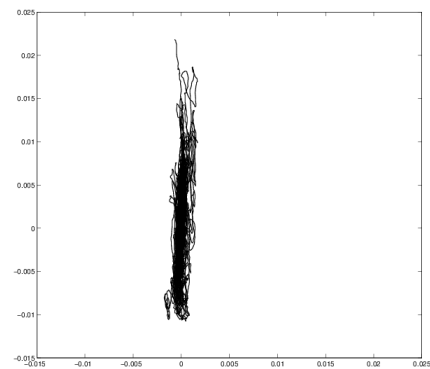


Figure 12. Constant polarization of the string observed for a V-shaped boundary condition place in the symmetry axis of the bridge. The photodetectors are situated in the vertical and horizontal planes, respectively. The duration of the observed decay is 5 s.

tion of the mass M . An initial vertical velocity is imposed (string motion in blue). The polarization axis then turns progressively (motion in green). After some time, the locus of the mass is elliptic (in red) and the orientation of the largest axis of the ellipse remains constant. The simulations show in addition that the main axis of this ellipse remains closer to the vertical axis if the angle α increases. This is coherent with the experimental observations. The horizontal component shown in this simulation is about three times larger than the one observed in Figure 14. This might be due to the fact that we observe here the motion of a point at one-third of the string, whereas the experimental data were recorded close the end, because of limited linearity range of the photodetectors. More accurate simulations must be conducted with a larger number of lumped elements in order to compare better the results with the experiments. The simulation of the string's motion at point Z under the needle shows that this point slips most of the time along the needle (in green), with a few erratic jumps (loss of contact) of small amplitude at the end of the sequence

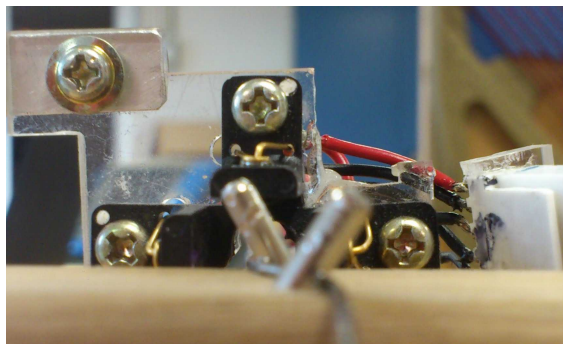


Figure 13. Zig-zag end condition.

(see Figure 18). One conclusion from this simple model is that the string might be subjected to micro-slipping at the zig-zag end at each arrival of the transverse pulses. The amplitude of these slipping depend on the friction. As a consequence, the string progressively takes an horizontal transverse component, and thus its polarization changes. In a piano, this horizontal component “sees” an input admittance at the bridge which is usually much higher than the admittance seen by the transverse vertical component. As a result, the vertical component is damped quicker and the horizontal component is dominant at the end of the tone.

5. CONCLUSION

The presented model of a grand piano is able to reproduce the main features of piano sounds and vibrations. The non-linear string model and the string-soundboard coupling, in particular, account for both the dependence of tone quality upon string amplitude and transmission of longitudinal components to the rest of the instrument. This allows us to simulate the spectral richness of piano sounds efficiently, including the presence of soundboard modes and phantom partials. The frequencies of these partials are predicted with great accuracy. However, some discrepancies exist on the prediction of their amplitudes, which incites us to reconsider the string-soundboard coupling conditions at the bridge in the future. Auditory evaluation of the simulated sound pressure shows, in addition, that the model fails in reproducing the bass/treble balance of the lowest notes accurately. More investigations are needed at this stage to properly evaluate the internal causes of losses in strings and soundboard, and reconsider the possible transmission of energy from the soundboard to other parts of the body (rim, keybed) in this frequency range. Another limitation of the model is that it does not allow energy exchange between the two transverse polarizations of the string. The experiments and simulations presented here show that the zig-zag end condition settled on many pianos might be a cause for such exchange. With the help of a simple lumped model, we were able to reproduce, at least qualitatively, the main properties observed on real strings. The rocking motion of the bridge also might a possible candidate at the origin of polarization change: its influence will be examined in the near future.

Acknowledgments

The authors wish to thank Lahcène Cherfa for the development of the monochord, Marc Duruflé for his valuable help in the simulations, and René Caussé for making the Steinway D piano available to us for the measurements.

6. REFERENCES

- [1] J. Chabassier, A. Chaigne, and P. Joly, “Modeling and simulation of a grand piano,” *J. Acoust. Soc. Am.*, 2013, (submitted).
- [2] J. Chabassier and P. Joly, “Energy preserving schemes for nonlinear Hamiltonian systems of wave equations. Application to the vibrating piano string,” *Computer Methods in Applied Mechanics and Engineering*, vol. 199, pp. 2779–2795, 2010.
- [3] J. Chabassier and S. Imperiale, “Stability and dispersion analysis of improved time discretization for simply supported prestressed Timoshenko systems. Application to the stiff piano string,” *Wave Motion*, 2012.
- [4] H. A. Conklin, “Generation of partials due to nonlinear mixing in a stringed instrument,” *J. Acoust. Soc. Am.*, vol. 105, no. 1, pp. 536–545, 1999.
- [5] A. Stulov, “Dynamic behavior and mechanical features of wool felt,” *Acta Mechanica*, vol. 169, no. 1, pp. 13–21, 2004.
- [6] P. Morse and K. Ingard, *Theoretical Acoustics*. Princeton, New Jersey: Princeton University Press, 1968, ch. 14, pp. 856–863.
- [7] E. Bécache, S. Fauqueux, and P. Joly, “Stability of perfectly matched layers, group velocities and anisotropic waves,” *Journal of Computational Physics*, vol. 188, pp. 399–403, 2003.
- [8] L. Rhaouti, A. Chaigne, and P. Joly, “Time-domain modeling and numerical simulation of a kettledrum,” *J. Acoust. Soc. Am.*, vol. 105, pp. 3545–3562, 1999.
- [9] G. Derveaux, A. Chaigne, P. Joly, and E. Bécache, “Time-domain simulation of a guitar: Model and method,” *J. Acoust. Soc. Am.*, vol. 114, pp. 3368–3383, 2003.
- [10] A. H. Nayfeh and D. T. Mook, *Nonlinear oscillations*, ser. Wiley classics library. Hoboken, New Jersey: J. Wiley, 1979.
- [11] C. Touzé and A. Chaigne, “Lyapunov exponents from experimental time series: Application to cymbal vibrations,” *Acustica united with Acta Acustica*, vol. 86, no. 3, pp. 557–567, 2000.

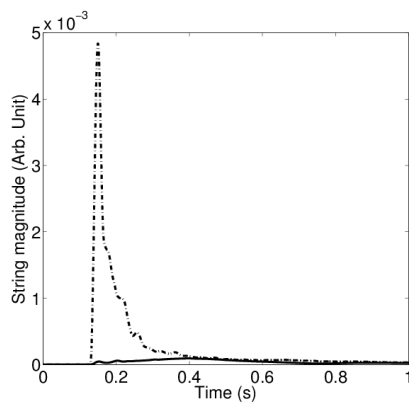
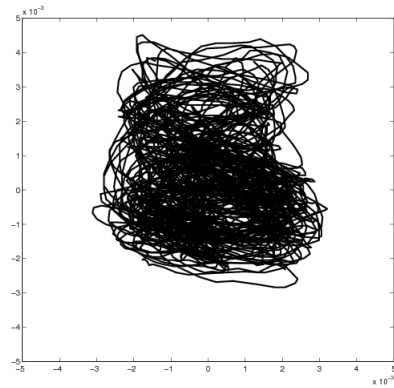
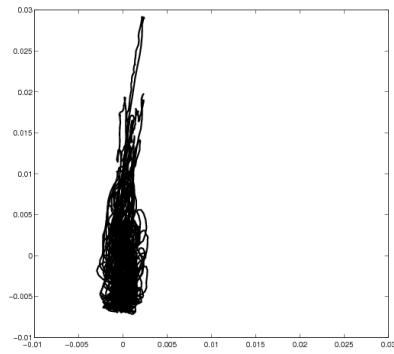


Figure 14. Zig-zag end condition. Comparison between string polarization during the first 100 ms (top) and after 0.35 s (middle) showing a progressive counterclockwise orientation of the axis to the string motion. The figure at the bottom displays the evolution of the envelope of both polarizations with time, showing a progressive increase of the horizontal component (solid line) during the first 200 ms of the string's oscillation. At time $t=0.4$ s, the magnitude of both components are comparable.

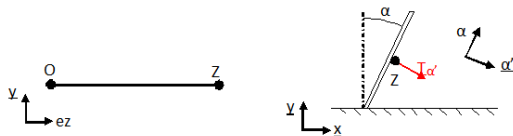


Figure 15. Model of the zig-zag end.

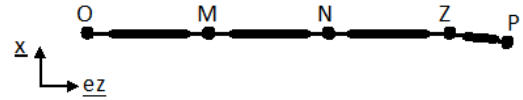


Figure 16. Discrete model of the string with one zig-zag end condition (top view).

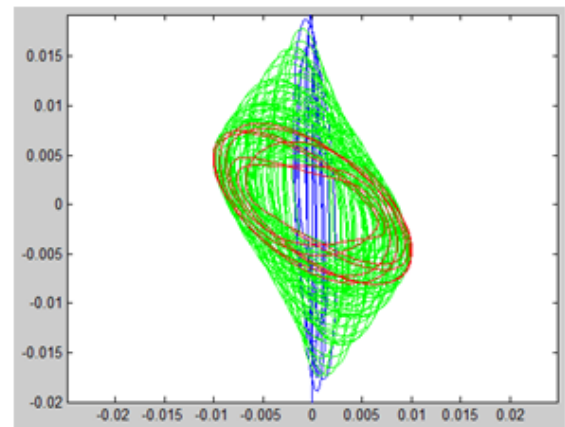


Figure 17. Simulated change of polarization of the simulated string with one zig-zag end. Motion of the mass M after the blow (in blue), during transition (in green), and after a few seconds (in red).

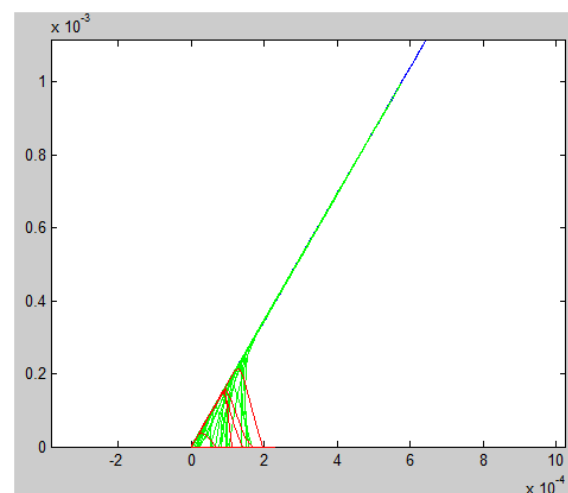


Figure 18. Simulated motion of the string close to the needle (Z-point in Figure 16).

Large Scale Physical Modeling Sound Synthesis

**Stefan Bilbao, Brian Hamilton,
Alberto Torin and Craig Webb**
Acoustics Group, University of Edinburgh
sbilbao@staffmail.ed.ac.uk

**Paul Graham, Alan Gray,
Kostas Kavoussanakis and James Perry**
EPCC, University of Edinburgh

ABSTRACT

Sound synthesis based on physical models of musical instruments is, ultimately, an exercise in numerical simulation. As such, for complex systems of the type seen in musical acoustics, simulation can be a computationally costly undertaking, particularly if simplifying hypotheses, such as those of traveling wave or mode decompositions are not employed. In this paper, large scale time stepping methods, such as the finite difference time domain and finite volume time domain methods are explored for a variety of systems of interest in musical acoustics, including brass instruments, percussion instruments based on thin plate and shell vibration, and also their embeddings in 3D acoustic spaces. Attention is paid here to implementation issues, particularly on parallel hardware, which is well-suited to time stepping methods operating over regular grids. Sound examples are presented.

1. INTRODUCTION

Physical modeling sound synthesis has been approached in a variety of ways; perhaps the best known methods are the lumped mass spring network methodology, developed by Cadoz and associates [1], modal synthesis, developed at IRCAM [2] and digital waveguide methods developed by Smith at CCRMA [3], and subsequently greatly extended [4, 5, 6]. A direct comparison of these methods is difficult (but possible [7])—all possess distinct advantages: lumped network methods allow an extreme degree of control over the system, to the level of individual masses and springs; modal methods offer the possibility of exact solutions when modal data is available in closed form, or easily computed, and waveguides are extremely efficient when the system under consideration behaves (nearly) as the 1D wave equation, which is roughly true for a variety of systems of interest in musical acoustics.

Direct numerical simulation techniques, essentially time stepping methods such as the finite difference time domain method (FDTD) [8], or finite volume time domain method [9], though mainly used in other mainstream domains can also be applied to the problem of sound synthesis; indeed, they were proposed as long ago as 1969 by Ruiz [10], in the

case of the vibrating string, and, in the scattering context by Kelly and Lochbaum [11] in the case of the vocal tract, and have been applied to a variety of problems, particularly in the musical acoustics setting, first by Chaigne and associates [12, 13, 14], and then others for synthesis [15]. Such brute force methods are often more computationally costly than the other techniques mentioned above; and yet, one advantage they possess is generality, in the sense that one may approach a great variety of systems, without the need for considering simplifying hypotheses (such as, e.g., linearity, or the availability of a modal decomposition or efficient traveling wave decomposition).

Until recently, it has been infeasible, computationally, to approach very large scale physical modeling synthesis using time stepping methods; however, the emergence into the mainstream of parallel hardware such as, e.g., general purpose graphical processing units (GPGPUs), has allowed the exploration of more complex systems. A project currently under way at the University of Edinburgh (NESS) is an exploratory attempt at producing synthetic sound for a variety of model systems, ultimately in 3D. Though not real time, computation time is becoming reasonable for certain systems; for other, particularly large systems (such as, e.g., the emulation of room acoustics), it is only now becoming possible to perform such simulations at all at audio rates, and on relatively cheap commercial hardware. GPGPUs offer one approach to optimisation among many; whilst it would be feasible to use large-scale grid computing to run such simulations, the emphasis here is on hardware that is generally available for desktop computing. The overall aim of the project is to develop systems which can be used by musicians and composers, rather than to run simulations on a supercomputing machine.

There are many complications in designing such systems. Some are the usual difficulties in working with time stepping methods, such as, e.g., determining sufficient conditions for numerical stability, particularly under nonlinear conditions (the case of most interest in musical sound synthesis), or determining appropriate numerical boundary conditions when the geometry of the object under consideration is not simple. Others involve a reexamination of the properties of general numerical methods when audio synthesis is the goal, leading to various algorithmic constraints. Finally the question of parallelizability must be addressed if one is to eventually make the most of parallel hardware such as GPGPUs. There has been a good deal of recent work in modeling of room acoustics on GPGPU [16, 17, 18], but sound synthesis, both through physical

modeling [19] and other methods [20] has also been approached.

In this article, the emphasis is on design and implementation issues for FDTD methods for sound synthesis, rather than on musical acoustics. To this end, a simple family of model problems is introduced in Section 2. A brief overview of simple time stepping methods is given in Section 3, followed by a discussion of algorithm design issues, geared towards audio applications, in Section 4, and further parallelization issues on GPU in Section 5. Simulation results are presented in Section 6.

2. MODEL SYSTEMS

2.1 Second Order Systems

As a representative system, consider an object for which the dynamics are described by a partial differential equation of the form

$$\frac{\partial^2 U}{\partial t^2} = F \quad (1)$$

Here, t is a time variable, and $U = U(t, \mathbf{x})$ is the variable of interest to be solved for, such as a displacement of a bar, or plate, or pressure in an acoustic tube, and \mathbf{x} is a spatial coordinate in d dimensions (normally $d = 1, 2$ or 3); generally the object is defined over a region $\mathbf{x} \in \mathcal{V} \subset \mathbb{R}^d$. F generally depends on U and its spatial derivatives (usually even) or temporal derivatives, or their combinations.

Such a model equation, though extremely simple, serves as a good first approximation to various systems of interest in musical acoustics. For example, consider the choices of F of

$$F = c^2 \nabla^2 U \quad (2)$$

$$F = -\kappa^2 \nabla^2 \nabla^2 U \quad (3)$$

$$F = \frac{c^2}{S(x)} \frac{\partial}{\partial x} (S(x)U) \quad (4)$$

$$F = c^2 \left(1 + \alpha \int_{\mathcal{V}} \left(\frac{\partial U}{\partial x} \right)^2 dx \right) \frac{\partial^2 U}{\partial x^2} \quad (5)$$

Here, c , κ and α are constants, and ∇^2 is the d -dimensional Laplacian operator, defined as

$$\nabla^2 = \sum_{\eta=1}^d \frac{\partial^2}{\partial x_{\eta}^2} \quad (6)$$

Equation 2 is the wave equation; when $d = 1$, it serves as an approximation to the vibration of a uniform string, where U is transverse displacement, and to the dynamics of a uniform acoustic tube, where U is pressure. When $d = 2$, it corresponds to the vibration of an ideal membrane, where U is transverse displacement, and when $d = 3$, it approximates wave propagation in an acoustic enclosure, where U is a variable such as a pressure or velocity potential.

Equation 3 is a linear model of vibration of a thin uniform bar ($d = 1$) or plate ($d = 2$), where U is transverse displacement.

Equation 4 is sometimes referred to as Webster's equation [21], and is a lossless 1D model of wave propagation, at

speed c , in a tube of variable cross section $S(x)$, where U represents pressure or velocity potential.

Equation 5 is perhaps the simplest possible distributed nonlinearity in musical acoustics, sometimes referred to as the Kirchhoff-Carrier equation [22, 23], and is used to model tension modulation effects [24] in strings, where again, U represents transverse displacement.

The model systems presented above are extremely simple, and in fact much too simple for good quality synthesis. For brevity, various features have been neglected here:

All the systems above are lossless; real systems possess a variety of internal dissipation mechanisms, such as thermal/viscous effects, and also transfer energy to their surroundings through radiation. In fully 3D synthesis models, radiation is neatly handled through direct coupling to the acoustic field; modeling of thermal/viscous effects, however, is much more involved using time stepping methods. See Section 6.1 for more on this topic.

Boundary conditions have not been specified here; in some cases in musical acoustics, these can be trivial (as in, e.g., the case of an ideal rigidly terminated string), but in others can require extreme care as in, e.g., free edges of thin structures such as cymbals or gongs, radiation impedance conditions in 1D models of wind instruments, and absorbing boundary conditions in 3D acoustic simulations.

Equation 1 represents the behaviour of an unforced system; that is to say, it is lacking, as yet, an excitation term. For most systems of interest, the excitation is nonlinearly dependent on U , but acts at a single location (or very small region) on the object; nonlinearity in F itself has a much greater impact on analysis and numerical design for the system, and is sometimes referred to as a distributed nonlinearity.

Not all systems of interest take the above form—one example is the system describing nonlinear wave propagation in an acoustic tube, which is most naturally written as a first order system [9]. In a modular setting, when dealing with systems in contact with the acoustic field, and also in more elaborate models of stiff systems such as bars and plates the variable U is coupled to other variables—alternatively, in such cases, one could interpret U as a vector variable.

3. TIME STEPPING METHODS

The first approximation necessary is discretization in time. A simple choice, and a natural one in audio applications, is an approximation at equally spaced intervals of k seconds ($F_s = 1/k$ is the sample rate), and thus in discrete time, $u^n(\mathbf{x})$ represents an approximation to $U(nk, \mathbf{x})$, where n is an integer (the time index).

Considering first the second time derivative operator, a multitude of approximation methods are available, and in particular multistep methods such as Runge Kutta, Adams Bashfort, etc. [25]. Given that in large simulations in parallel hardware, memory use may be a bottleneck, depending on accuracy requirement it may be useful to make use of the simplest possible approximation:

$$\frac{\partial^2 U}{\partial t^2} \Big|_{t=nk} \implies \frac{1}{k^2} (u^{n+1} - 2u^n + u^{n-1}) \quad (7)$$

Such an approximation, though only second order accurate [26] is minimal in terms of memory usage. (For some systems, particularly if there are long memory effects associated with viscothermal losses requiring more memory, then the need for such simple approximations is less urgent—see Section 6.1.)

The character of the resulting algorithm depends greatly on the discretization of F . If the simple approximation $F^n = f(U(nk, \mathbf{x}))$ is used, then an update for 1 will be of the form

$$u^{n+1} = 2u^n - u^{n-1} + k^2 f^n \quad (8)$$

and, after suitable discretization over a grid (see Section 3.1), will be fully explicit: at each time step, values of u^{n+1} at the next time step may be computed directly from previously calculated values of u and f at time steps n and $n - 1$.

In some cases, however, for reasons having to do with stability [26], and also in reducing artifacts resulting from numerical dispersion [15], it may be preferable to use a different approximation, such as, e.g., $f^n = \frac{1}{2} (f^{n+1} + f^{n-1})$; in this case, the update for 1 becomes

$$u^{n+1} = 2u^n - u^{n-1} + \frac{k^2}{2} (f^{n+1} + f^{n-1}) \quad (9)$$

which is implicit— u^{n+1} and f^{n+1} (which depends on u^{n+1}) must be computed together at each time step. If the system under consideration is linear, then this amounts to a linear system solution, which often complicates implementation in parallel hardware. If the system is nonlinear, then existence and uniqueness issues in the update above may appear. For more on issues related to implicit methods, see Section 3.2.

3.1 Grids

In the FDTD setting, numerical solutions are often represented over uniform grids—see Figure 1, showing simple Cartesian grids in 1D, 2D and 3D. Working over such regular grids has, of course, strengths and weaknesses. The most significant benefit is in terms of parallelizability, particularly when the problem under consideration is uniform over space, as the action of a given approximation to a differential operator is the same at each point in the domain. Such is the case, for example, for membranes and plates of constant thickness, in tubes of uniform cross section, and for 3D acoustic wave propagation. Difficulties emerge, however, when treating irregular boundaries, both in terms of analysis, as well as in parallel implementation. The former difficulty, however, can be addressed by an appeal to methods over unstructured grids, such as, e.g., finite volume techniques [9]; see [27] for more on this topic in the case of the 3D wave equation and room acoustics applications.

Supposing operation over a regular Cartesian grid, of spacing h between adjacent grid points, then grid functions \hat{u} are fully discrete approximations to the solution U . For example, the grid functions \hat{u}_l^n , $\hat{u}_{l,m}^n$ and $\hat{u}_{l,m,p}^n$ are approximations to the solution $U(t, \mathbf{x})$ at time $t = nk$, and at locations $\mathbf{x} = lh$, $\mathbf{x} = [l\ m]h$ and $\mathbf{x} = [l\ m\ p]h$ in 1D, 2D and 3D respectively, for integer l , m and p .

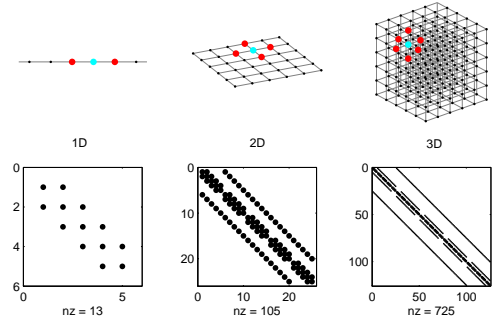


Figure 1. Top row: Cartesian grids in 1D, 2D and 3D. At a given grid point (in blue), the family of neighboring points over which a simple approximation to the Laplacian operates is indicated in red. Bottom row: sparsity plots for Laplacian approximations, when the operation is represented as a matrix multiplication.

Difference operations approximating differential operators are straightforward to obtain. Consider the Laplacian operator, which plays a central role in many physical models. The simplest approximations in 1D, 2D and 3D, written here as $\delta_{\nabla^2}^{(1D)}$, $\delta_{\nabla^2}^{(2D)}$ and $\delta_{\nabla^2}^{(3D)}$, respectively, are

$$\delta_{\nabla^2}^{(1D)} u_l^n = \frac{1}{h^2} (u_{l+1}^n - 2u_l^n + u_{l-1}^n) \quad (10)$$

$$\delta_{\nabla^2}^{(2D)} u_{l,m}^n = \frac{1}{h^2} (u_{l+1,m}^n + u_{l-1,m}^n + u_{l,m+1}^n + u_{l,m-1}^n - 4u_{l,m}^n) \quad (11)$$

$$\delta_{\nabla^2}^{(3D)} u_{l,m,p}^n = \frac{1}{h^2} (u_{l+1,m,p}^n + u_{l-1,m,p}^n + u_{l,m+1,p}^n + u_{l,m-1,p}^n + u_{l,m,p+1}^n + u_{l,m,p-1}^n - 6u_{l,m,p}^n) \quad (12)$$

Such approximations employ nearest neighbours only on a regular grid—see Figure 1 for a graphical representation of the region of operation of such operators (or stencils) over Cartesian grids.

3.2 Sparse Vector Matrix Representations and Recursions

For computing purposes, is often convenient to represent the grid functions *hatu*, which are multidimensional arrays, as vectors $\hat{\mathbf{u}}$ —see Figure 2, showing concatenation of columns of a 2D array into a vector. In this formalism, difference operators are represented as matrix multiplications, where, due to the local character of finite difference operations, the matrices are sparse. See the bottom row of Figure 1, showing sparsity plots of matrix representations of the Laplacian operator, in 1D, 2D and 3D.

Using this formalism, in many cases, it is then possible to write recursions such as (8) and (9) in vector matrix form as

$$\mathbf{A}\hat{\mathbf{u}}^{n+1} = \mathbf{B}\hat{\mathbf{u}}^n + \mathbf{C}\hat{\mathbf{u}}^{n-1} \quad (13)$$

where here, \mathbf{A} , \mathbf{B} and \mathbf{C} are update matrices incorporating the effects of various finite difference approximations over the grid, and are generally sparse.

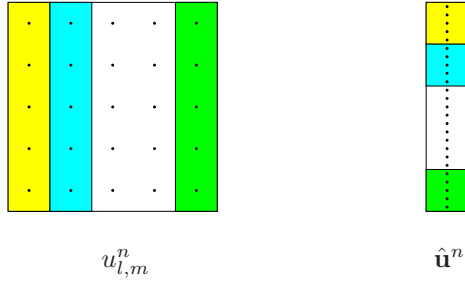


Figure 2. Reorganisation of a 2D array $u_{l,m}^n$ as a vector, by concatenation of columns.

Various special cases emerge at this point.

If the scheme is explicit, as in, say, the case of (8), then \mathbf{A} is simply the identity matrix, and thus the solution may be updated solely through sparse matrix multiplication operations (of \mathbf{B} and \mathbf{C} in (13)). This is the ideal case in a parallel realization.

If the scheme is implicit, but linear, then \mathbf{A} will not be the identity, and a linear system solution will be required in the loop in order to solve for $\hat{\mathbf{u}}^{n+1}$. However, as it is constant, it is possible to some perform preconditioning once, offline.

If the scheme is implicit, and also nonlinear, then it may not be possible to write a recursion such as the above—in some cases, and particularly when the nonlinearity is of a simple form (such as, e.g., cubic, as occurs frequently in models of vibration of strings and plates, and as exhibited in (5)), then one may arrive at such a form—but generally \mathbf{A} is dependent on previously computed values of the solution $\hat{\mathbf{u}}^n$, and thus must be constructed anew at each time step. This is certainly the most challenging case in a parallel realization.

4. ALGORITHM DESIGN ISSUES

4.1 Linear System Solutions

For implicit methods, as described above, linear system solutions are necessary in updating the state of the system. For methods defined locally over a grid, such as FDTD, such matrices are sparse and possess a banded structure (see Figure 1).

Many efficient methods are available in approaching the solution of sparse banded systems; among the best known are methods such as the Thomas algorithm (and extensions) [26]; but such methods generally require diagonal dominance, which is not always the case for matrix representations of FDTD schemes. Also, they are inherently serial—one must proceed, step by step, along the bands of the matrix in order to arrive at a solution. For an efficient realization in parallel hardware, parallelizable methods must clearly be employed. There are many such methods available—particularly well-suited to sparse systems are iterative methods such as the conjugate gradient family, employing a sparse preconditioner (such as incomplete Cholesky factorization). The problem of determining this preconditioner (possibly anew at each time step) must be

weighed against the computational cost of determining the preconditioner.

In other cases, such as, e.g., those involving interpolation between distinct grids (as in the case of systems coupled to the acoustic field), this banded structure may be lost; in compensation, however, the system to be solved may be strongly diagonally dominant, allowing the use of very simple iterative methods (such as, e.g., Gauss-Seidel [26]). For an example of such interpolation between grids in matrix form, see, e.g., [19].

4.2 Stability

For explicit schemes, which are of most interest in parallel implementation, stability conditions for schemes for the model system 1 are best framed in terms of a lower bound on the grid spacing h in terms of k , the time step (which, in audio applications, is normally set a priori): in other words,

$$h \geq h_{\min}(k) \quad (14)$$

One way of arriving at such conditions, for linear problems, is through the use of frequency domain techniques, or von Neumann analysis [26]; such methods, however, do not apply directly to nonlinear systems, nor do they allow the determination of sufficient stability conditions when boundary are included. As an alternative, methods based on discrete energy conservation/dissipation are employed [15, 28], which do allow such stability conditions to be determined under very general conditions.

4.3 Bandwidth Limitation

Given the lower bound on h in (14), it may be tempting to choose a value of h which is larger than the minimum, in the interest of reducing computational complexity; this, however, leads to severe limitations on output bandwidth—see Figure 3. Accompanying this is a phenomenon known as numerical dispersion, leading to a mistuning of modal frequencies; for this reason, in audio applications, it is generally best to choose the grid spacing as close to the stability condition as possible.

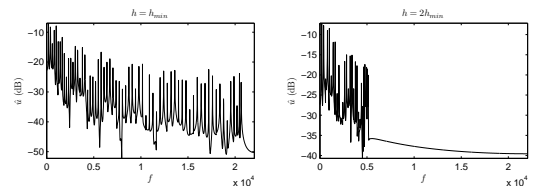


Figure 3. Spectrum of sound output from a simple stiff string physical model, at 44.1 kHz. Left: where the grid spacing is chosen at the minimum allowable value from the stability condition, and right, at twice the minimum value, illustrating a severe numerical cutoff at approximately 5 kHz.

4.4 Computational Complexity

Computational complexity, as expected, scales strongly with the dimension of the system; in particular, in 3D, computational costs for large volume simulations are extreme. See

Table 1 for typical computational costs for FDTD methods, in terms of memory required to hold the state, and floating point operations per second output, at a typical audio rate. Such costs motivate an examination of acceleration in parallel hardware, as discussed in the next section.

	Memory (B)	Flop/s
Tube, length 1 m	3.11×10^3	2.86×10^7
String, steel, length 1 m radius 1×10^{-3} m, tension 100 N	1.66×10^4	1.53×10^8
Plate, steel, area 1 m^2 , thickness 5×10^{-4} m	3.46×10^9	1.08×10^{10}
Small enclosure (1 m^3)	1.00×10^7	1.67×10^{11}
Large room (10^4 m^3)	1.00×10^{11}	1.67×10^{15}

Table 1. Memory requirements, in B, and floating point operations per second output for several typical systems, for standard FDTD schemes operating at $F_s = 44.1$ kHz, and using double precision arithmetic.

5. IMPLEMENTATION ON GPU

In this section, the focus is on the use of GPGPUs, using Nvidia's CUDA architecture.

The starting point for the implementation of the time stepping schemes described above is generally a prototype code in a high level language such as MATLAB that is especially useful for dealing with the matrix structures and operations that arise. In order to move to a GPU implementation, the prototype models are rewritten first as serial C code, and then the time critical elements are threaded to run in parallel using the CUDA language (see Figure 4). Threads perform a kernel operation in a SIMD manner (Single Instruction Multiple Data) on the GPU device. These threads are grouped into blocks of up to 1024 threads, and then multiple blocks form the thread grid.

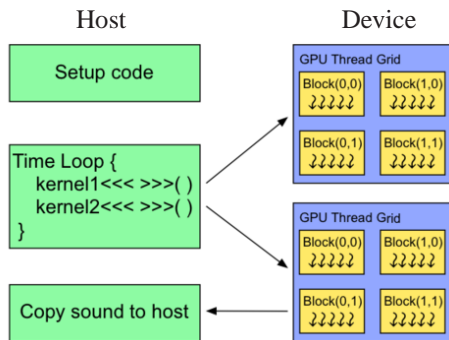


Figure 4. Code design for host and device

Following from the discussion in Section 3.2, from an implementation perspective, the systems can be grouped into four categories:

1. Uniform explicit schemes.
2. Non-uniform, but explicit, schemes.
3. Implicit schemes, with constant update matrices.
4. Implicit schemes, with update matrices constructed at each time step.

The first case, of uniform schemes, are explicit and where the update matrices **B** and **C** contain bands of constant coefficients. They are often of Toeplitz or block Toeplitz form. For such systems, the matrix form can be ‘unrolled’ such that grid points can be updated by a simple equation using scalar coefficients and neighbouring points. This greatly reduces the amount of memory access required, which is one of the main concerns for the GPU implementation. FDTD schemes generally suffer from a low compute-to-memory access ratio, and so efficiency is always limited by data transfer rather than peak computing performance. For even the simplest scheme there are still many possible approaches to thread design and various other optimisations, and therefore a certain amount of experimentation is usually required to find the most efficient CUDA solution.

The second case, where the update matrices are not uniform but the scheme is still explicit (such as, e.g., the case of schemes for Webster’s equation (4)), requires either some storage system for holding the coefficients, or operating directly with sparse matrix objects. A suitable code library is required to handle sparse matrices in C code and perform basic linear algebra operations. These operations then need to be performed, and hopefully accelerated, on the GPU. Whilst some libraries are available for this purpose, such as Nvidia’s cuSparse [29], custom-designed functions may well provide a more efficient solution.

The majority of the matrices involved are of multi-banded, or block banded form. Of the various sparse matrix formats available, the DIA (or diagonal) format is particularly suitable. Each band is set as a column in a table, with a small integer array to indicate the distance of each column away from the centre diagonal, as shown in Figure 5. This format

$$\begin{bmatrix} 1 & 2 & . & . & . \\ . & 7 & 8 & . & . \\ 11 & . & 13 & 14 & . \\ . & 17 & . & 19 & 20 \\ . & . & 23 & . & 25 \end{bmatrix} \rightarrow \begin{bmatrix} [-2 & 0 & 1] \\ . & 1 & 2 \\ . & 7 & 8 \\ 11 & 13 & 14 \\ 17 & 19 & 20 \\ 23 & 25 & . \end{bmatrix}$$

Figure 5. Sparse matrix representation in DIA format.

provides an efficient representation, and a library of linear algebra functions is being developed that are optimised specifically for the systems that arise. Some elements, such as the interpolation between 2D and 3D systems, result in matrices that are not banded in structure. In this case a more general format, such as CSR, is still required.

The third and fourth categories of code are implicit, and thus require a solution to a system of linear equations at each time step of the simulation. This may use either a persistent, unchanging matrix, or may require the construction of a matrix system at each iteration. In prototyping it is simple enough to use MATLAB’s backslash operator in this situation. Accelerating this in C code and on the GPU requires more complex libraries such as PETSc [30], or the development of custom functions based on iterative methods [31].

6. SYSTEMS AND SIMULATIONS

In this section, simulation results using FDTD methods are presented for several families of systems. Accompanying sound examples and video demonstrations appear on the NESS project website: www.ness-music.eu

6.1 Brass Instruments

As mentioned in Section 2, a starting point for brass instrument synthesis is a 1D model such as Webster's equation [4]. In the lossless case, simple, provably stable numerical methods are available [15]; for good quality synthesis, however, various refinements are necessary. Chief among these is the modeling of viscothermal boundary layer losses in the tube, which is typically described in terms of input impedance [32, 33]; when translated to the time domain, fractional time derivative terms appear, which in FDTD requires recursions of order higher than two, and thus the memory requirement/operation count is increased. A second feature of interest involves user-controlled time variation (through, e.g., slides or valves); such time variation is handled locally within an FDTD scheme, requiring no additional precomputation (in contrast with, e.g., modal methods, which would require a recalculation of modal shapes and frequencies for every valve configuration). Finally, if nonlinear effects (due to shock formation in long cylindrical bore instruments such as the trombone [34]) are to be introduced, variants of Webster's equation are no longer suitable, and it is best to revert to a first order system, as is common in the mainstream finite volume literature [9]; numerical stability is difficult to maintain under such conditions, without introducing artificial viscosity, which can impact negatively on sound quality.

6.2 Percussion: Nonlinear Plate and Shell Vibration

The high amplitude vibration of thin rigid structures such as plates and shells features strongly in many percussion instruments, including cymbals, gongs and tamtams, and also impacts to a lesser extent on the sound of drums. In general, for such instruments, under a striking excitation, there is a migration of energy from low frequencies to high, giving rise to crash like effects in cymbals, and slow swells in gongs—see Figure 6, showing a spectrogram of sound output for a typical nonlinear flat plate model. Linear models, such as that of, e.g., Kirchhoff Love [35] do not capture such effects, and thus nonlinear models are necessary—that of von Kármán [36] is probably the simplest to adequately render such effects. In the modal setting, the nonlinearity leads to the transfer of energy between modes—see Figure 7.

From an implementation perspective, the main difficulty is in performing linear system solutions in the time recursion; though sparse, due to the nonlinearity, the linear system to be solved must be constructed anew at each time step, as described in Section 3.2, and thus linear system solution techniques can become computationally quite heavy.

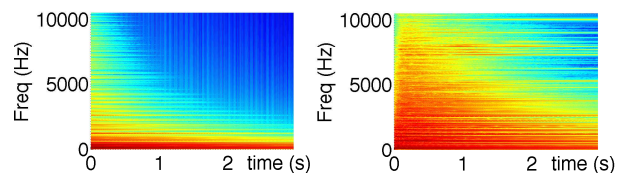


Figure 6. Spectrogram of sound output for a rectangular plate under linear conditions (left) and nonlinear conditions (right).

Linear

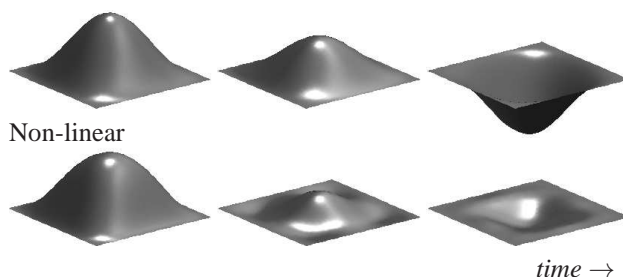


Figure 7. Linear (top) and nonlinear (bottom) time evolution of the displacement of a thin square plate, under simply supported boundary conditions. In both cases, the initial condition is set to the lowest linear mode of vibration.

6.3 Modular Environments

One of the main goals of many physical modeling sound synthesis paradigms (including the CORDIS environment [37], MOSAIC/Modalys [38], and BlockCompiler [39]) is modular construction of new virtual instruments. The goal is no different in the present case of FDTD methods, where here, the canonical elements are distributed, possibly nonlinear objects such as bars, plates, strings, membranes and acoustic tubes, all represented over distinct grids, and coupled through a variety of connection types.

In Figure 8, at left, a plate based percussion instrument is shown, with connections (of mass/spring/damper type, as in CORDIS) indicated by blue lines. In Figure 8, at right, a modular instrument constructed from a set of acoustic tubes is shown—here, as discussed in Section 6.1, the configuration is time varying, allowing for half-valve effects [40].

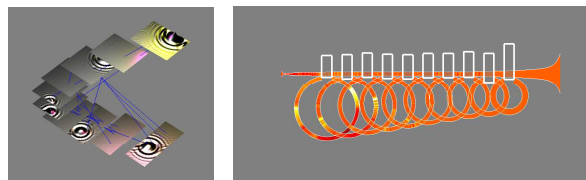


Figure 8. Left: Modular percussion instrument made up of a collection of thin plates, with connections indicated by blue lines. Right: Modular brass instrument made up of a collection of acoustic tubes.

6.4 3D Acoustic Spaces

Modeling of 3D acoustic spaces using FDTD, for room acoustics modeling and artificial reverberation applications

is a well-researched topic [41, 42, 43]. Acceleration using GPGPUs is of great importance in this case, given the computational scale of the problem.

Table 2 shows benchmark times for the standard rectilinear 3D FDTD scheme at various grid sizes; the simulation was computed for 44,100 samples at 44.1kHz. Serial C code was tested on an Intel Xeon E5-2620 with -O3 compiler optimisation, whilst the CUDA code was tested on an Nvidia Tesla K20. Both double (DP) and single (SP) precision floating-point arithmetic tests were performed.

Test	Serial C Intel	CUDA Tesla	Speedup
$1m^3$ DP	55 sec	8.7 sec	x6.3
$1m^3$ SP	55 sec	7.2 sec	x7.3
$50m^3$ DP	52.7 min	3.3 min	x16.0
$50m^3$ SP	51.6 min	2.0 min	x25.8
$500m^3$ DP	531.3 min	31.9 min	x16.6
$500m^3$ SP	528.7 min	18.9 min	x28.0

Table 2. Benchmarks for serial C code vs CUDA.

Beyond the question of raw acceleration, numerous other features in room acoustics required a detailed examination from a numerical perspective. One is air viscosity, leading to damping of wave propagation at high frequencies [44], and which is especially important to avoid unnatural ringing in computed solutions in large spaces. Another is that of stable boundary termination, especially over irregular geometries, and when realistic wall impedance conditions are taken into account. Finally, care must be taken when choosing the grid to be used in 3D space—see Figure 9 for two such choices. There are great variations in the effects of numerical dispersion depending on the grid—and also on GPU implementation.

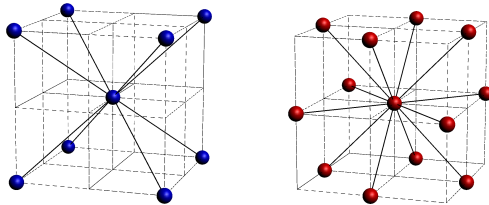


Figure 9. Computational grids in 3D: Left, body-centred cubic, and right, face-centred cubic.

6.5 Embeddings of Instruments in 3D

The ultimate goal of physical modeling is, perhaps, the full emulation of musical instruments in a 3D enclosure. New issues emerge here regarding the coupling between the acoustic field (described by a variant of the 3D wave equation 2, as detailed above) and the object at hand.

See Figure 10, illustrating the time evolution of the acoustic field surrounding a set of timpani drums, within a 3D enclosure.

7. CONCLUSIONS AND PERSPECTIVES

This paper is intended as an overview of the use of FDTD methods in sound synthesis, particularly as applied to large,

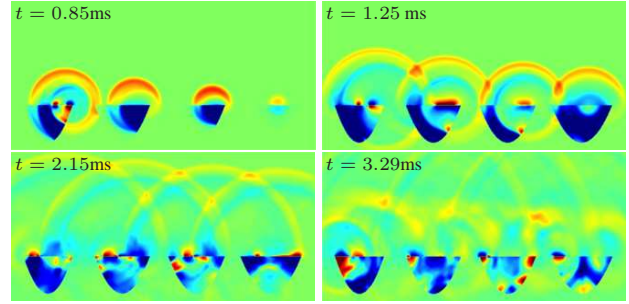


Figure 10. Snapshots of the time evolution of the acoustic field surrounding a set of four struck timpani drums.

real-world physical models, and ultimately in 3D. Such methods, which involve local or nearest neighbour updates, are a good match to implementation in parallel hardware. As one might expect, when a state space representation is employed, such methods reduce to large sparse matrix operations performed in a time loop—and in particular multiplications, and inversions (or linear system solutions). In general, it is the latter which poses the most problems in a parallel implementation, though due to matrix sparsity, specialized methods are available. For very large problems, when a linear system solution is unnecessary (as is the case for large 3D simulations), an implementation on GPU leads to great accelerations in computation time.

From an algorithm design point of view, time stepping methods represent, in some respects, a last resort—for sufficiently complex systems, involving distributed nonlinearities, and nontrivial couplings among disparate components (including the acoustic field itself), there does not appear to be any other avenue of approach. One great strength of such an approach not mentioned in the Introduction is that, ultimately, in 3D simulations, one has complete access to the acoustic field surrounding such as instrument—and thus complete control over spatialization of the resulting sound output. The difficulties, however, are many—great care must be taken at the design stage, compared with other methodologies, to ensure that numerical artifacts (such as, e.g., bandwidth limitation and numerical dispersion) do not impact negatively on sound output.

One issue which has not been broached here in any detail (and which is premature, given current compute times) is, as is usual in physical modeling synthesis, user control. Particularly in the setting of modular instrument constructions, which may be quite a bit more complex than real world acoustic instruments, finding a meaningful and parsimonious means of both designing and playing such instruments presents a daunting challenge.

Acknowledgments

This work was supported by the European Research Council, under grant StG-2011-279068-NESS.

8. REFERENCES

- [1] C. Cadoz, A. Luciani, and J.-L. Florens, “Responsive input devices and sound synthesis by simulation of instrumental mechanisms,” *Computer Music Journal*, vol. 8, no. 3, pp. 60–73, 1983.
- [2] J.-M. Adrien, “The missing link: Modal synthesis,” in *Representa-*

- tions of Musical Signals, G. DePoli, A. Piccilli, and C. Roads, Eds. Cambridge, Massachusetts: MIT Press, 1991, pp. 269–297.
- [3] J. O. Smith III, “Physical modelling using digital waveguides,” *Computer Music Journal*, vol. 16, no. 4, pp. 74–91, 1992.
- [4] P. Cook, “Tbone: An interactive waveguide brass instrument synthesis workbench for the NeXT machine,” in *Proceedings of the International Computer Music Conference*, Montreal, Canada, 1991, pp. 297–299.
- [5] M. Karjalainen, V. Välimäki, and T. Tolonen, “Plucked-string synthesis: From the Karplus-Strong algorithm to digital waveguides and beyond,” *Computer Music Journal*, vol. 22, no. 3, pp. 17–32, 1998.
- [6] V. Välimäki, M. Laurson, and C. Erkut, “Commutated waveguide synthesis of the clavichord,” *Computer Music Journal*, vol. 27, no. 1, pp. 71–82, 2003.
- [7] V. Välimäki, J. Pakarinen, C. Erkut, and M. Karjalainen, “Discrete time modeling of musical instruments,” *Reports on Progress in Physics*, vol. 69, pp. 1–78, 2006.
- [8] A. Taflov, *Computational Electrodynamics*. Boston, Massachusetts: Artech House, 1995.
- [9] R. Leveque, *Finite Volume Methods for Hyperbolic Problems*. Cambridge, UK: Cambridge University Press, 2002.
- [10] P. Ruiz, “A technique for simulating the vibrations of strings with a digital computer,” Master’s thesis, University of Illinois, 1969.
- [11] J. Kelly and C. Lochbaum, “Speech synthesis,” in *Proceedings of the Fourth International Congress on Acoustics*, Copenhagen, Denmark, 1962, pp. 1–4, paper G42.
- [12] A. Chaigne, “On the use of finite differences for musical synthesis. Application to plucked stringed instruments,” *Journal d’Acoustique*, vol. 5, no. 2, pp. 181–211, 1992.
- [13] A. Chaigne and A. Askenfelt, “Numerical simulations of struck strings. I. A physical model for a struck string using finite difference methods,” *Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 1112–1118, 1994.
- [14] L. Rhaouti, A. Chaigne, and P. Joly, “Time-domain modeling and numerical simulation of a kettledrum,” *Journal of the Acoustical Society of America*, vol. 105, no. 6, pp. 3545–3562, 1999.
- [15] S. Bilbao, *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*. Chichester, UK: John Wiley and Sons, 2009.
- [16] A. Southern, D. Murphy, G. Campos, and P. Dias, “Finite difference room acoustic modelling on a general purpose graphics processing unit,” in *Audio Engineering Society Convention 128*, 2010. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=15325>
- [17] R. Mehra, N. Raghuvanshi, L. Savioja, M. Lin, and D. Manocha, “An efficient GPU-based time domain solver for the acoustic wave equation,” *Applied Acoustics*, vol. 73, no. 2, pp. 83–94, 2012.
- [18] C. Webb and S. Bilbao, “Computing room acoustics with cuda - 3d ftdt schemes with boundary losses and viscosity,” in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 317–320.
- [19] S. Bilbao and C. Webb, “Timpani drum synthesis in 3d on gpgpus,” in *Proceedings of the 15th International Conference on Digital Audio Effects*, York, UK, September 2012.
- [20] L. Savioja, V. Valimäki, and J. Smith, “Audio signal processing using graphics processing units,” *J. Audio Eng. Soc.*, vol. 59, no. 1/2, Feb 2011.
- [21] A. Webster, “Acoustical impedance, and the theory of horns and of the phonograph,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 5, no. 7, pp. 275–282, 1919.
- [22] G. Kirchhoff, *Vorlesungen über Mechanik*. Leipzig: Tauber, 1883.
- [23] G. Carrier, “On the nonlinear vibration problem of the elastic string,” *Quarterly of Applied Mathematics*, vol. 3, pp. 157–165, 1945.
- [24] T. Tolonen, V. Välimäki, and M. Karjalainen, “Modelling of tension modulation nonlinearity in plucked strings,” *IEEE Transactions in Speech and Audio Processing*, vol. 8, pp. 300–310, 2000.
- [25] R. S. A. Quarteroni and F. Saleri, *Numerical Mathematics*. Springer, 2007.
- [26] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*. Pacific Grove, California: Wadsworth and Brooks/Cole Advanced Books and Software, 1989.
- [27] S. Bilbao, “Modeling of complex geometries and boundary conditions in finite difference/finite volume time domain room acoustics simulation,” *IEEE Transactions on Audio Speech and Language Processing*, 2013, in press.
- [28] B. Gustaffson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods*. New York, New York: John Wiley and Sons, 1995.
- [29] Nvidia, “CUSPARSE library,” *CUDA toolkit documentation*. [Online]. [Cited: 29th Mar 2013.] <http://docs.nvidia.com/cuda/>, 2013.
- [30] S. Balay, W. Gropp, L. McInnes, and B. Smith, “Efficient management of parallelism in object oriented numerical software libraries,” in *Modern Software Tools in Scientific Computing*, E. Arge, A. M. Bruaset, and H. P. Langtangen, Eds. Birkhäuser Press, 1997, pp. 163–202.
- [31] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2003.
- [32] D. Keefe, “Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximations for isothermal and nonisothermal boundary conditions,” *Journal of the Acoustical Society of America*, vol. 75, no. 1, pp. 58–62, 1984.
- [33] R. Caussé, J. Kergomard, and X. Lurton, “Input impedance of brass musical instruments—comparison between experiment and numerical models,” *Journal of the Acoustical Society of America*, vol. 75, no. 1, pp. 241–254, 1984.
- [34] A. Hirschberg, J. Gilbert, R. Msallam, and A. Wijnands, “Shock waves in trombones,” *Journal of the Acoustical Society of America*, vol. 99, no. 3, pp. 1754–1758, 1996.
- [35] K. Graff, *Wave Motion in Elastic Solids*. New York, New York: Dover, 1975.
- [36] A. Nayfeh and D. Mook, *Nonlinear Oscillations*. New York, New York: John Wiley and Sons, 1979.
- [37] C. Cadoz, A. Luciani, and J.-L. Florens, “Cordis-anima: A modeling and simulation system for sound and image synthesis,” *Computer Music Journal*, vol. 17, no. 1, pp. 19–29, 1993.
- [38] D. Morrison and J.-M. Adrien, “Mosaic: A framework for modal synthesis,” *Computer Music Journal*, vol. 17, no. 1, pp. 45–56, 1993.
- [39] M. Karjalainen, “Block-compiler: Efficient simulation of acoustic and audio systems,” presented at the 114th Audio Engineering Society Convention, Amsterdam, the Netherlands, May, 2003. Preprint 5756.
- [40] S. Bilbao, “Modeling of brass instrument valves,” in *Proceedings of the 14th International Conference on Digital Audio Effects*, Paris, France, September 2011.
- [41] L. Savioja, T. Rinne, and T. Takala, “Simulation of room acoustics with a 3-D finite-difference mesh,” in *Proceedings of the International Computer Music Conference*, Århus, Denmark, September 1994, pp. 463–466.
- [42] D. Botteldooren, “Finite-difference time-domain simulation of low-frequency room acoustic problems,” *Journal of the Acoustical Society of America*, vol. 98, no. 6, pp. 3302–3308, 1995.
- [43] D. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley, “Acoustic modelling using the digital waveguide mesh,” *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 55–66, 2007.
- [44] P. Morse and U. Ingard, *Theoretical Acoustics*. Princeton, New Jersey: Princeton University Press, 1968.

SIMULATED EFFECTS OF COMBINED CONTROL APPLIED TO AN EXPERIMENTALLY IDENTIFIED SOUNDBOARD

Simon Benacchio

IRCAM - UMR-CNRS 9912
1 Place Igor Stravinsky 75004 Paris
simon.benacchio@ircam.fr

Adrien Mamou-Mani

IRCAM - UMR-CNRS 9912
Adrien.Mamou-Mani@ircam.fr

Baptiste Chomette

Institut Jean Le Rond d'Alembert - UMR-CNRS 7190
4 Place Jussieu 75252 Paris
baptiste.chomette@upmc.fr

François Ollivier

Institut Jean Le Rond d'Alembert - UMR-CNRS 7190
francois.ollivier@upmc.fr

ABSTRACT

This paper presents an approach using combined state and derivative state active control applied to an experimentally identified model of a vibrating structure. Time simulations are made on the simplified soundboard of a string instrument and then discussed to study the effects of this control.

1. INTRODUCTION

For string instruments, mechanical characteristics of the soundboard are believed to be very important. The study of modal parameters, such as frequency and damping of their vibrating parts, garners a lot of attention. Over the last few decades, some studies such as [1] or [2] have dealt with the active control of these parameters in order to change the timbre of musical instruments. According to Richardson [3], the effective masses of soundboards are very important for string instruments since they affect the sound radiation by modifying the interaction between the strings and the soundboard. To make the radiated sound louder, active control can be used on a string instrument by changing its effective masses. However this modification must not make the system unstable. Such a control is applied using a method called combined state and derivative state control which is a modal active control method working in the state space. This optimal control method is particularly adapted to musical instruments due to the fact that the modal parameters and the modal amplitude can be modified simultaneously.

In the first section, this paper presents the proposed method for applying the combined control in the case of an arbitrary structure. Several steps are proposed to realise the design of the desired control system. A second part describes the application of this method on a soundboard model obtained from experimental data. Modeling and control system design steps are applied to this simplified structure. Then, a recorded string signal is used to perform simulations in the time domain in order to study the efficiency of

this new method of control of musical instruments. The results of these simulations are presented and discussed in the last section.

2. THE PROPOSED METHOD

2.1 Modeling of the active structure

For instruments, modal parameters are understood to be essential to the quality of the sound. According to [4], modal active control is a very convenient way to modify modal parameters of a structure and enables the control targeting of modes of vibration. On a string instrument this enables the modification of its timbre. Moreover, to increase the vibration amplitude while keeping the system stable, the combined control method, described in a following section, can be applied. Except for the controller, the control system remains similar to that described by Preumont in [5]. The dynamic of a linear system may be described by a set of first order linear differential equations

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t) + \mathbf{B}u(t) + \mathbf{G}w(t), \quad (1)$$

$$y(t) = \mathbf{C}\mathbf{X}(t), \quad (2)$$

where $\mathbf{X}(t)$ is the state vector, $u(t)$ the control signal and $y(t)$ the measured signal. \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{G} are the system, the input, the output and the disturbance signal input matrices respectively. Using the state space model, the dynamics of the Luenberger observer [6] can be written

$$\dot{\hat{\mathbf{X}}}(t) = \mathbf{A}\hat{\mathbf{X}}(t) + \mathbf{B}u(t) + \mathbf{L}(y(t) - \hat{y}(t)), \quad (3)$$

where \mathbf{L} is the optimal gain vector of the observer and with

$$\hat{y}(t) = \mathbf{C}\hat{\mathbf{X}}(t). \quad (4)$$

It is used to estimate the state vector $\hat{\mathbf{X}}$ and the derivative state vector $\dot{\hat{\mathbf{X}}}$ of the structure. According to [7], two control gain vectors \mathbf{K}_1 and \mathbf{K}_2 are then applied to the estimated state and derivative state vectors to give the control signal

$$u(t) = -\mathbf{K}_1\hat{\mathbf{X}}(t) - \mathbf{K}_2\dot{\hat{\mathbf{X}}}(t), \quad (5)$$

which is finally sent in the structure. Control scheme of this combined method is shown in Fig. 1.

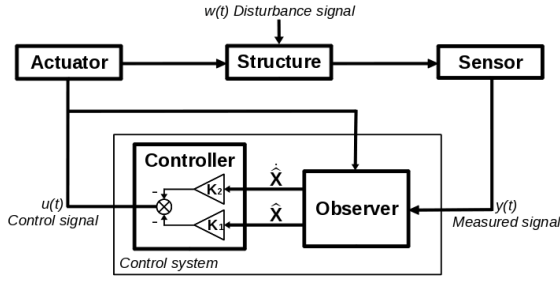


Figure 1. Combined control feedback loop.

In this study, the control system is based on one actuator and one sensor. Consequently, \mathbf{B} and \mathbf{C} are one column and one row matrices respectively. Details on state matrices are given in the section 2.2.

2.2 State model based on rational fraction polynomial identification

In the majority of cases, the actuator and the sensor modal parameters cannot be easily measured, particularly in the case of piezoelectric transducers with low electromechanical coupling coefficient. Furthermore, if the structure material is sensitive to external conditions, time can modify the modal frequencies and dampings of the structure, which are used to fill out state matrices. In this case, a good alternative is to use the Rational Fraction Polynomial (RFP) algorithm given in [8] to identify the state model (\mathbf{A} , \mathbf{B} , \mathbf{C}) used by the control system as in [9, 10]. The RFP algorithm is based on the partial fraction form which can be written for n modes

$$H(s) = \sum_{k=1}^n \left(\frac{r_k}{s - p_k} + \frac{r_k^*}{s - p_k^*} \right), \quad (6)$$

where p_k and r_k are the poles and the residues of the system respectively, $s = j\omega$ is the Laplace variable and $*$ denotes the conjugate. (6) can be expressed using the polynomial fraction form

$$H(s) = \sum_{k=1}^n \frac{A_k + B_k s}{s^2 - 2\text{Re}(p_k)s + |p_k|^2}, \quad (7)$$

where A_k and B_k can be developed as functions of poles and residues

$$A_k = -2[\text{Re}(p_k)\text{Re}(r_k) + \text{Im}(p_k)\text{Im}(r_k)], \quad (8)$$

$$B_k = 2\text{Re}(r_k). \quad (9)$$

In the case of piezoelectric transducers, the transfer function between one actuator and one sensor can be written

$$H(j\omega) = \sum_{k=1}^n \frac{\Pi_k^a \Pi_k^c}{\omega_k^2 - \omega^2 + 2j\xi_k \omega_k \omega}, \quad (10)$$

where ω_k and ξ_k are the natural frequency and the modal damping respectively, Π_k^a and Π_k^c are the actuator and sensor coefficients in modal coordinates. Assuming small and

diagonal damping, the modal parameters can be obtained using (7) and (10)

$$\omega_k = |p_k|, \quad \xi_k = -\frac{\text{Re}(p_k)}{|p_k|}, \quad (11)$$

$$\Pi_k^a \Pi_k^c = A_k + j\omega B_k \approx A_k \quad (12)$$

Assuming that the total contribution of actuator and sensor is contained in matrix \mathbf{B} and that matrix \mathbf{C} is equal to the unit matrix, the state model associated to the state vector \mathbf{X} can be written using identified modal parameters

$$\mathbf{A} = \begin{pmatrix} \mathbf{0}_{n,n} & \mathbf{Id}_{n,n} \\ -\text{diag}(\omega_k^2) & -\text{diag}(2\xi_k \omega_k) \end{pmatrix}, \quad (13)$$

$$\mathbf{B} = \begin{pmatrix} \mathbf{0}_{n,1} \\ \Pi^a \Pi^c \end{pmatrix}, \quad \mathbf{C} = (\mathbf{1}_{1,n} \quad \mathbf{0}_{1,n}), \quad \mathbf{X} = \begin{pmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{pmatrix}, \quad (14)$$

where \mathbf{q} is the modal displacement vector. This proposed state model will be used below to design the control system.

2.3 Design of the control system using state and derivative state combined control

The next step of the control system design is to find the gain vectors of the observer and of the controller. Firstly, a linear quadratic estimator (LQE) algorithm is used to find the observer gain vector \mathbf{L} . Then, a double pole placement process, described in [7], is used to find controller gain vectors. A pole placement method given in [11] enables the computing of the gain vector \mathbf{K}_2 which is then applied to $\dot{\mathbf{X}}$. Afterwards, a standard pole placement method enables the computing of the second gain vector \mathbf{K}_1 is applied to \mathbf{X} .

The designed control application is presented in section 3.

3. APPLICATION ON AN EXPERIMENTALLY IDENTIFIED SOUNDBOARD

In this section, the proposed method described in section 2 is applied on a model built with data from an experimental setup. Then, the effects of this combined control method are simulated and described.

3.1 The experimental setup

Fig. 2 presents the experimental setup. It consists of a rectangular spruce plate identical to those used by luthiers to make acoustic guitar soundboards. Its edges are under clamped boundary conditions. Its dimensions are given in Table 1.

A single string is tied on and connected to the bridge almost parallel with the soundboard plane. Its tension is chosen to set the fundamental frequency of the note at approximately 70 Hz. This frequency roughly matches the first vibration mode of the soundboard. The control of modes 1, 2, 5, 9 and 10 enables the variation of very perceptible partials of the string signal.

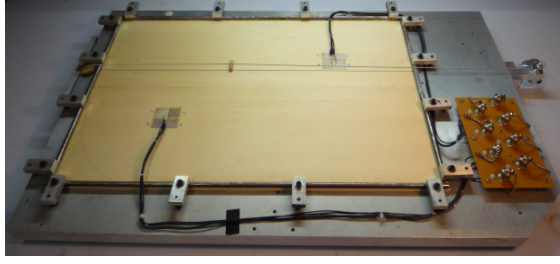


Figure 2. Experimental setup.

Plate dimensions	
Wide	4.10^{-1} m
Long	6.10^{-1} m
Thick	4.10^{-3} m
Piezo dimensions	
Wide	2.10^{-2} m
Long	2.10^{-2} m
Thick	2.10^{-4} m

Table 1. Plate and piezoelectric patches dimensions.

The transducers are piezoelectric patches made in PZT-5H. Their dimensions are given in Table 1. For each of the two positions shown in Fig. 2, four patches are glued side by side. Their positions are chosen to have high coupling coefficients for modes 1, 2, 5, 9 and 10 according to the Hać and Liu criterion given in [12]. Characteristics of wood and piezoelectric material, given in Table 2 and 3, were used in a numerical model of the soundboard to find these coupling coefficients.

Wood characteristics	
E_x [GPa]	11.5
E_y, E_z [GPa]	0.47
G_{xy}, G_{xz}, G_{yz} [GPa]	0.5
$\nu_{xy}, \nu_{xz}, \nu_{yz}$	0.005
ρ [kg.m ⁻³]	392

Table 2. Plate characteristics, with E the Young modulus, G the shear modulus, ν the Poisson coefficient and ρ the density.

Piezo characteristics	
c_{11}^E, c_{22}^E [GPa]	127.2
c_{33}^E [GPa]	117.4
c_{12}^E [GPa]	802.1
c_{13}^E [GPa]	846.7
e_{31} [C.m ⁻²]	-6.6
e_{33} [C.m ⁻²]	23.2
e_{33}^S [nC.F ⁻¹]	8.85

Table 3. Elastic, piezoelectric and dielectric constants of piezoelectric materials.

For the rest of the study, only two piezoelectric transduc-

ers are used. The first one, located in the bottom left quarter of the plate in Fig. 2, is used as the actuator in the control system. The second one, located in the top right quarter of the plate in Fig. 2, is used as the sensor.

3.2 Identification and modeling of the active structure

Prior to the control implementation, the RFP algorithm is used to identify the structure modal parameters. A frequency response function (FRF) (○) shown in Fig. 3, is used to find the state matrices. It is obtained measuring the transfer function between the actuator and the sensor piezoelectric patches. The excitation signal sent to the actuator is a chirp signal from 20 to 1600 Hz, with a 22050 Hz sampling rate.

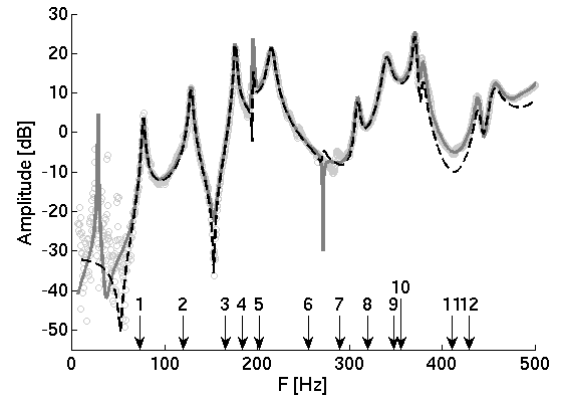


Figure 3. Experimental FRF (○), modal reconstruction (—) and state space model (- - -). Arrows indicate the modes' number.

This transfer function is used to find the estimated FRF given by (6) thanks to the RFP algorithm. 20 modes are used for the identification but only 12 are kept for the modal reconstruction. The representation of the modal model given by (7) is (—). Finally, (- - -) is the FRF of the state space model after the approximation of (12) and after having kept the 12 available identified modes whose numbers are indicated by arrows in Fig. 3. It can be noted that the frequency, the damping and the amplitude of each mode correspond closely. The first peak at approximately 30 Hz is a low frequency artefact due to noise. It is therefore not used in the state space FRF model. Moreover, the decrease of gain in the high frequencies might be due to the fact that higher modes are not used to do the identification. This first identification step allows the setting of matrices **A**, **B** and **C**. However, the matrix **G**, which matches the characteristics of the disturbance location, still remains unknown. A second identification step has to be conducted to find this last matrix on a different measured FRF. It uses the bridge admittance which is another FRF measured between the bridge and the piezoelectric sensor patch using an impact hammer.

3.3 Design of the control system

The next step for the control system design is to set the gain vectors **K**₁, **K**₂ and **L**. Transfer functions of the simulated structure for two different control configurations are

presented in Fig. 4. The first control configuration changes only the global gain value of the soundboard FRF. The first pole placement giving \mathbf{K}_2 is selected to modify the imaginary parts of the poles of modes 10 and 11 by about 40 %. The second pole placement giving \mathbf{K}_1 is useful for maintaining the same modal frequency and damping as in the uncontrolled structure. Fig. 4(a) gives the uncontrolled FRF (—) and the controlled FRF (---) for this first control configuration. The second control configuration changes not only the global gain value of the soundboard FRF but also three modal frequencies and one modal damping. The first pole placement giving \mathbf{K}_2 remains the same while the second pole placement giving \mathbf{K}_1 is chosen to apply the modifications of Table 4. Fig. 4(b) gives the uncontrolled FRF (—) and the controlled FRF (---) for this second control configuration.

Mode	Frequency shift [%]	Damping shift [%]
1	-5	0
2	+10	0
5	0	-50
9	-1	0

Table 4. Modifications applied to modes 1, 2, 5 and 9 thanks to combined control.

The arrows in Fig. 4, indicate the frequency of the partials of the string's force signal used in this study. It may be noted that several modes are absent from the soundboard FRF. For example modes 4 or 6 which were identified in section 3.2 are not visible here. This is due to the choice of actuator and sensor locations described in section 3.1.

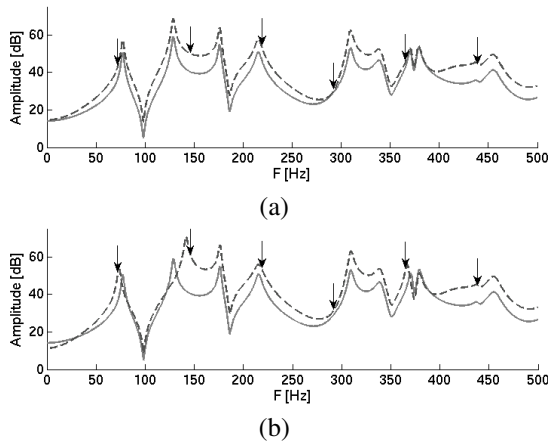


Figure 4. (a) Soundboard FRF without (—) and with (---) combined control for the first control configuration. (b) Soundboard FRF without (—) and with (---) combined control for the second control configuration. Arrows give the partials' frequency of the string's force signal.

Fig. 4 shows that the combined control increases the global gain of the soundboard FRF. However, gain increase is not equal for each mode. This is due to the fact that modal control is not an independent control and that the terms of the \mathbf{B} and \mathbf{C} matrices depend on the features and

location of the actuator and sensor. For the second control configuration, modifications are selected to modify the perceptible partials of the string signal. 4. The frequency shifts of the first, second and ninth modes should increase the first, second and fifth partial amplitudes respectively. The decrease of the fifth mode damping should not have a significant effect on the amplitude of the third partial since both FRF gains seem to be similar at this frequency for the two control configurations.

All matrices and vectors used to design the control system, are now known. In the next part, they are used to simulate the different control configurations on the soundboard model.

3.4 Combined control simulation based on experimental data

The effects of the combined control applied on a soundboard model are studied in this part. A modeling software is used to build the numerical model thanks to the \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{G} matrices found in section 3.2. Then, the matrices \mathbf{A} , \mathbf{B} and \mathbf{C} and the gain vectors \mathbf{K}_1 , \mathbf{K}_2 and \mathbf{L} are used to build the numerical control system. It may be noted that these last numerical block could be used in an experimental case without any changes.

The disturbance signal used in the simulation is a recording of a bowed string's force signal measured thanks to a piezoelectric sensor set under the string on a cello bridge. It is given in Fig. 5(a).

Then, a time simulation is conducted sending this measured time signal in the controlled soundboard model. This signal, filtered by the soundboard model, is given in Fig. 5(b). It matches the simulated vibration of the soundboard, measured by the piezoelectric sensor and without control of the structure.

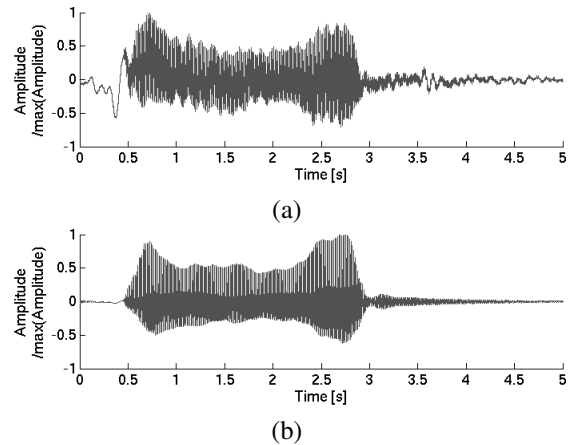


Figure 5. String's force signal used as disturbance on the numerical soundboard model applied to the bridge location (a) and simulation given by piezoelectric sensor (b).

It is then possible to measure this signal for each control configuration given in Fig. 4. In these cases, the shape of the signal curves is very similar to the shape of the curve of Fig. 5(b) but with a larger amplitude. This means that these two control configurations make the sound louder while remaining stable. Fig. 6 shows the Fourier transform of

the signal as a sum of harmonic peaks whose amplitude is filtered by the uncontrolled soundboard FRF and its shape (—).

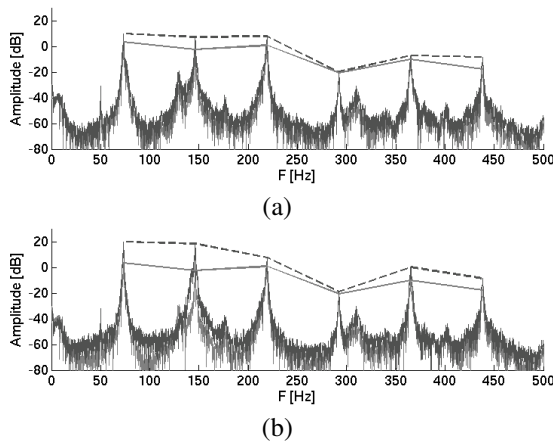


Figure 6. (a) Fourier transform shape of the measured signal without (—) and with (---) combined control for the first configuration.

(b) Fourier transform shape of the measured signal without (—) and with (---) combined control for the second configuration.

Fig. 6(a) gives the Fourier transform shape of the resulting sound (---) for the first control configuration. This control increases the amplitude of the global soundboard FRF. The amplitude increase of each peak is given in Table 5.

Partial	Amplitude increase [dB]
1	6.5
2	9.7
3	6.9
4	0.8
5	2.7
6	9.2

Table 5. Partials amplitude increase for the first control configuration compared to uncontrolled configuration.

As mentioned in section 3.3, the amplitude increase is not equal for each peak, particularly for the fourth partial. This is not very problematic since it is not the most perceptible partial.

Fig. 6(b), gives the Fourier transform shape of the resulting sound (---) for the second control configuration. The amplitude increase in each peak and the amplitude difference between the two control configurations are given in Table 6.

The effects of the frequency shifts are very clear. Indeed, partials 1, 2 and 5 increase approximately ten decibels. The effects of damping control on the fifth mode are less clear. Even if the effect on the FRF in Fig. 4 is very visible, it is too localised to impact upon the sound partial. However, a modification of the frequency with the damping control should enable the decrease of this partial. It can be noted

Partial	Amplitude increase [dB]	Difference [dB]
1	16	9.5
2	20.8	11.1
3	6.5	-0.4
4	2	1.2
5	10.2	7.5
6	8.9	-0.3

Table 6. Partials amplitude increase for the second control configuration compared to uncontrolled configuration and amplitude difference between the first and the second control configuration.

that the frequency and damping control have an effect on the entire soundboard FRF since the amplitude of the partials 4 and 6 have also changed. So this second control configuration not only enables a change in the amplitude of the measured signal but also the modification of its timbre thanks to the changes in the frequency and damping of the modes.

4. DISCUSSION

This combined control method has to be tested on an experimental structure. The first comment that could be made pertains to the control signal. The amplitude of this signal must not be too high in order to keep a reasonable tension level at the amplifier output. Thanks to the time simulations, it is possible to observe the control signal amplitude. For the first control configuration, this amplitude is near to 300 V for its highest value. This is an acceptable value since the tension amplifier used in this control system is able to deliver a maximum value of 350 V. However, for the second control configuration, this amplitude is near to 500 V. It is too high for the amplifier. The control signal will be saturated and the control system will not reach the control target. Modifications on the soundboard FRF have to be smaller. Which leads to the second comment. It is difficult to accurately set small amplitude modifications of the soundboard FRF since the influence of the modes on each other seems to be important. This could be due to the fact that a single input - single output (SISO) system has been used in this study. Indeed, a one sensor and one actuator system is not able to control a lot of modes at the same time. Understanding of this phenomenon, modeling of higher modes and the use of a multi input - multi output (MIMO) system could be a promising area of investigation.

5. CONCLUSION

This paper proposes an approach using combined control method, enabling the change of the modal frequencies, dampings and amplitudes of a structure. The advantage of this method is to ensure the stability of the system. An approach is proposed to design a control system using experimental data. The effects of this new control method are studied thanks to time simulations on a simplified soundboard model. The result of a first simulation shows that

the combined control method enables the change in the modal amplitudes of the soundboard FRF. A second simulation shows that, in addition to this first modification, the modal frequencies and dampings can be shifted independently. Applied to musical instruments, this combined control method allows the modification of their timbres as well as making their sound louder.

Acknowledgments

This work has been done during the PhD of Simon Benacchio, funded by the Agence National de la Recherche (ANR IMAREV project).

6. REFERENCES

- [1] S. Griffin, “Acoustic replication in smart structures using active structural/acoustic control”, *Phd thesis*, Georgia Institute of Technology, 1995.
- [2] H. Boutin, “Méthodes de contrôle actif d’instruments de musiques acoustiques”, *Phd thesis*, Université Pierre et Marie Curie, 2011.
- [3] B. Richardson, “Simple models as basis for guitar design”, *Journal Catgut Acoustical Society*, vol. 4, pp. 30–36, 2002.
- [4] B. Chomette, D. Remond, S. Chesné, and L. Gaudiller, “Semi-adaptive modal control of on-board electronic boards using identification method”, *Smart Material and Structures*, vol. 17, pp. 1–8, 2008.
- [5] A. Preumont, *Vibration Control of Active Structures: An Introduction*. Springer Ed., 2001.
- [6] D. Luenberger, “An Introduction to Observers” *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, 1971.
- [7] S. Benacchio, A. Mamou-Mani, B. Chomette, and F. Ollivier, “Modal active control applied to simplified string musical instrument” in *Proc. Int. Conf. of Acoustics*, Montreal, 2013.
- [8] M.H. Richardson, D.L. Formenti, “Parameter estimation from frequency response measurements using rational fraction polynomials” in *Proc. IMAC Conference*, Orlando, FL, 1982.
- [9] V. Lhuillier, “Contrôle actif de la transparence acoustique d’une double paroi”, *Phd thesis*, Institut National des Sciences Appliquées de Lyon, 2009.
- [10] S. Chesné, C. Jean-Mistral, and L. Gaudiller, “Modal active control of composite structure using identification techniques”, in *Vibrations, Shocks and Noise*, Clamart, 2012.
- [11] T. Abdelaziz, “A direct algorithm for pole placement by state-derivative feedback for single-input linear systems” *Acta Polytechnica*, vol. 43, 2003.
- [12] A. Hać, and L. Liu, “sensor and actuator location in motion control of flexible structures” *Journal of Sound and Vibration*, vol. 167(2), pp. 239–261, 1993.

SOUND SYNTHESIS OF GONGS OBTAINED FROM NONLINEAR THIN PLATES VIBRATIONS: COMPARISON BETWEEN A MODAL APPROACH AND A FINITE DIFFERENCE SCHEME.

Michele Ducceschi
UME - ENSTA Paristech
ducceschi@ensta.fr

Cyril Touzé
UME - ENSTA Paristech

Stefan Bilbao
University Of Edinburgh

ABSTRACT

The sound of a gong is simulated through the vibrations of thin elastic plates. The dynamical equations are necessarily nonlinear, crashing and shimmering being typical nonlinear effects. In this work two methods are used to simulate the nonlinear plates: a finite difference scheme and a modal approach. The striking force is approximated to the first order by a raised cosine of varying amplitude and contact duration acting on one point of the surface. It will be seen that for linear and moderately nonlinear vibrations the modal approach is particularly appealing as it allows the implementation of a rich damping mechanism by introducing a damping coefficient for each mode. In this way, the frequency-dependent decay rates can be tuned to get a very realistic sound. However, in many cases cymbal vibrations are found in strongly nonlinear regimes, where an energy cascade through lengthscales brings energy up to high-frequency modes. Hence, the number of modes retained in the truncation becomes a crucial parameter of the simulation. In this sense the finite difference scheme is usually better suited for reproducing crash and gong-like sounds, because this scheme retains all the modes up to (almost) Nyquist. However, the modal equations will be shown to have useful symmetry properties that can be used to speed up the off-line calculation process, leading to large memory and time savings and thus giving the possibility to simulate higher frequency ranges using modes.

1. INTRODUCTION

Thin plates are common mechanical elements found in several contexts in physics and engineering, from fluid-structure interaction, to aeronautics, civil engineering, wave turbulence [1–3], and others. The context of musical acoustics does not represent an exception, including linear and nonlinear examples. Plates are constitutive parts of several musical instruments: the soundboards of pianos and guitars are thin plates vibrating usually in a linear regime. Plates have been used in the past as reverberation units in music performances before the advent of digital software. Large

metallic plates were found at times in theatres as they can reproduce quite conveniently the sound of a storm when shaken.

When a thin plate is struck with a mallet or a hammer at large amplitudes, it produces a gong-like sound [4, 5]. Nonlinear effects caused by the large amplitudes of vibration are responsible for the crashing and shimmering sound similar to that of a gong [6]. These effects are reproduced by the von Kármán equations, who have proved to be an effective model for weakly nonlinear vibrations despite the introduction of a single second-order correction in the in-plane strain tensor with respect to the linear plate equations [7, 8].

Here the focus is on the resolution of the von Kármán equations in the context of sound synthesis obtained through physical modelling. A finite difference code developed by Bilbao [9, 10] is used as a benchmark for testing a code based on modal projection. The finite difference scheme is energy conserving and was used before in the analysis of thin plates in chaotic and turbulent regimes [1]. Although modal schemes have been employed several times in physical modelling of linear instruments, examples of their use for the production of sounds in a nonlinear context are rare. This is explained by the fact that the nonlinear modal equations can be quite involved, and therefore require a lot of memory and computational time, considering that the number of modes to be kept for synthesis is typically a few hundred.

In this work, some symmetries of the modal nonlinear plate equations are shown to help achieve memory and computational savings: in turn, this can help in the creation of faster synthesis algorithms. The most appealing feature of modal synthesis is the possibility of adding a rich damping mechanism with practically no extra effort: it is sufficient to add a damping term to each mode. The damping terms can be tuned at will, and they can be estimated in a real experiment, at least in a first approximation. In a finite difference code, on the contrary, damping can only be introduced in the time domain, thus limiting the implementation of a rich mechanism.

2. MODEL EQUATIONS

Consider a rectangular domain S of lengths L_x, L_y with boundary ∂S . Cartesian coordinates $\mathbf{x} \equiv (x, y)$ will be used to identify a point over S . Weakly nonlinear vibrations $w(\mathbf{x}, t)$ of the order of the plate thickness h are de-

scribed by the von Kármán equations. These are:

$$\rho h \ddot{w} = -D \Delta \Delta w - 2\sigma_0 \dot{w} + L(w, F) + P(\mathbf{x}, t); \quad (1a)$$

$$\Delta \Delta F = -\frac{Eh}{2} L(w, w). \quad (1b)$$

The function $F(\mathbf{x}, t)$ is an auxiliary function that describes the in-plane motion; it is usually referred to as Airy's stress function. The symbol Δ is the Laplacian; therefore in Cartesian coordinates $\Delta \Delta w \equiv (w_{,xx} + w_{,yy})^2$. $D = \frac{Eh^3}{12(1-\nu^2)}$ is the rigidity of the plate, where E is Young's modulus and ν is Poisson's ratio; ρ is the volume density. $P(\mathbf{x}, t)$ is a forcing term of some kind acting perpendicularly to the plate surface, and σ_0 is a damping coefficient. $L(\cdot, \cdot)$ is the nonlinear coupling term known as von Kármán operator, who reads:

$$L(w, F) = w_{,xx} F_{,yy} + w_{,yy} F_{,xx} - 2w_{,xy} F_{,xy}. \quad (2)$$

The system must be provided with boundary conditions along ∂S . In this work, simply supported boundary conditions with movable in-plane edges are chosen, *i.e.*

$$w = w_{,nn} = 0 \quad \forall \mathbf{x} \in \partial S, \quad (3a)$$

$$F = F_{,n} = 0 \quad \forall \mathbf{x} \in \partial S, \quad (3b)$$

where n is the direction normal to the boundary. These conditions describe edges fixed in the transverse direction, but free to rotate. In the in-plane direction the plate is free of loads.

2.1 Linear Modes and Modal Projection

A solution to system (1) can be obtained by projecting the functions w and F onto their linear modes. These are defined as:

$$w = S_w \sum_{i=1}^{N_w} \frac{\Phi_i(\mathbf{x})}{\|\Phi_i\|} q_i(t); \quad (4a)$$

$$\Delta \Delta \Phi_i(\mathbf{x}) = \frac{\rho h}{D} \omega_i^2 \Phi_i(\mathbf{x}). \quad (4b)$$

$$F = S_F \sum_{i=1}^{N_F} \frac{\Psi_i(\mathbf{x})}{\|\Psi_i\|} \eta_i(t); \quad (4c)$$

$$\Delta \Delta \Psi_i(\mathbf{x}) = \zeta_i^4 \Psi_i(\mathbf{x}). \quad (4d)$$

(4b) and (4d) are completed, respectively, by their boundary conditions (3a) and (3b). Note that N_w and N_F are, in theory, infinite. However they must be truncated to finite numbers for obvious computational reasons. As usual for linear problems, the modes are orthogonal with respect to a suitable scalar product. The scalar product between two modes Φ_i, Φ_j can be defined as

$$\langle \Phi_i, \Phi_j \rangle_S = \int_S d\mathbf{x} \Phi_i \Phi_j, \quad (5)$$

and the orthogonality condition imposes

$$\langle \Phi_i, \Phi_j \rangle_S = \|\Phi_i\|^2 \delta_{ij}, \quad (6)$$

with δ being Kronecker's delta. The norm of each mode is thus imposed by the scalar product. However, the constant S_w appearing in (4a) can be chosen so that the norm of the mode $S_w \Phi_i / \|\Phi_i\|$ becomes, precisely, S_w . The same is true for the modes Ψ_i and the constant S_F .

To obtain the modal equations, (4) is inserted into (1). Then one takes inner products of (1a) and (1b) with, respectively, $\Phi_s(\mathbf{x})$ and $\Psi_k(\mathbf{x})$ to get:

$$\ddot{q}_s + 2\chi_s \omega_s \dot{q}_s + \omega_s^2 q_s = -\frac{ES_w^2}{2\rho} \sum_{n,p,q,r=1}^{\infty} \frac{H_{q,r}^n E_{p,n}^s}{\zeta_n^4} q_p q_q q_r + \frac{\langle \Phi_s, P(\mathbf{x}, t) \rangle_S}{\|\Phi_s\| \rho h S_w}. \quad (7)$$

Note that the original σ_0 coefficient is replaced here by suitable χ_s coefficients, which are the modal damping coefficients. Regarding the forcing, one usually chooses a pointwise impulsion at the point \mathbf{x}_0 , therefore

$$P(\mathbf{x}, t) = \delta(\mathbf{x} - \mathbf{x}_0) p(t). \quad (8)$$

The form of $p(t)$ can be chosen as a raised cosine of the form

$$p(t) = \begin{cases} \frac{p_0}{2} (1 + \cos(\pi(t - t_0)/\Delta t)), & |t - t_0| \leq \Delta t \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

This creates a raised cosine of maximum amplitude p_0 centered around t_0 and of length Δt . This is a first approximation to a striking impulsion on the plate. Typically, for a timpani mallet one may choose $p_0 \approx 5 - 35$ N and $\Delta t \approx 5$ ms; for drum sticks $p_0 \approx 17 - 200$ N and $\Delta t \approx 0.15 - 0.3$ ms. Examples are given in fig. 1(a).

Note that, with the current choice for $P(\mathbf{x}, t)$ one gets

$$\langle \Phi_s, P(\mathbf{x}, t) \rangle_S = \Phi_s(\mathbf{x}_0) p(t). \quad (10)$$

Moreover, two third order tensors appear in (7). These are:

$$H_{q,r}^n = \frac{\langle \Psi_n, L(\Phi_q, \Phi_r) \rangle_S}{\|\Psi_n\| \|\Phi_q\| \|\Phi_r\|}; \quad (11a)$$

$$E_{p,n}^s = \frac{\langle \Phi_s, L(\Phi_p, \Psi_n) \rangle_S}{\|\Phi_p\| \|\Phi_s\| \|\Psi_n\|}. \quad (11b)$$

The two tensors can be combined to give

$$\Gamma_{p,r,q}^s \equiv \sum_{n=1}^{N_F} \frac{H_{q,r}^n E_{p,n}^s}{\zeta_n^4}; \quad (12)$$

this tensor is the tensor of coupling coefficients for the modal equations. It is fourth order because the modal equations are cubic with respect to the modal coordinates $q_s(t)$.

2.1.1 Solutions to The Eigenvalue Problems

A solution to (4b) with boundary conditions (3a) is obtained immediately by considering

$$\Phi_i(\mathbf{x}) = \sin\left(\frac{i_1 \pi x}{L_x}\right) \sin\left(\frac{i_2 \pi y}{L_y}\right), \quad (13)$$

for integers i_1, i_2 . This gives the following eigenfrequencies of vibration:

$$\omega_i^2 = \frac{D}{\rho h} \left[\left(\frac{i_1 \pi}{L_x} \right)^2 + \left(\frac{i_2 \pi}{L_y} \right)^2 \right]^2. \quad (14)$$

On the other hand, for the Airy stress function modes there is no analytical solution. It is worth noticing that the eigenvalue problem (4d) with boundary conditions (3b) corresponds mathematically to the problem of a clamped Kirchhoff plate, even though it describes the physical situation of in-plane motion free of loads at the boundary. The question is then how to find the modes of a Kirchhoff plate with clamped edges. A possible strategy is to construct an appropriate algebraic eigenvalue problem starting from energy considerations. This is known as Galerkin's method. This method is based on the assumption that the generic eigenfunction Ψ_k can be written as a weighted sum of chosen expansion functions, hence:

$$\Psi_k(\mathbf{x}) = \sum_{n=1}^{N_c} a_k^n \Lambda_n(\mathbf{x}). \quad (15)$$

The rate of convergence and accuracy of such a method relies heavily on the expansion functions used to approximate the sought solution, as well as on the total number of functions, N_c . Obviously the expansion functions must form a complete set over the domain of interest; in addition they need to satisfy all the geometric boundary conditions. The case of the clamped plate presents two such conditions, namely zero displacement and zero slope at the boundary. Usually one resorts to modification of Fourier series, for which completeness follows directly from the Fourier theorem. In addition, satisfaction of the boundary conditions can be achieved by adding a fourth order polynomial to the Fourier series, as explained in [11]. Hence for the clamped plate problem one may use

$$\Lambda_n(\mathbf{x}) = X_{n_1}(x)Y_{n_2}(y), \quad (16)$$

where

$$X_{n_1}(x) = \cos\left(\frac{n_1\pi x}{L_x}\right) + \frac{15(1 + (-1)^{n_1})}{L_x^4}x^4 - \frac{4(8 + 7(-1)^{n_1})}{L_x^3}x^3 + \frac{6(3 + 2(-1)^{n_1})}{L_x^2}x^2 - 1, \quad (17)$$

and similarly for $Y_{n_2}(y)$. The eigenvalue problem may be written in the form

$$K\mathbf{a} = \zeta^4 M\mathbf{a}; \quad (18)$$

this gives the expansion coefficients a_k^n along with the eigenvalues ζ_k . The stiffness and mass matrices are obtained from energy considerations [11], and they read:

$$K_{ij} = \langle \Delta \Lambda_i, \Delta \Lambda_j \rangle_S - \int_S d\mathbf{x} L(\Lambda_i, \Lambda_j), \quad (19a)$$

$$M_{ij} = \langle \Lambda_i, \Lambda_j \rangle_S. \quad (19b)$$

2.2 The Finite Difference Approximation

Time and space are discretised so that the continuous variables (x, y, t) are approximated by their discrete counterparts (lh_x, mh_y, nh_t) , where (l, m, n) are integer indices and (h_x, h_y, h_t) are the steps. Boundedness of the domain implies that $(l, m) \in [0, N_x] \times [0, N_y]$ so that the grid size

is given by $(N_x + 1) \times (N_y + 1)$. The continuous variables $w(\mathbf{x}, t)$, $F(\mathbf{x}, t)$ are then approximated by $w_{l,m}^n$, $F_{l,m}^n$ at the discrete time n for the grid point (l, m) . Time shifting operators are introduced as

$$e_{t+}w_{l,m}^n = w_{l,m}^{n+1}, \quad e_{t-}w_{l,m}^n = w_{l,m}^{n-1}. \quad (20)$$

Time derivatives can then be approximated by:

$$\delta_{t-} = \frac{1}{2h_t}(e_{t+} - e_{t-}), \quad \delta_{t+} = \frac{1}{h_t}(e_{t+} - 1),$$

$$\delta_{t-} = \frac{1}{h_t}(1 - e_{t+}), \quad \delta_{tt} = \delta_{t+}\delta_{t-}. \quad (21)$$

Time averaging operators are introduced as:

$$\mu_{t+} = \frac{1}{2}(e_{t+} + 1), \quad \mu_{t-} = \frac{1}{2}(1 + e_{t-}),$$

$$\mu_t = \frac{1}{2}(e_{t+} + e_{t-}), \quad \mu_{tt} = \mu_{t+}\mu_{t-}. \quad (22)$$

Similar definitions hold for the space operators. Hence, the Laplacian Δ and the double Laplacian $\Delta\Delta$ are given by:

$$\delta_{\Delta} = \delta_{xx} + \delta_{yy}, \quad \delta_{\Delta\Delta} = \delta_{\Delta}\delta_{\Delta}. \quad (23)$$

The von Kármán operator at interior points $L(w, F)$ can then be discretised as:

$$l(w, F) = \delta_{xx}w\delta_{yy}F + \delta_{yy}w\delta_{xx}F - 2\mu_x\mu_y(\delta_{x+y}w\delta_{x+y}F). \quad (24)$$

Thus the discrete counterpart of (1) is:

$$D\delta_{\Delta\Delta}w + \rho h\delta_{tt}w = l(w, \mu_t F) - 2\sigma_0\delta_{t-}w + P_{l,m}^n; \quad (25a)$$

$$\mu_{t-}D\delta_{\Delta\Delta}F = -\frac{Eh}{2}l(w, e_{t-}w). \quad (25b)$$

Such a scheme is energy conserving, where the discrete energy is positive definite and yields a stability condition, as proved in [9, 10]. Implementation of boundary conditions is explained thoroughly in [10]. For the sound synthesis of a struck plate, however, the constraint of energy conservation may be relaxed: if the initial amplitude of vibration is not too large (typically less than $10\sqrt{L_x L_y}$), the damping effects will make sure that the time series will not become unstable. A faster scheme than can then be implemented, and this is:

$$D\delta_{\Delta\Delta}w + \rho h\delta_{tt}w = l(w, F) - 2\sigma_0\delta_{t-}w + P_{l,m}^n; \quad (26a)$$

$$D\delta_{\Delta\Delta}F = -\frac{Eh}{2}l(w, w). \quad (26b)$$

2.3 Convergence of Γ coefficients

Table 1 presents a convergence test of the Γ tensor for a plate of aspect ratio $\xi = L_x/L_y = 2/3$. The convergence in this case depends on two factors: the first is the amount of stress function modes retained in the definition of Γ (N_F in eq. (12)); the second is the accuracy on the Airy stress function modes and frequencies (quantified by the number N_c in eq. (15)). For clarity, in the following tables

k	N_F		
	100	144	225
1	20.033	20.034	20.034
20	$7.5605 \cdot 10^3$	$9.4893 \cdot 10^3$	$9.4960 \cdot 10^3$
50	$1.3928 \cdot 10^4$	$1.3929 \cdot 10^4$	$1.3937 \cdot 10^4$
100	$1.4847 \cdot 10^4$	$2.7360 \cdot 10^4$	$1.2413 \cdot 10^5$
k	400	484	625
	20.034	20.034	20.034
20	$9.4970 \cdot 10^3$	$9.4975 \cdot 10^3$	$9.4977 \cdot 10^3$
50	$1.3937 \cdot 10^4$	$1.3937 \cdot 10^4$	$1.3937 \cdot 10^4$
100	$1.3334 \cdot 10^5$	$2.2100 \cdot 10^5$	$2.2108 \cdot 10^5$

Table 1. Convergence of coupling coefficients, $\Gamma_{k,k,k}^k (L_x L_y)^3$, $\xi = 2/3$.

k	Grid Points		
	51×76	161×241	280×419
1	20.728	20.381	20.189
20	$9.7935 \cdot 10^3$	$9.6413 \cdot 10^3$	$9.5567 \cdot 10^3$
50	$1.4440 \cdot 10^4$	$1.4234 \cdot 10^4$	$1.4080 \cdot 10^4$
100	$2.0223 \cdot 10^5$	$2.0246 \cdot 10^5$	$2.0286 \cdot 10^5$

Table 2. Convergence of coupling coefficients, FD scheme, $\Gamma_{k,k,k}^k (L_x L_y)^3$, $\xi = 2/3$, $N_F = 225$

N_F is always the same as N_C . It is seen that a four-digit convergence up to the $\Gamma_{100,100,100}^{100}$ coefficient is obtained when $N_F = 484$. The same coefficients can be calculated using the finite difference code. Results are summarised in table 2. It is seen that the coupling coefficients can be calculated to very high precision when using the modal description. The major drawback of the modal approach is the limited number of modes that one can keep for the simulations (≈ 200). On the other hand, the finite difference scheme produces simulations including a vast number of modes (≈ 10000), at the expense of numerical precision. Comparing tables 1 and 2 suggests that a four-digit convergence is out of reach for the finite difference scheme with typical grid sizes. The question is then which one of the two methods is better suited for sound synthesis.

3. EXAMPLE: A STRUCK PLATE

In this section an example is investigated to compare the two methods. The reference plate is a steel plate of dimensions $L_x \times L_y = 0.4 \times 0.6$ m², Young's modulus $E = 2 \cdot 10^{11}$ Pa, density $\rho = 7860$ kg/m³, Poisson's ratio $\nu = 0.3$ and thickness $h = 1$ mm. The plate is excited with a temporal raised cosine at one point, with $\Delta t = 0.1$ ms and varying amplitude p_0 . The damping factor is $\omega_s \chi_s = 0.75$ s⁻¹, resulting in $\sigma_0 = 0.75 \rho h$ kg/m².

The finite difference scheme is run at 100 kHz, resulting in a grid size of 51×76 points. Scheme (25) takes about two hours of calculation per second of simulation, on a machine running MATLAB equipped with Intel Core i5 CPU 650 @ 3.20GHz. However, the simpler scheme (26)

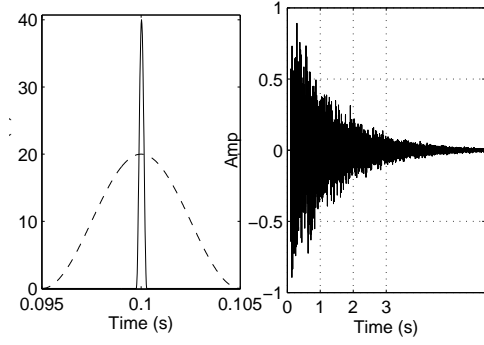


Figure 1. (a) Raised cosine simulating the impacts of a timpani mallet (dashed line) and a wooden drumstick (solid line). (b) Normalised time series obtained from the modal scheme when the plate is excited with a raised cosine of amplitude 15 N and $\Delta t = 0.1$ ms, and having a damping factor $\sigma_0 = 0.75 \rho h$ kg/m².

is faster and requires about an hour of calculation for the same simulation parameters.

On the other hand, for the modal approach the biggest issue results in the calculation of the Γ tensor. However, this calculation is performed off-line once and for all. It is worth noticing that the same tensor can be used for all problems sharing the same boundary conditions and aspect ratio. The typical maximum size achievable for a fourth order tensor of dimension N_w^4 is reached when N_w (the total number of displacement modes) is about 150; in this case the tensor occupies about 1 GB of memory. From section 2.3 it is seen that a reasonable convergence of the Γ tensor is obtained when $N_F = 484$. The calculation of Γ with $N_F = 484$, $N_w = 150$ takes about 0.5 hours in MATLAB. Once the tensor is ready, one can design a basic Störmer-Verlet scheme for the time integration [12]. This is achieved by replacing $\frac{d^2}{dt^2}$, $\frac{d}{dt}$ with their discrete counterparts δ_{tt} , δ_t in (7). For sound synthesis, this scheme can be run at lower sampling rates, typically 20kHz; doing so results in a calculation time of about 10 minutes per second of simulation. The time series produced by the modal scheme for $p_0 = 15$ N is shown in fig.1(b). Although the envelope resembles that of a decaying impulsion, the high frequency content of the time series is really poor: this is where the truncation to $N_w = 150$ modes rears its head. Fig.2 compares the spectrograms obtained from finite differences and modes, in the case of three forcing cases: $p_0 = 15, 100, 200$ N (top to bottom). As the excitation grows, energy is passed to higher parts of the spectrum, which bear significant perceptual information. The flow of energy towards smaller scales should be interpreted in the realm of wave turbulence [2]; in actual fact, wave turbulence in vibrating plates is the subject of extensive studies [1,3,8]. A turbulent flow is characterised by the flow of energy to higher frequencies up to the dissipation scales; however the modal truncation does not allow the energy to flow past 5000 Hz, a limit which makes the plate sound dull

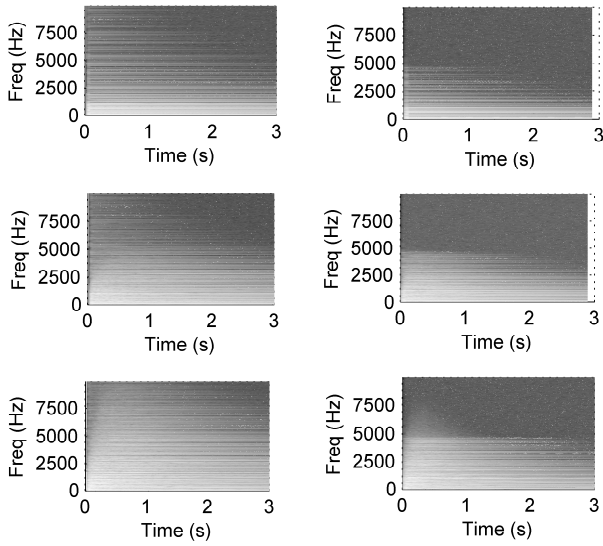


Figure 2. Comparison between FD (left) and modes (right) for $p_0 = 15$ N (top), 100 N (middle), 200 N (bottom). The modal truncation at $N_w = 150$ modes is evident in all cases.

and colourless. The finite difference scheme, on the contrary, produces a much brighter and sharp sound, because the frequency spectrum is much larger (for a sampling rate of 100kHz, the upper limit is about 40kHz). The question is then how to spare memory space in the calculation of the Γ tensor, and possibly how to speed up its calculation. This is the subject of the next section.

4. CALCULATION SHORTCUTS

4.1 Symmetry Properties

Useful symmetries can be derived for the Γ tensor. A first obvious property is the following:

$$H_{p,q}^i = H_{q,p}^i. \quad (27)$$

This follows directly from the definition of $L(\cdot, \cdot)$, which is bilinear in its entries. A second, less straightforward property is obtained when integrating by parts the tensor E from (11b). This gives

$$E_{p,q}^n = H_{p,n}^q + \oint \mathbf{n} \cdot \delta \mathbf{S} G_{p,q}^n \quad (28)$$

where G is a third order tensor depending on Ψ_n , Φ_p , Φ_q and their derivatives along the boundary. It was noted in [7] that the selected boundary conditions make the integral vanish. In this way, the tensor Γ may then be conveniently written as

$$\Gamma_{p,q,r}^s = \sum_{n=1}^{N_F} \frac{H_{p,q}^n H_{r,s}^n}{\zeta_n^4}. \quad (29)$$

Basically the symmetry properties for Γ mean the following sets of indices will produce the same numerical value:

$$(s, p, q, r), (r, p, q, s), (s, q, p, r), (r, q, p, s),$$

$$(q, r, s, p), (p, r, s, q), (q, s, r, p), (p, s, r, q).$$

These symmetry properties can lead to large memory savings when the number of transverse and in-plane modes is a few hundred.

4.2 Null Coupling Coefficients

For the sake of numerical computation, it would be interesting to know *a priori* which coupling coefficients are null. In actual fact, empirical observations of the Γ tensor suggest that only a smaller fraction of coefficients is not zero. As an example, consider fig.3. In this figure some of the coefficients $\Gamma_{p,q,r}^n$ are plotted for given n, p and for $q = r = 1 : 10$. A black slot corresponds to a value of zero. It is clear that more than half of the coefficients is vanishing. Thus, an *a priori* knowledge of their occurrences could lead to huge memory and computational savings. Referring to fig.3, one may notice that the patterns depicted in the diagrams of the left column are repeated exactly by the adjacent diagrams in the right column. This suggests that the modes can be grouped in families whose members share the same coupling rules with respect to members of another family. A straightforward way to divide the modes on the plate is to consider the symmetry of their shape with respect to a coordinate system with origin at the centre of the plate. Four families exist, and these are: doubly symmetric (SS), antisymmetric-symmetric (AS and SA) and doubly antisymmetric (AA). For instance, the first mode is a doubly-symmetric mode because it presents one maximum at the centre of the plate, and is thus symmetric with respect to the two orthogonal directions departing from the centre of the plate in the x and y directions. The first few modes for the case under study may be classified in the following groups:

- SS: 1,4,8,11,12,20,...
- SA: 2,7,9,14,16,...
- AS: 3,6,13,15,19,...
- AA: 5,10,17,18,...

Observing the black slots of fig.3 permits to state the following heuristic rule:

the indices (s, p, q, r) will give a nonzero value for $\Gamma_{p,q,r}^s$ if and only if modes s, p, q, r come all from distinct modal shape groups or if they come from the same group two by two.

For example, the combinations (SS, SS, AS, SA) and (SS, SS, SS, AS) will definitely give a zero value; on the other hand the combinations (SS, SS, SS, SS), (SS, AA, SS, AA) and (SS, AS, SA, AA) will give a nonzero value. A rigorous mathematical proof is not carried out as it involves a rather lengthy development which is beyond the scope of the present work. For the sake of generality, it should be mentioned that similar symmetries hold for other boundary conditions in rectangular and circular geometries [7, 13]. Thanks to this interpretation it is easy to see why the left column diagrams are reproduced symmetrically in the right

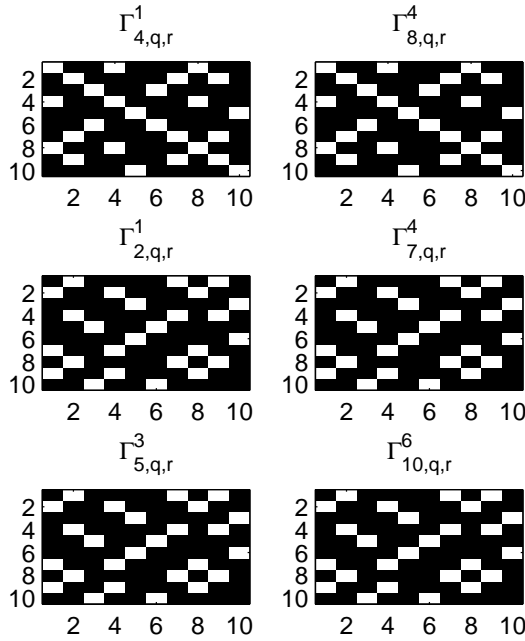


Figure 3. The null coupling coefficients (black squares) are placed in the slots predicted by the heuristic rule.

column: the modes chosen for the two columns come from the same families, hence they give rise to the same zero coefficients. This rule, in combination with the previous remarks on symmetry, can be used to speed up the calculation of the Γ tensor (for example by pre-allocating the zero entries when using a sparse matrix description). Ideally, a tensor comprising 400 - 500 modes should be used for sound synthesis, in order to be able to simulate frequencies up to 10kHz.

5. CONCLUSIONS

This work presented a comparison between a finite difference and a modal scheme for the reproduction of the nonlinear vibrations of a thin, metallic plate, aiming at simulating the sound emitted by a gong struck by an impulsive force (like a mallet). The two methods were shown to yield the same results, within the bounds imposed by numerical approximation. The problems induced by modal truncation were highlighted and their influence on the sound produced described. It was seen that symmetry and coupling rules are desirable to the extent of reducing the burden deriving from the calculation of the coupling coefficient tensor. These properties were shown, leading to an a priori knowledge of the null coupling coefficients. These coefficients constitute more than a half of the total coefficients. This work opens the possibility to construct a very effective and precise modal scheme for nonlinear vibrations: the coupling coefficients can be calculated to very high precision, and a fine damping mechanism can be implemented in the modal equations. Sound samples from the finite difference and modal schemes will be played at the conference, for

both rectangular and circular geometries.

6. REFERENCES

- [1] C. Touzé, S. Bilbao, and O. Cadot, "Transition scenario to turbulence in thin vibrating plates," *Journal of Sound and Vibration*, vol. 331, no. 2, pp. 412–433, 2011.
- [2] G. Düring, C. Josserand, and S. Rica, "Weak turbulence for a vibrating plate: Can one hear a kolmogorov spectrum?" *Phys. Rev. Lett.*, vol. 97, p. 025503, 2006.
- [3] A. Boudaoud, O. Cadot, B. Odille, and C. Touzé, "Observation of wave turbulence in vibrating plates," *Phys. Rev. Lett.*, vol. 100, p. 234504, 2008.
- [4] C. Touzé, S. Bilbao, L. Longo-Mucciante, O. Cadot, and A. Boudaoud, "Vibrations chaotiques de plaques minces: application aux instruments de type cymbale," in *Proceedings of CFA 2010, 10eme Congrès Français d'Acoustique*, Lyon, April 2010.
- [5] K. A. Legge and N. H. Fletcher, "Nonlinearity, chaos, and the sound of shallow gongs," *The Journal of the Acoustical Society of America*, vol. 86, no. 6, pp. 2439–2443, 1989.
- [6] A. Chaigne, C. Touzé, and O. Thomas, "Nonlinear vibrations and chaos in gongs and cymbals," *Acoustical Science and Technology*, vol. 26, pp. 403–409, 2005.
- [7] O. Thomas and S. Bilbao, "Geometrically nonlinear flexural vibrations of plates: In-plane boundary conditions and some symmetry properties," *Journal of Sound and Vibration*, vol. 315, no. 3, pp. 569–590, 2008.
- [8] C. Touzé, O. Thomas, and M. Amabili, "Transition to chaotic vibrations for harmonically forced perfect and imperfect circular plates," *International Journal of Non-Linear Mechanics*, vol. 46, no. 1, pp. 234 – 246, 2011.
- [9] S. Bilbao, "A family of conservative finite difference schemes for the dynamical von Kármán plate equations," *Numerical Methods for Partial Differential equations*, vol. 24, no. 1, pp. 193–216, 2008.
- [10] —, *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*. Wiley, 2009.
- [11] W. Li, "Vibration analysis of rectangular plates with general elastic support," *Journal of Sound and Vibration*, vol. 273, no. 3, pp. 619–635, 2003.
- [12] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*. Springer, 2006.
- [13] O. Thomas, C. Touzé, and A. Chaigne, "Non-linear vibrations of free-edge thin spherical shells: modal interaction rules and 1:1:2 internal resonance," *International Journal of Solids and Structures*, vol. 42, no. 1112, pp. 3339 – 3373, 2005.

A STRUCTURED APPROACH TO USING A RECTANGULAR BRACE TO DESIGN A SOUNDBOARD SECTION FOR A DESIRED NATURAL FREQUENCY

Patrick Dumond

pduomo057@uottawa.ca

Department of Mechanical Engineering

University of Ottawa

161 Louis Pasteur, CBY A205

Ottawa, Canada K1N 6N5

Natalie Baddour

nbaddour@uottawa.ca

ABSTRACT

The manufacture of acoustically consistent wooden musical instruments remains economically demanding and can lead to a great deal of material waste. To address this, the problem of design-for-frequency of braced plates is considered in this paper. The theory of inverse eigenvalue problems seeks to address the problem by creating representative system matrices directly from the desired natural frequencies of the system. The goal of this paper is to demonstrate how the generalized Cayley-Hamilton theorem can be used to find the system matrices. In particular, a simple rectangular brace-plate system is analyzed. The radial stiffness of the plate is varied in order to model variations typically found in wood which is quartersawn. The corresponding thickness of the brace required to keep the fundamental natural frequency of the brace-plate system at a desired value is then calculated with the proposed method. It is shown that the method works well for such a system and demonstrates the potential of using this technique for more complex systems, including soundboards of wooden musical instruments.

1. INTRODUCTION

Many aspects of the manufacture of wooden musical instruments have been addressed and rendered consistent. However, acoustical consistency still remains unattainable in most situations [1]. This is a consequence of the fact that the material of choice for many musical instruments is wood, a natural material that exhibits high variability in its material properties. By definition, wood has inconsistent properties because its growth is directly related to the highly variable climate of its environment. Luthiers have been compensating for these inconsistent material properties in guitars by means of various methods for years. The most prominent method currently in use is to adjust the shape of the soundboard's braces in order to attain a more consistent frequency spectrum from this part of the instrument [2]. These methods have had varying degrees of success and mostly depend on the skill and experience of the luthier. Worsening the situation is

the fact that these methods are very labor intensive, rendering them cost-prohibitive, as well as difficult to implement into a structured manufacturing process. For these reasons, most manufacturers only use material that has mechanical properties which fit within their set criteria. Such an approach leads to much waste [3].

Like most design problems, a design is first created from experience and then iteratively refined in order to achieve the desired parameters. This is especially true of systems in which certain eigenvalues are desired [4]. For guitar soundboards, luthiers begin with a certain design, remove material from the braces in small increments and then check the system's natural frequencies (eigenvalues) until a desired solution converges. It has been shown in previous work that it is indeed possible to alter certain frequencies of a soundboard system by simply adjusting the shape of the braces [5]. While effective, this trial-and-error method is not optimal. A better approach would be to design/construct the system directly from the desired natural frequencies (eigenvalues).

In order to achieve this, we turn to the field of study known as inverse eigenvalue problems, which deals specifically with finding matrices from a set of given eigenvalues [6], [7]. A rather young area, inverse eigenvalue problems use knowledge of matrix algebra and numerical methods to create matrices that yield a desired frequency spectrum (set of eigenvalues) or a partial spectrum. It is well known that inverse eigenvalue problems are ill-posed, meaning there generally exists many solutions [7].

In design, the existence of many solutions is potentially beneficial, giving the designer options. However, physical constraints do need to be applied in order to make a system physically realizable. Most methods for inverse eigenvalue problems involve the use of well-developed matrix theory for matrices with a specific structure (e.g. Jacobi and band matrices) and then apply appropriate numerical algorithms to solve for the unknown matrices from the known desired eigenvalues [8], [9], [10], [11], [12], [13], [14]. The structure of the matrices generally implies various physical constraints. However, there exist very few methods that can solve for matrices having a more general unstructured form. The goal of this paper is to demonstrate the use of a technique that has been recently proposed using the generalized Cayley-Hamilton

theorem [15]. In particular, this method is interesting since it allows the use of *any* matrix structure. Therefore, a suitable matrix structure can be determined by other means, for instance from modeling the forward dynamics of the problem. Implementing particular desired constraints thus becomes an exercise of the forward modeling process.

In this paper, we apply the Cayley-Hamilton technique to a simple rectangular brace-plate system in order to design the combined brace-plate system to a desired natural frequency. In doing so, we demonstrate that a braced plate can be designed directly from knowledge of the desired fundamental frequency. This approach is novel because it would allow the construction of wooden soundboards having a consistent set of natural frequencies via design of the braces.

2. MODEL

2.1 Problem Statement

Given a desired fundamental frequency, construct a brace-plate system as described by a mass matrix M and a stiffness matrix K .

All mechanical properties of the system are a function of the radial stiffness E_R of the wooden specimen, which is assumed known and given (and which tends to vary from specimen to specimen). All dimensional (geometric) properties of the brace-plate system are assumed to be specified and fixed except for the thickness of the brace h_c , the design variable for which we must solve.

2.2 Forward Model

The model is based on a typical section of a guitar soundboard, where a single brace is used to structurally reinforce the weaker plate direction. The model is shown in Figure 1.

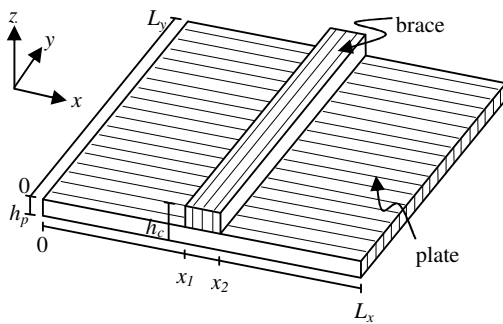


Figure 1. Orthotropic plate reinforced with a rectangular brace.

The forward model is discretized using the assumed shape method, similar to the model used in [5]. The assumed shape method is an energy method which uses global plate elements within the kinetic and strain energy plate equations in order to determine the system's equations of motion, from which the mass and stiffness matrices are extracted [16]. The system is assumed simply

supported, conservative and the material properties are assumed orthotropic.

The forward model is created assuming Sitka spruce and all material properties are related to the radial stiffness of the wood specimen (i.e. the Young's Modulus in the radial direction, E_R) as indicated in [17]. The use of E_R is chosen because quality control practice observed in industry use the stiffness across the grain of the wooden soundboards, measured as in Figure 2.

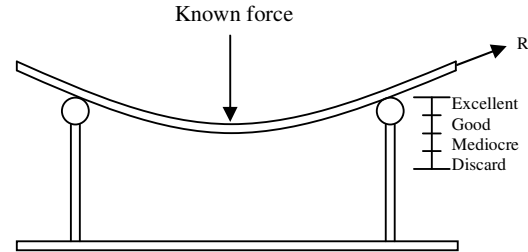


Figure 2. Quality control measurement of soundboard plate used in industry

It is assumed that an exact measurement of E_R could be obtained in a similar fashion as that described above and which could be used in the calculations.

2.3 Inverse Model

The goal of this paper is to reconstruct the brace-plate system from a desired fundamental frequency. The generalized Cayley-Hamilton theorem inverse eigenvalue method is used [15]. The generalized Cayley-Hamilton theorem states that if $p(\lambda)$ is the characteristic polynomial of the generalized eigenvalue problem (K, M) , where K and M are square matrices obtained from $p(\lambda) = \det(K - \lambda M)$, then substituting $(M^{-1}K)$ for λ in the polynomial gives the zero matrix [18], [19]. Thus, by building the model in the forward sense, and by leaving relevant design parameters as variable, it is possible to design the brace-plate model for the fundamental frequency.

It is shown in [20] that in order to adjust the fundamental frequency of the brace-plate system to a desired value, it is necessary to adjust the thickness of the brace. A cross section of the fundamental modeshape is shown in Figure 3. It is clear that the brace affects the maximum amplitude of this modeshape, thus also affecting the associated frequency [21].

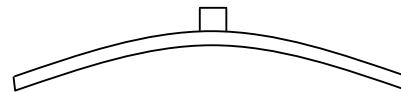


Figure 3. Cross section of the brace-plate system's fundamental modeshape

2.4 Cayley-Hamilton Algorithm

Using the facts that the wood's mechanical properties vary based on its radial stiffness, and that the brace thickness controls the brace-plate system's fundamental frequency, the forward model is created using the assumed

shape method while leaving these two parameters as variables. Thus, the mass matrix M is a function of h_c , the height of the brace-plate system at the (assumed fixed) location of the brace, and the stiffness matrix K is a function of h_c and also E_R , the plate's radial stiffness. Here, we use 2×2 trial functions in the assumed shape method. Hence, 4th order square matrices are created. The trial functions used are those of the simply supported rectangular plate such that

$$w(x, y, t) = \sum_{n_x=1}^{m_x} \sum_{n_y=2}^{m_y} \sin\left(\frac{n_x \pi x}{L_x}\right) \sin\left(\frac{n_y \pi y}{L_y}\right) q_{n_x n_y}(t) \quad (1)$$

where m are the modal numbers, q the time function and w is the displacement variable normal to the plate. The displacement variable w is then used directly in creating the kinetic and strain energy equations of the simply supported rectangular plate. These equations are broken into three sections as shown in Figure 1 in order to take into account the brace. This procedure is well described in [5].

It is assumed that the E_R of the wood selected for manufacture is measured during the manufacturing process and used as input information into the stiffness matrix. This leaves h_c as the only unknown parameter, appearing in both the mass and stiffness matrices.

In order to solve these matrices from the desired fundamental frequency, the Cayley-Hamilton theorem is used. To do so, the characteristic polynomial is created using the desired frequency,

$$p(\lambda) = (\lambda - a) \cdot (\lambda - b_1) \cdot (\lambda - b_2) \cdot (\lambda - b_3) \quad (2)$$

where a is the desired frequency and b_1 - b_3 are unknown values which need to be found. Since we have assumed 2×2 trial functions so that the mass and stiffness matrices are both 4×4 , the characteristic polynomial must be fourth order, as shown in equation (2). Subsequently, $p(\lambda)$ is expanded so that the polynomials coefficients can be found. Once the polynomial is created, the Cayley-Hamilton equation can be written by substituting $(M^{-1}K)$ for λ into equation (2).

$$p(K, M) = c_4 (M^{-1}K)^4 + c_3 (M^{-1}K)^3 + c_2 (M^{-1}K)^2 + c_1 (M^{-1}K) + c_0 I = 0 \quad (3)$$

where c_n are the coefficients of λ in $p(\lambda)$ determined via equation (2). As stated in [15], equation (3) produces sixteen equations, of which only four are independent. Solving the equations on the main diagonal for the four unknowns (h_b , b_1 , b_2 , b_3) produces $4^4 = 256$ possible solutions, according to Bézout's theorem [22]. From the set of all possible solutions, complex solutions can be immediately eliminated as not being physically meaningful. Clearly, further constraints must be added to the solution in order to get a solution which fits within the desired physical limits. These physical limits are based on the maximum and minimum brace dimensions which are required to compensate for the range of plate stiffnesses

used during the analysis, as well as the range of natural frequencies which can be obtained using these system dimensions. Thus, the following constraints are implemented into the solution:

$$\begin{aligned} 0.013 &\leq h_b \leq 0.016 \text{ m} \\ 1 \times 10^7 &\leq b_1 \leq 9 \times 10^8 \text{ rad/s} \\ 1 \times 10^7 &\leq b_2 \leq 9 \times 10^8 \text{ rad/s} \\ 1 \times 10^7 &\leq b_3 \leq 9 \times 10^8 \text{ rad/s} \end{aligned} \quad (4)$$

Solving the four equations obtained from equation (3) within the constraints provided by (4) yields a physically realistic solution which satisfies the desired fundamental frequency, as well as the system's parameters.

3. RESULTS

3.1 Material Properties

The material used for the brace-plate system during the analysis is assumed to be Sitka spruce, due to its widespread use in the industry. The mechanical properties of Sitka spruce are obtained from [17] and are given in Table 1.

Material properties	Values
Density – μ (kg/m ³)	403.2
Young's modulus – E_R (MPa)	850
Young's modulus – E_L (MPa)	$E_R / 0.078$
Shear modulus – G_{LR} (MPa)	$E_L \times 0.064$
Poisson's ratio – ν_{LR}	0.372
Poisson's ratio – ν_{RL}	$\nu_{LR} \times E_R / E_L$

Table 1. Material properties of Sitka spruce [17].

Since wood is an orthotropic material, the 'R' and 'L' subscripts used in Table 1 refer to the radial and longitudinal property directions of wood, respectively. The wood used in making instrument soundboards is generally quartersawn. Therefore the tangential property direction can be neglected. Since wood properties are highly variable, the values presented in Table 1 represent a statistical average and are hence used as a benchmark for further analysis.

3.2 Model Dimensions

The dimensions used for the model throughout the analysis of the brace-plate system are shown in Table 2.

Dimensions	Values
Length – L_x (m)	0.24
Length – L_y (m)	0.18
Length – L_b (m)	0.012
Reference – x_1 (m)	$L_x / 2 - L_b / 2$
Reference – x_2 (m)	$x_1 + L_b$
Thickness – h_p (m)	0.003
Thickness – h_b (m)	0.012
Thickness – h_c (m)	$h_p + h_b$

Table 2. Dimensions of brace-plate model.

These dimensions refer to those shown in Figure 1, where ‘ p ’ refers to the plate’s dimensions, ‘ b ’ refers to the brace’s dimensions and ‘ c ’ refers to the dimensions of the combined system. These dimensions are chosen based on the dimensions of a typical guitar soundboard section reinforced by a single brace.

3.3 Benchmark Values

In order to demonstrate the working values of the model, a benchmark is set using the statistical averages for the properties of Sitka spruce as defined in Table 1. Therefore, a plate with a radial stiffness of $E_R = 850$ MPa to which a brace is attached with a combined brace-plate thickness of $h_c = 0.015$ m is investigated. Using these values and the forward model, the eigenvalue problem can be solved to find a fundamental natural frequency of 687 Hz for the brace-plate system. In order to see the effect on the overall system, the four frequencies provided using 2×2 trial functions are shown in Table 3. In this case m_x and m_y represent the mode numbers in the x and y directions respectively.


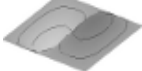
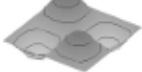

m_x	m_y	Natural frequency (Hz)	Modeshapes
1	1	687	
2	1	790	
2	2	1366	
1	2	2648	

Table 3. System’s natural frequencies at benchmark values.

3.4 Analysis

After determining the benchmark values, an analysis is performed using the inverse method described in the previous section. As the plate’s radial stiffness varies, the thickness of the brace-plate section is calculated such that the fundamental frequency of the brace-plate system is kept consistent at 687 Hz. The results of the computations can be found in Table 4.

Young’s modulus E_R (MPa)	Brace thickness h_c (m)	Fundamental Frequency a (Hz)
750	0.01576	687
800	0.01536	687
813	0.01527	687
850	0.01500	687
900	0.01466	687
950	0.01435	687

Table 4. Results of the inverse model analysis.

Clearly, adjusting the thickness of the brace also has an effect on the other natural frequencies. These can be seen in Table 5.

Young’s modulus E_R (MPa)	Brace thickness h_c (m)	b_1 (Hz)	b_2 (Hz)	b_3 (Hz)
750	0.01576	774	1360	2653
800	0.01536	782	1363	2650
813	0.01527	784	1364	2650
850	0.01500	790	1366	2648
900	0.01466	798	1370	2645
950	0.01435	806	1374	2642

Table 5. Calculated frequencies of the inverse model analysis.

Interestingly, the constraints indicated in equation (4), although physically strict, allow for more than one solution in certain cases. An example is shown in Table 6.

Young’s modulus E_R (MPa)	Brace thickness h_c (m)	a (Hz)	b_1 (Hz)	b_2 (Hz)	b_3 (Hz)
750	0.01359	687	570	1149	2185

Table 6. Alternate brace thickness solution satisfying the physical constraints.

These results, along with their importance are discussed below.

4. DISCUSSION

From these results, it is evident that designing a brace-plate system starting with a desired fundamental frequency is possible. Table 4 clearly shows that by adjusting the thickness of the brace by small increments (10^{-5} m, machine limit), it is possible to compensate for the variation in the radial stiffness of the plate (E_R) so that the fundamental frequency of the combined system is equal to that of the benchmark value of 687 Hz. The results obtained using the Cayley-Hamilton theorem algorithm match those values obtained using the forward model exactly when compared to the benchmark values in Table 3.

In modifying the thickness of the brace, the fundamental frequency is not the only frequency which is modified. Table 5 shows that frequencies b_1 to b_3 are affected, with a varying degree of magnitude. To be precise, while fundamental frequencies a remain consistent, frequencies b_1 to b_3 vary by 2%, 0.6% and 0.2% respectively based on the variation of E_R . Therefore, it is important to ensure that there is a good understanding of what the brace can control. A detailed discussion of varying the shape of the brace in order to simultaneously control two natural frequencies is found in [5]. This involves the use of scalloped shaped braces such as the one shown in Figure 4. Using such a brace, one can control both the fundamental frequency, as well as one of its higher partials simultaneously. Furthermore, by increasing the number of variable parameters such as brace position or number of braces, many more system frequencies could be controlled.

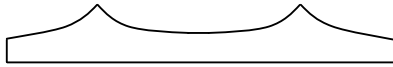


Figure 4. Scalloped shaped brace

Moreover, it was a surprise to find that within the strict physical constraints of (4), there is more than one brace-plate system (solution) that satisfies the Cayley-Hamilton theorem of equation (3). From Table 6 it can be seen that an alternate solution to the system exists, different from the one presented in Table 4, for a plate having a radial stiffness of 750 MPa. In this case, by reducing the thickness of the brace, it is still possible to achieve a system having the desired frequency of 687Hz. However, the desired frequency is no longer the fundamental frequency but rather becomes the second frequency and the fundamental has been replaced with a fundamental frequency of 570 Hz. It is important to keep this phenomenon in mind while designing a system. This is especially true if the order in the spectrum of a certain frequency associated with a certain modeshape is absolutely critical.

In situations requiring the specification of a fundamental frequency, it is clear that to achieve consistency in this desired frequency, the thickness of the brace must increase when the stiffness of the plate decreases and vice-versa. The method demonstrated in this paper provides for a precise way in which these values can be found.

5. CONCLUSION

In this paper, we demonstrate a direct method for adjusting the natural frequencies of various systems including musical instruments. The method also directly demonstrates the ability to compensate for variations in the radial stiffness of wooden plates, thereby making it possible to create brace-plate systems having a consistent fundamental frequency. This structured approach represents a significant improvement over heuristic methods currently in use. Although only a simple model was presented in this paper, the concept of design-for-frequency using the Cayley-Hamilton method was demonstrated as a proof of concept for future work in the field of musical acoustics. It is clear that much work needs to be done in order to apply this method of design to actual musical instrument soundboards. However, this technique holds great potential for creating system matrices of complex systems from a set of desired frequencies. In doing so, it promises to greatly benefit the advancement of the manufacturing of acoustically consistent wooden musical instrument soundboards.

Acknowledgments

The authors would like to acknowledge the generous support provided by the Natural Sciences and Engineering Research Council of Canada.

6. REFERENCES

- [1] R. M. French, "Engineering the Guitar: Theory and Practice," in *Engineering the Guitar: Theory and Practice*, 1st ed., New York: Springer, 2008, pp. 159–208.
- [2] R. H. Siminoff, *The Luthier's Handbook: A Guide to Building Great Tone in Acoustic Stringed Instruments*. Milwaukee, WI: Hal Leonard, 2002.
- [3] M. French, R. Handy, and M. J. Jackson, "Manufacturing sustainability and life cycle management in the production of acoustic guitars," *International Journal of Computational Materials Science and Surface Engineering*, vol. 2, no. 1, pp. 41–53, Jan. 2009.
- [4] A. Schoofs, F. V. Asperen, P. Maas, and A. Lehr, "I. Computation of Bell Profiles Using Structural Optimization," *Music Perception: An Interdisciplinary Journal*, vol. 4, no. 3, pp. 245–254, Apr. 1987.
- [5] P. Dumond and N. Baddour, "Effects of using scalloped shape braces on the natural frequencies of a brace-soundboard system," *Applied Acoustics*, vol. 73, no. 11, pp. 1168–1173, Nov. 2012.
- [6] M. T. Chu and G. H. Golub, *Inverse Eigenvalue Problems: Theory, Algorithms, and Applications*. Oxford University Press, USA, 2005.
- [7] G. M. L. Gladwell, *Inverse problems in vibration*. Kluwer Academic Publishers, 2004.
- [8] D. Boley and G. H. Golub, "A survey of matrix inverse eigenvalue problems," *Inverse Problems*, vol. 3, no. 4, p. 595, 1987.
- [9] M. T. Chu and J. L. Watterson, "On a Multivariate Eigenvalue Problem: I. Algebraic Theory and a Power Method," *SIAM J. Sci. Comput.*, vol. 14, pp. 1089–1106, 1993.
- [10] R. Erra and B. Philippe, "On some structured inverse eigenvalue problems," *Numerical Algorithms*, vol. 15, no. 1, pp. 15–35, 1997.
- [11] F. W. Biegler-König, "Construction of band matrices from spectral data," *Linear Algebra and its Applications*, vol. 40, pp. 79–87, Oct. 1981.
- [12] G. H. Golub and R. R. Underwood, "The block Lanczos method for computing eigenvalues," in *Mathematical Software III*, J.R. Rice., New York: Springer, 1977.
- [13] F. W. Biegler-König, "A Newton iteration process for inverse eigenvalue problems," *Numerische Mathematik*, vol. 37, no. 3, pp. 349–354, 1981.
- [14] M. T. Chu, "A Fast Recursive Algorithm for Constructing Matrices with Prescribed Eigenvalues and Singular Values," *SIAM Journal on Numerical Analysis*, vol. 37, no. 3, pp. 1004–1020, Jan. 2000.

- [15] P. Dumond and N. Baddour, "A Structured Approach to Design-for-Frequency Problems Using the Cayley-Hamilton Theorem," *Computers & Structures*, vol. Submitted: CAS-D-13-00301, Apr. 2013.
- [16] L. Meirovitch, "Principles and Techniques of Vibrations," in *Principles and Techniques of Vibrations*, Upper Saddle River, NJ: Prentice Hall, 1996, pp. 542–543.
- [17] Forest Products Laboratory (US), "Wood Handbook, Wood as an Engineering Material," Madison, WI: U.S. Department of Agriculture, Forest Service, 1999, pp. 4.1–13.
- [18] A. W. Knapp, "Basic Algebra," in *Basic Algebra*, Boston: Birkhäuser, p. 219.
- [19] F. R. Chang and H. C. Chen, "The generalized Cayley-Hamilton theorem for standard pencils," *Systems & Control Letters*, vol. 18, no. 3, pp. 179–182, Mar. 1992.
- [20] P. Dumond and N. Baddour, "Toward improving the manufactured consistency of wooden musical instruments through frequency matching," in *Transactions of the North American Manufacturing Research Institution of SME*, 2010, vol. 38, pp. 245–252.
- [21] P. Dumond and N. Baddour, "Effects of a Scalloped and Rectangular Brace on the Modeshapes of a Brace-Plate System," *International Journal of Mechanical Engineering and Mechatronics*, vol. 1, no. 1, pp. 1–8, 2012.
- [22] J. L. Coolidge, *Treatise on Algebraic Plane Curves*. Mineola, New York: Dover Publications, 1959.

A new method for the identification of the original modes of damped three-dimensional axi-symmetric structures subjected to constraining boundary conditions

Vincent Debut, Miguel Carvalho and Jose Antunes

Applied Dynamics Laboratory

Campus Tecnológico e Nuclear, Instituto Superior Técnico/Universidade Técnica de Lisboa

2686-953 Sacavem, Portugal

mailto:vincentdebut@ctn.ist.utl.pt

ABSTRACT

The motivation of this work comes from the need to diagnose the tuning of large historical carillon bells which, for security reasons, had to be provisionally supported using scaffolds at several locations of their rim. Since the presence of these additional supporting fixtures significantly changes the modes of the free structure, a direct estimation of the original bell vibrational properties from measurements is not feasible. However, as recently presented by the present authors, the dynamical problem can be formulated in terms of structural modification techniques, from which the original modal frequencies of the bells, as well as the local mass and stiffness constraints, can be recovered. In this paper, we extend to a more realistic situation our identification strategy developed for a 1D conceptual system by considering a continuous 3D axisymmetric shell and accounting for dissipation phenomena. We start by briefly presenting the relevant dynamical formulations and then illustrate the technique on a simulated realistic axisymmetric shell. Our identification results highlight the robustness of the proposed technique. Interestingly for a practical situation, the method is not prone to disturbing effects from modal identification nor truncation errors since it operates directly on the constrained transfer function measured at the constrained locations.

1. INTRODUCTION

Structural modification is usually referred to as a technique to analyze the effects of additional local physical elements on the vibrational properties of a structure. For a practical configuration, the problem can be addressed by using a direct or an inverse formulation. The direct problem aims to predict the dynamical property changes once the physical parameters of the system have been modified, whereas the inverse problem aims to determine the necessary structural changes that would result in a

desired dynamic behaviour. Although less frequent, the *reverse* problem can also be of interest: it consists of recovering the unconstrained modes of a structure from the dynamical information measured (or computed) under constrained configuration.

The motivation for this work stems from an interest in knowing the tuning of large historical carillon bells, which had to be provisionally supported using scaffolds, constraining the bell rim at several locations. We are aware that the stringent precision demanded on instrument tuning is, in practice, possibly beyond the estimation capabilities of structural modification techniques. Indeed, just a few cents of mistuning are already clearly perceived by any trained ear [1]. However, in that form, the difficulties created by the additional anchoring of the bells appeared as an opportunity to formulate the dynamical problem in terms of structural modification techniques which might enable us to produce (more or less crude) estimations of the original bell modes, before installation of the constraining supports.

Following such an objective, the present authors recently described a method for recovering the free modes of axisymmetric systems from frequency response data (FRFs) stemming from a constrained configuration [2]. Instead of solving an eigenvalue problem, the proposed approach, developed along the lines of Li and He [3], involves a set of linear equations with readily available solutions. The method, applied for a conceptual system of discrete mass-stiffness elements arranged in a ring shape, seems promising and extension to three-dimensional structures has been presented [4] very recently. One central feature is to impose some specific properties of axisymmetric bodies during the inversion, which also allows to identify both mass and stiffness local constraints.

In this paper, we turn to an even more realistic situation, by accounting for internal dissipation in the structure as well as for the additional dissipation imposed by the supports. Obviously, the introduction of local damping by the constraints leads to complex modes for the constrained system, and consequently to additional difficulties as far as the dynamics of the constrained system is concerned. One should notice however that such complexity will be encapsulated in the measured (or computed) transfer functions, and one can assume light (or proportional) damping for the unconstrained system,

with real modeshapes and modal frequencies close to those of the undamped system. Compared to the procedure presented in [2, 4], another significant improvement is that, beyond these modelling aspects, the identification technique no longer requires the knowledge of the transfer functions of the constrained system under near-static conditions ($\omega \rightarrow 0$): both mass and stiffness constraints are now identified simultaneously by simple inversion. Due to the poor quality of the data collected at such a low frequency in a practical situation, this is a strong feature of the new proposed technique.

We start by briefly presenting the relevant formulations of structural modifications and then describe the identification technique for recovering the unconstrained modes as well as the constraint parameters from constrained measurements. The inversion procedure for the case of three-dimensional axi-symmetric structure is then detailed and finally, numerical examples are provided to demonstrate the feasibility of the method developed for the given system.

2. BASIC MODELLING APPROACH FOR STRUCTURAL MODIFICATIONS

The dynamics of the original unconstrained damped structure are generally modeled mathematically in terms of a mass matrix $[\mathbf{M}_u]$, a damping matrix $[\mathbf{C}_u]$ and a stiffness matrix $[\mathbf{K}_u]$ which can be written as:

$$[\mathbf{M}_u]\{\ddot{\mathbf{Y}}(t)\} + [\mathbf{C}_u]\{\dot{\mathbf{Y}}(t)\} + [\mathbf{K}_u]\{\mathbf{Y}(t)\} = \{\mathbf{F}(t)\} \quad (1)$$

If local mass, damping and stiffness modifications are applied by means of additional devices at discrete locations, then the dynamical equations of the constrained system can be written as:

$$[\mathbf{M}_c]\{\ddot{\mathbf{Y}}(t)\} + [\mathbf{C}_c]\{\dot{\mathbf{Y}}(t)\} + [\mathbf{K}_c]\{\mathbf{Y}(t)\} = \{\mathbf{F}(t)\} \quad (2)$$

where matrices $[\mathbf{M}_c] = [\mathbf{M}_u] + [\mathcal{M}_c]$, $[\mathbf{C}_c] = [\mathbf{C}_u] + [\mathcal{C}_c]$ and $[\mathbf{K}_c] = [\mathbf{K}_u] + [\mathcal{K}_c]$ now contain the effects of the modifications m_c , c_c and k_c in the matrices $[\mathcal{M}_c]$, $[\mathcal{C}_c]$ and $[\mathcal{K}_c]$. In the present paper, the modification matrices are considered diagonal for simplicity. Nevertheless, for practical purpose, the approach can accommodate more sophisticated structural modification configuration. In such a case, the problem is not so immediate and demands a proper rewriting of the modification matrices before applying the technique as done in [3].

To analyse the effects of the modifications on a structure, several techniques have been proposed, involving either modal parameter or frequency response data. However, the use of FRFs is particularly attractive for real structures since only the FRFs at the constrained locations are needed in practice [3, 5]. By definition, the transfer function is related to the mass, damping and stiffness matrices by a simple inversion, given by:

$$[\mathbf{H}_X(\omega)] = [-\omega^2[\mathbf{M}_X] + j\omega[\mathbf{C}_X] + [\mathbf{K}_X]]^{-1} \quad (3)$$

where X refers to the free or constrained configuration. Other advantage of the FRFs technique lies in the

fact that it is neither prone to modal truncation or modal identifications errors, offering robustness to the identification results. Of practical interest is that one can derive the constrained transfer functions directly from the unconstrained ones through the relation [6]:

$$[\mathbf{H}_c(\omega)] = [[\mathbf{I}] + [\mathbf{H}_u(\omega)] (-\omega^2[\mathcal{M}_c] + j\omega[\mathcal{C}_c] + [\mathcal{K}_c])]^{-1} [\mathbf{H}_u(\omega)] \quad (4)$$

or, conversely:

$$[\mathbf{H}_u(\omega)] = [[\mathbf{I}] + [\mathbf{H}_c(\omega)] (\omega^2[\mathcal{M}_c] - j\omega[\mathcal{C}_c] - [\mathcal{K}_c])]^{-1} [\mathbf{H}_c(\omega)] \quad (5)$$

where $[\mathbf{I}]$ is the identity matrix. Once known the modification matrices, Eq.(5) allows to deduce the transfer functions of the free system from computed or measured FRFs.

3. INVERSE IDENTIFICATION PROCEDURE

When we now turn to consider the given identification problem, it seems attractive to reformulate the dynamics of the original system in terms of the constrained configuration:

$$([\mathbf{M}_c] - [\mathcal{M}_c])\{\ddot{\mathbf{Y}}_u(t)\} + ([\mathbf{C}_c] - [\mathcal{C}_c])\{\dot{\mathbf{Y}}_u(t)\} + ([\mathbf{K}_c] - [\mathcal{K}_c])\{\mathbf{Y}_u(t)\} = \{\mathbf{F}(t)\} \quad (6)$$

Assuming an external harmonic excitation, the free system executes an exponentially damped harmonic oscillation, satisfying the following equation:

$$(-\omega^2[\mathbf{M}_c] + j\omega[\mathbf{C}_c] + [\mathbf{K}_c])\{\mathbf{Y}_u(\omega)\} - (-\omega^2[\mathcal{M}_c] + j\omega[\mathcal{C}_c] + [\mathcal{K}_c])\{\mathbf{Y}_u(\omega)\} = \{\mathbf{F}(\omega)\} \quad (7)$$

which can be recast into the form, by using Eq.(3):

$$([\mathbf{I}] - [\mathbf{H}_c(\omega)] [-\omega^2[\mathcal{M}_c] + j\omega[\mathcal{C}_c] + [\mathcal{K}_c]])\{\mathbf{Y}_u(\omega)\} = [\mathbf{H}_c(\omega)]\{\mathbf{F}(\omega)\} \quad (8)$$

where $[\mathbf{H}_c(\omega)]$ are FRFs which contain all the effects related to the physical mass, damping and stiffness properties of the system under constrained configuration. An interesting feature in Eq.(8) is that, under no external excitation, it represents an eigenvalue problem

$$\{\mathbf{Y}_u(\omega)\} = [[\mathbf{H}_c(\omega)] (-\omega^2[\mathcal{M}_c] + j\omega[\mathcal{C}_c] + [\mathcal{K}_c])] \{\mathbf{Y}_u(\omega)\} \quad (9)$$

from which the modal frequencies ω_{un} and modeshapes $\{\phi_{un}\}$ of the free system can be computed. Based on this relation, the modal properties of the free system can be related with the data obtained from the constrained configuration since

$$\{\phi_{un}\} = [[\mathbf{H}_c(\omega_{un})] (-\omega_{un}^2[\mathcal{M}_c] + j\omega_{un}[\mathcal{C}_c] + [\mathcal{K}_c])] \{\phi_{un}\} \quad (10)$$

At this point, it should be recalled that structural modification is carried out locally so that only the FRFs information at the modifications points are needed [3]. In practice, this approach, originally presented by Li and He [3] for conservative systems ($c_p=0, \forall p$), allows to determine either mass or stiffness modifications to score a desired modal behaviour. As noted in [3], the matrices $[\mathcal{M}_c]$ and $[\mathcal{K}_c]$, and thus also $[\mathcal{C}_c]$ for the given case, may commute with $\{\phi_{un}\}$ so that,

$$\{\phi_{un}^*\} = [\mathbf{H}_c^*(\omega_{un})][\phi_{un}^*] (-\omega_{un}^2 \{m_c\} + j\omega_{un} \{c_c\} + \{k_c\}) \quad (11)$$

where $[\mathbf{H}_c^*(\omega_{un})]$ is a fully populated matrix, $[\phi_{un}^*]$ is a diagonal matrix built from the terms of $\{\phi_{un}^*\}$, while $\{m_c\}$, $\{c_c\}$ and $\{k_c\}$ are vectors containing the terms of the diagonal constraints matrices $[\mathcal{M}_c]$, $[\mathcal{C}_c]$ and $[\mathcal{K}_c]$, and in general, $*$ denotes quantities at the constraint coordinates. In the form of Eq.(11), the problem of determining the modifications needed to change a dynamical behaviour does not involve to solve the original eigenvalue problem (9) but only a set of linear equations. Accounting for the fact that damping has only a small control on the determination of the modal frequencies for most of systems of interest, a reasonable estimation for mass-only modifications is given by:

$$\{m_c\} = \Re \left(-\frac{1}{\omega_{un}^2} [[\mathbf{H}_c^*(\omega_{un})][\phi_{un}^*]]^{-1} \{\phi_{un}^*\} \right) \quad (12)$$

Similarly, if only stiffness modifications are considered, then Eq.(11) leads to:

$$\{k_c\} = \Re \left([[\mathbf{H}_c^*(\omega_{un})][\phi_{un}^*]]^{-1} \{\phi_{un}^*\} \right) \quad (13)$$

which are the results given by Li and He [3].

4. APPLICATION TO PERFECT 3D AXISYMMETRIC STRUCTURES

4.1 Dynamical modelling of cylindrical shell

Shell vibration generally involves a combination of longitudinal and transverse motions. The normal modes $u(x, \theta, t)$ of the shell wall in the transverse direction can be described as the product of a beam function in the axial direction $U_m(x)$ (with appropriate boundary conditions), with a circumferential function $\varphi_n(\theta)$ in the horizontal plane, taking the form:

$$\vec{u}(x, \theta) = A_{mn} U_m(x) \vec{\varphi}_n(\theta) \quad (14)$$

where A_{mn} is an amplitude coefficient, and m and n relate to the number of nodal meridians extending over the top of the shell and the number of nodal circles respectively. To be more accurate, circumferential motions present both normal and tangential displacements since an excitation along one direction is able to excite motions in both directions, so that:

$$\vec{\varphi}_n(\theta) = \varphi_n^r(\theta) \vec{e}_r + \varphi_n^t(\theta) \vec{e}_t \quad (15)$$

where $\varphi_n^r(\theta) = \cos(n\theta)$ and $\varphi_n^t(\theta) = -\sin(n\theta)/n$. One must also remind that perfectly axi-symmetrical structures

exhibit normal modes occurring in degenerate pairs, which become non-degenerate when slight alteration of the symmetry is introduced. Of course, these two orthogonal modal families are an important aspect of the modelling for the correct dynamical description of axisymmetric bodies. Assuming of maximum of M nodal lines in the axial direction and N nodal lines in the circumferential plane, the size of the modal basis is therefore $M \times 2N$.

In what follows, the original structure considered is a free cylinder, open on the lower end and fitted with a clamped cover on the upper end, for which a modal model is considered. The original structure is then constrained at several locations of the open end, using point masses, dampers and stiffnesses. For both constrained and unconstrained configurations, the useful transfer functions $[\mathbf{H}(\omega)]$ are computed from:

$$[\mathbf{H}(\omega)] = [\mathbf{U}] [-\omega^2 [\mathbf{M}] + j\omega [\mathbf{C}] + [\mathbf{K}]]^{-1} [\mathbf{U}]^T \quad (16)$$

where $[\mathbf{M}]$, $[\mathbf{C}]$ and $[\mathbf{K}]$ are diagonal matrices of the modal parameters and $[\mathbf{U}]$ is the modal matrix of the corresponding modeshapes. For the unconstrained system, approximations of the natural frequencies and modeshapes used in the simulations, solutions of the Donnel-Mushtari-Vlasov equations, are found in [7]. However, it should be mentioned that the Donnel-Mushtari-Vlasov model is here used for convenience, in order to provide the data to perform the identifications, and is clearly not intended to describe the bell vibrations for which the thin shell theory cannot apply. Based on this, examples of computed modeshapes of the free structure are presented in Figure 1 (see Section 5 for detail). The natural frequencies

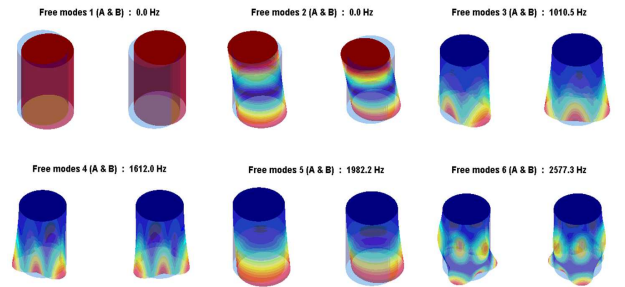


Figure 1. Natural modes of the unconstrained system.

and modeshapes for the constrained configuration are then computed using a penalty formulation of the modal parameters of the unconstrained system. As can be shown, they are solutions of the following eigenvalue problem:

$$\begin{aligned} & (-\omega^2 ([\mathbf{M}_u] + [\mathbf{U}_u]^T [\mathcal{M}_c] [\mathbf{U}_u]) + \\ & + j\omega ([\mathbf{C}_u] + [\mathbf{U}_u]^T [\mathcal{C}_c] [\mathbf{U}_u]) + \\ & + [\mathbf{K}_u] + [\mathbf{U}_u]^T [\mathcal{K}_c] [\mathbf{U}_u]) \{ \mathbf{Y} \} = 0 \end{aligned} \quad (17)$$

from which one deduces the corresponding transfer functions by using Eq.(16).

4.2 Identification of the unconstrained modes and constraints

Based on the ideas developed in section 3, Debut et al. [2] recently demonstrated the feasibility of such an approach for conservative 2D axi-symmetric discrete structures. Interestingly, by enforcing essential properties of all perfect axisymmetric unconstrained bodies during the inversion, identification of the local mass and stiffness constraints and modal properties of the free system, can be achieved. Indeed, implicit in the technique is the fact that all free axisymmetric structures display rigid-body modes, as well as modal doublets, with known similar mode shapes and near-identical frequencies.

For the case of 3D axi-symmetric structures, it should be obvious by now that an important aspect of the modelling is to account for both radial and tangential motions of the structure and also to include the two orthogonal families of modes in the modal basis. In particular, this requires to account for a set of FRFs functions relating all constraint coordinates in accordance with each possible type of excitation and motion. It thus leads to consider *radial-to-radial*, *radial-to-tangential*, *tangential-to-radial* as well as *tangential-to-tangential* point-to-point transfer functions, each family of FRFs having the generic form:

$$H_{ij}^{pq}(\omega) = \sum_n \frac{\phi_n^p(\theta_i)\phi_n^q(\theta_j)}{m_n(\omega_n^2 + j\omega 2\omega_n\zeta_n - \omega^2)} \quad (18)$$

where H_{ij}^{pq} is the frequency response of the structure measured at point θ_j in the direction \vec{e}_q considering an excitation at point θ_i in the direction \vec{e}_p , with $p, q=r$ or t .

As presented in [4], the identification scheme can be further generalized by writing Eq.(9) in the expanded form:

$$\begin{Bmatrix} \{\phi_{un}^{r*}\} \\ \{\phi_{un}^{t*}\} \end{Bmatrix} = \begin{bmatrix} \mathbf{H}_c^{rr*}(\omega) & \mathbf{H}_c^{rt*}(\omega) \\ \mathbf{H}_c^{tr*}(\omega) & \mathbf{H}_c^{tt*}(\omega) \end{bmatrix} \begin{bmatrix} \Phi_{un}^{r*} & 0 \\ 0 & \Phi_{un}^{t*} \end{bmatrix} \left(-\omega^2 \begin{Bmatrix} \{m_c^r\} \\ \{m_c^t\} \end{Bmatrix} + j\omega \begin{Bmatrix} \{c_c^r\} \\ \{c_c^t\} \end{Bmatrix} + \begin{Bmatrix} \{k_c^r\} \\ \{k_c^t\} \end{Bmatrix} \right) \quad (19)$$

where \mathbf{H}_c^{pq*} is a submatrix of FRFs relating the responses of the constrained structure in the q -direction with the excitations in the p -direction, between all constraint coordinates, Φ_{un}^{r*} and Φ_{un}^{t*} are diagonal matrices built from the terms of $\{\phi_{un}^{r*}\}$ and $\{\phi_{un}^{t*}\}$ respectively, while $\{m_c^r\}$, $\{m_c^t\}$, $\{c_c^r\}$, $\{c_c^t\}$, $\{k_c^r\}$ and $\{k_c^t\}$ are the mass, damping and stiffness modifications in the radial and tangential directions. The inversion now proceeds as follows:

1. First, one computes the modal frequencies of the unconstrained system by using the fact that the free system displays double vibrational modes occurring in orthogonal pairs with identical frequencies. We therefore consider two similar orthogonal modeshapes of near-identical frequencies (e.g $\omega_{un} \approx \omega_{u(n+1)}$), and then write Eq.(19) for each mode to

finally impose a compatibility condition so that:

$$\begin{aligned} & \left(\begin{bmatrix} \mathbf{H}_c^{rr*}(\omega_{u(n+1)}) & \mathbf{H}_c^{rt*}(\omega_{u(n+1)}) \\ \mathbf{H}_c^{tr*}(\omega_{u(n+1)}) & \mathbf{H}_c^{tt*}(\omega_{u(n+1)}) \end{bmatrix} \begin{bmatrix} \Phi_{u(n+1)}^{r*} & 0 \\ 0 & \Phi_{u(n+1)}^{t*} \end{bmatrix} \right)^{-1} \begin{Bmatrix} \{\phi_{u(n+1)}^{r*}\} \\ \{\phi_{u(n+1)}^{t*}\} \end{Bmatrix} \\ & - \left(\begin{bmatrix} \mathbf{H}_c^{rr*}(\omega_{un}) & \mathbf{H}_c^{rt*}(\omega_{un}) \\ \mathbf{H}_c^{tr*}(\omega_{un}) & \mathbf{H}_c^{tt*}(\omega_{un}) \end{bmatrix} \begin{bmatrix} \Phi_{un}^{r*} & 0 \\ 0 & \Phi_{un}^{t*} \end{bmatrix} \right)^{-1} \begin{Bmatrix} \{\phi_{un}^{r*}\} \\ \{\phi_{un}^{t*}\} \end{Bmatrix} = \{0\} \end{aligned} \quad (20)$$

which is a nonlinear problem but can be solved for ω_{un} through a systematic frequency sweep. Notice that, for free axisymmetric structures, the modeshape functions at the interface where constraints will be applied can always be safely assumed following Eq.(15). The process can then be repeated by assuming other orthogonal modeshapes so that their corresponding degenerate modal frequencies could be determined.

2. The second step consists in recovering simultaneously both mass and stiffness modifications. Again, the starting point is from Eq.(19), applied now to two distinct modeshapes (with index number n and m) with already identified modal frequencies. This yields to a set of two equations which can be recast in a compact matrix form as:

$$\begin{Bmatrix} \{\Phi_{un}^{r*}\} \\ \{\Phi_{um}^{r*}\} \end{Bmatrix} = \Re \left(\begin{bmatrix} [\Psi_n^*] & 0 \\ 0 & [\Psi_m^*] \end{bmatrix} \begin{bmatrix} -\Omega_n & I \\ I & -\Omega_m \end{bmatrix} \right) \begin{Bmatrix} \{M_c\} \\ \{K_c\} \end{Bmatrix} \quad (21)$$

denoting

$$\{\Phi_{uX}^{r*}\} = \begin{Bmatrix} \{\phi_{uX}^{r*}\} \\ \{\phi_{uX}^{t*}\} \end{Bmatrix}$$

and

$$[\Psi_X^*] = \begin{bmatrix} \mathbf{H}_c^{rr*}(\omega_{uX}) & \mathbf{H}_c^{rt*}(\omega_{uX}) \\ \mathbf{H}_c^{tr*}(\omega_{uX}) & \mathbf{H}_c^{tt*}(\omega_{uX}) \end{bmatrix} \begin{bmatrix} \Phi_{uX}^{r*} & 0 \\ 0 & \Phi_{uX}^{t*} \end{bmatrix}$$

where Ω_X is a diagonal matrix containing the modal frequencies of the considered mode. Finally, Eq.(21) allows to compute the vectors $\{M_c\}$ and $\{K_c\}$, which contain the mass $\{m_c\}$ and stiffness $\{k_c\}$ modifications in both directions, by a simple inversion,

$$\begin{Bmatrix} \{M_c\} \\ \{K_c\} \end{Bmatrix} = \Re \left(\left(\begin{bmatrix} [\Psi_n^*] & 0 \\ 0 & [\Psi_m^*] \end{bmatrix} \begin{bmatrix} -\Omega_n & I \\ I & -\Omega_m \end{bmatrix} \right)^{-1} \begin{Bmatrix} \{\Phi_{un}^{r*}\} \\ \{\Phi_{um}^{r*}\} \end{Bmatrix} \right) \quad (22)$$

3. The last step of the identification procedure concerns the estimations of the local damping modifications, which are given by applying Eq.(19) to one arbitrarily identified unconstrained modes (index n), and retaining the imaginary terms only,

$$\begin{aligned} \{c_c\} &= \Im \left(\frac{1}{j\omega_{un}} \right. \\ & \left. \left(\begin{bmatrix} \mathbf{H}_c^{rr*}(\omega_{un}) & \mathbf{H}_c^{rt*}(\omega_{un}) \\ \mathbf{H}_c^{tr*}(\omega_{un}) & \mathbf{H}_c^{tt*}(\omega_{un}) \end{bmatrix} \begin{bmatrix} \Phi_{un}^{r*} & 0 \\ 0 & \Phi_{un}^{t*} \end{bmatrix} \right)^{-1} \begin{Bmatrix} \{\phi_{un}^{r*}\} \\ \{\phi_{un}^{t*}\} \end{Bmatrix} \right) \end{aligned} \quad (23)$$

As seen, it should be pointed out that, in the above development, it was implicitly assumed that inertial and stiffness modifications only affects the real part of the complex frequency responses while local dissipation effects influences their imaginary part only.

One useful aspect of the above-described procedure is the possibility to specify arbitrarily the unconstrained modes to consider for the determination of the modal frequencies and constraints. Actually, Eq.(20) is valid for any pair of modes and one can determine several modal frequencies (step 1) to consequently ensure the reliability of the identification results of the constraints (steps 2-3). Furthermore, the form of Eq.(22) reveals that the technique can be used to prescribe several natural frequencies from a limited number of local modifications. In that case, the use of a pseudo-inversion is necessary and therefore best-fit estimates of the modifications will be computed. Overall, the technique has the significant advantage that it does not imply any modal parameter identification of the constrained system. Actually, some errors always occur from either a faulty modal data acquisition procedure or a bias in the numerical algorithm, which may consequently have perverse effects on the identification results. Also related to the modelling aspect, a third advantage is that the method is not prone to truncation error as common for modal-based approach identification technique [8]. Last but not least, notice that a fourth advantage of the proposed approach resides in the fact that the formulation is not limited to small perturbation effects.

5. ILLUSTRATIVE COMPUTATIONS

The system considered consists on a closed clamped-free circular shell with geometrical and physical data: mean radius 0.1m, height 0.25m, thickness 0.01m, module Young 10^{11} N.m⁻², density 8500 kg.m⁻³ and Poisson coefficient 0.3. The cylinder has been divided into 64 locations in the circumferential plane, and mass, damping and stiffness constraints have been applied in both radial and tangential directions, at locations 1, 16, 32, and 48. A modal damping of 0.01% was used for all modes. For illustration, Figure 2 displays some of the first modes of the constrained system (obtained from a conservative computation), which gives evidence of all the significant differences displayed by the computed modes (compare with Figure 1). Of course, a similar computation accounting for local dampers would result in complex modes: the larger the damping coefficient, the more complex the modes. The values of the imposed mass, damping and stiffness modifications are given in Table 1.

For illustration, Figure 3 shows the absolute values of radial-to-radial and tangential-to-tangential FRFs collected at node 16 by applying an excitation at node 1, for both the unconstrained and constrained configurations. Finally, Figures 4 and 5 display the superposition of the computed and identified transfer functions for the unconstrained system, measured in the same conditions. The numerical values of the identified mass, damping and

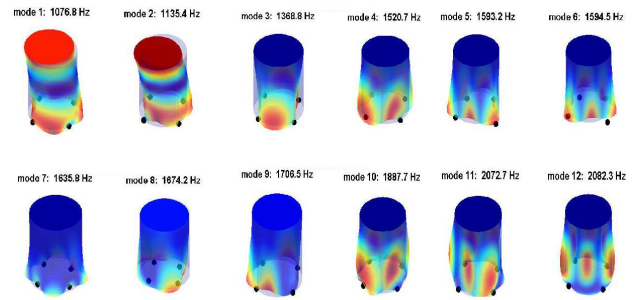


Figure 2. Natural modes of the conservative constrained system. Constraint locations appears as large black circles.

stiffness constraints at each node location are given in Table 2, as well as the modal frequencies identified and then used during the inversion procedure.

6. CONCLUSIONS

In this paper we proposed a method for the identification of the free modes of damped 3D axisymmetric structures - for instance carillon bells - as well as of the constraint mass, damping and stiffness parameters, based on the dynamical responses of the constrained system at the constraint locations. The problem was formulated in terms of structural modification concepts and the identification performed using a computationally straightforward scheme accounting for essential features of axisymmetric structures. Obviously, the overall results, based on simulated data, are over-impressive when compared with what can be reasonably expected from experimental data. Although such analysis may appear overambitious for real situations, the results presented in this paper are quite encouraging for using experimental data measured on a laboratory setup. Experiments are being prepared in order to validate the discussed identification technique.

7. ACKNOWLEDGEMENTS

The authors thank M. Haddar and F. Chaari for fruitful discussions. This work was supported by the Portuguese FCT, under a bi-lateral grant Ref. Tunisia124639209928430.

8. REFERENCES

- [1] Rossing, T. D. "Science of Percussion Instruments". World Scientific Publishing, Singapore, 2007.
- [2] V. Debut, M. Carvalho and J. Antunes. "Recovering the unconstrained modes of axisymmetric structures from measurements under constrained condition". *Proceeding of the 19th International Congress on Sound and Vibration*, Vilnius, Lithuania, 2012.
- [3] T. Li, and J. He. "Local structural modification using mass and stiffness changes". *Engineering Structures*, 1999, Vol.21, pp.1028-1037.
- [4] J. Antunes, V. Debut and M. Carvalho. "Identification of the unconstrained modes of 3D axisymmetric structures from measurements under constraining

Element	Mass (Kg)		Damping (Ns/m)		Stiffness (N/m)	
	rad.	tang.	rad.	tang.	rad.	tang.
1	10	5	100	50	10^9	5×10^8
16	5	2.5	500	250	5×10^8	2.5×10^8
32	10	5	200	100	10^9	5×10^8
48	5	2.5	300	150	5×10^8	2.5×10^8

Table 1. Physical mass, damping and stiffness modifications.

Location	Mass (Kg)		Damping (Ns/m)		Stiffness (N/m)		Modal frequencies (Hz)	
	rad.	tang.	rad.	tang.	rad.	tang.	Exact	Identified
1	9.99	4.99	101	54	10^9	5×10^8	1010.5	1009.2
16	4.99	2.50	500	253	5×10^8	2.5×10^8	1612.0	1611.3
32	9.99	5.00	201	103	10^9	5×10^8	2742.3	2741.5
48	4.99	2.49	301	154	5×10^8	2.5×10^8	4760.4	4759.0

Table 2. Left: identified physical mass, damping and stiffness modification values for both radial and tangential directions. Right: exact and identified modal frequencies of the free system computed from the inversion procedure

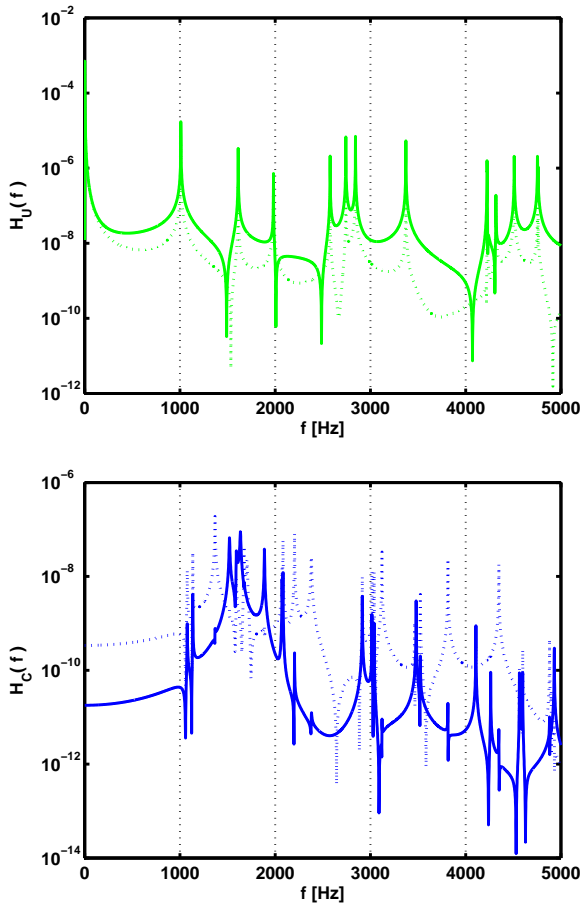


Figure 3. Computed radial-to-radial (solid) and tangential-to-tangential (dot) transfer functions. Up: unconstrained system. Bottom: constrained system. $x_i = 1$, $x_j = 16$.

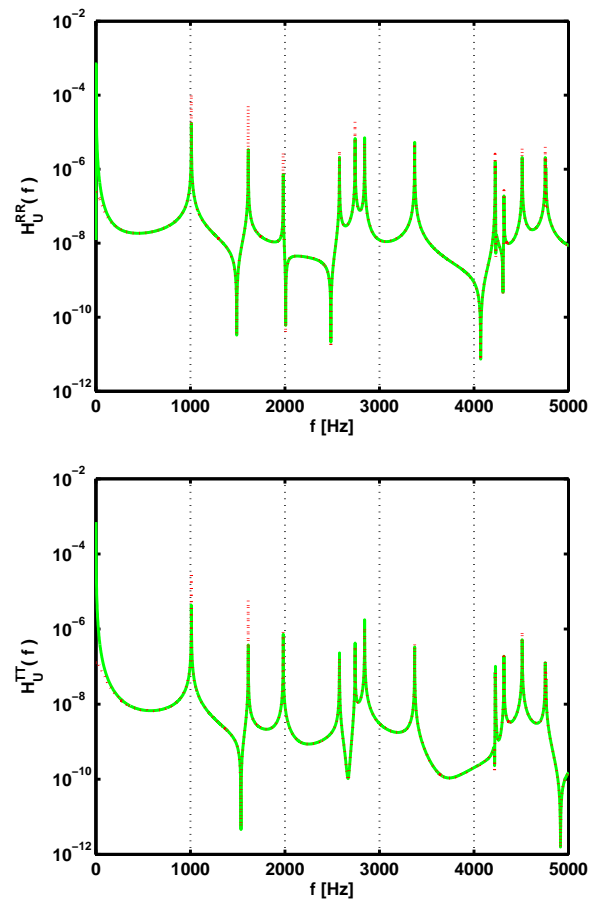


Figure 4. Computed (green) and identified (red) transfer functions of the unconstrained system. Up: radial-to-radial. Bottom: tangential-to-tangential $x_i = 1$, $x_j = 16$.

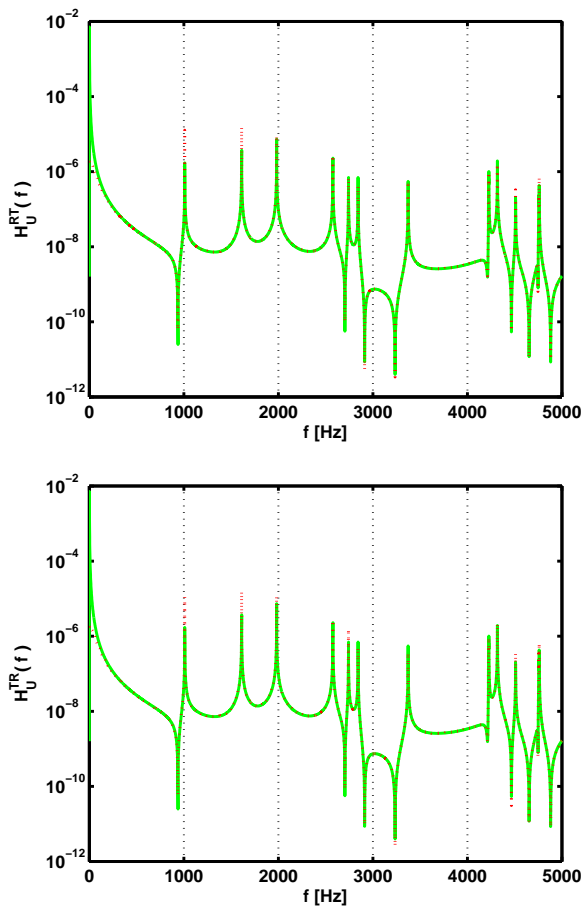


Figure 5. Computed (green) and identified (red) transfer functions of the unconstrained system. Up: radial-to-tangential. Bottom: tangential-to-radial. $x_i = 1$, $x_j = 16$.

support conditions”. To be published in *Proceeding of the 5th International Congress Design and Modelling of Mechanical Systems*, Djerba, Tunisia, 2013.

- [5] J. He. “Structural modification”. *Philosophical Transactions of the Royal Society*, 2001, Vol.A-359, pp.187-204.
- [6] H. N. Ozguven. “Structural modifications using frequency response functions”. *Mechanical Systems and Signal Processing*, 1990, Vol.4, pp.53-63.
- [7] W. Soedel. “Shells”. *Encyclopedia of Vibrations*, Elsevier, 2001.
- [8] P. Avitabile. “Twenty years of structural dynamic modification: a review”. *Sound and vibration*, 2003, Vol.37, pp.14-25.

MODELING A VIBRATING STRING TERMINATED AGAINST A BRIDGE WITH ARBITRARY GEOMETRY

Dmitri Kartofelev, Anatoli Stulov

Institute of Cybernetics at
Tallinn University of Technology,
Tallinn, Estonia
dima@cs.ioc.ee

Heidi-Maria Lehtonen, Vesa Välimäki

Department of Signal Processing and
Acoustics, Aalto University,
Espoo, Finland
heidi-maria.lehtonen@aalto.fi

ABSTRACT

This paper considers dynamic string motion in which the displacement is unilaterally constrained by the termination condition with an arbitrarily chosen geometry. A digital waveguide model is proposed for simulating the nonlinearity inducing interactions between the vibrating string and the contact condition at the point of string termination. The current work analyzes the resulting string motion influenced by the contact conditions with mostly flat but slightly curved geometries. The effect of a minute imperfection of the termination condition on the string vibration is investigated. It is shown that the lossless string vibrates in two distinct vibration regimes. In the beginning the string starts to interact in a nonlinear fashion with the bridge, and the resulting string motion is nonperiodic. The duration of that vibration regime depends on the geometry of the bridge. After some time of nonperiodic vibration, the string vibration settles in a periodic regime. Presented results are applicable for example in the physics-based sound synthesis of stringed musical instruments, such as the shamisen, biwa, sitar, tambura, veena or even the bray harp and the grand piano.

1. INTRODUCTION

In numerous musical instruments the collision of a vibrating string with rigid spatial obstacles, such as frets or a curved bridge, are crucial to the tonal quality of the produced sounds. Lutes such as the shamisen, biwa, sitar, tambura or veena have a very distinctive sound which can be described as *buzzing*. The form of the bridge used in these instruments is quite different from that usually found in most stringed instruments, since the profiles of the bridges are slightly curved, almost flat (see Fig. 1). The spatial extent of the bridges along the direction of the string is relatively large compared to the speaking length of the strings themselves [1].

A similar mechanism is also not unknown in Western instruments. The treble strings of a grand piano usually terminate at the capo bar (*capo d'astro*). The V-shaped sec-

tion of the capo bar has a parabolic curvature, and although the area to which the string rapidly touches while vibrating is small compared to the string's speaking length, it was shown by Stulov and Kartofelev [2] that the capo bar has a noticeable effect on the piano tone formation.

Also the Medieval and Renaissance bray harp has small bray-pins which provide a metal surface for the vibrating string to impact, increasing the upper partial content in the tone and providing a means for the harp to be audible in larger spaces and in ensemble with other instruments [3]. It is evident that for realistic physically informed modeling of these instruments such nonlinearity inducing interactions need to be examined and simulated accurately.

Raman was the first to study the effect and identify the bridge as the main reason for distinctive sounding of the tambura and veena [4]. Over the years many authors have solved this problem using different approaches [3], [5] – [15]. An overview and comparison of the existing methods that are proposed for modeling the interaction between the termination and the string are presented by Vyasarayani, *et al.* in [3].

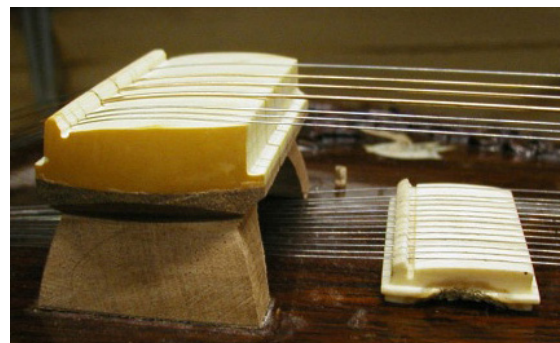


Figure 1. *Jawari*, the main bridge of the sitar and *taraf ka ghoraj*, the sympathetic string bridge.

The aim of the current paper is to model the vibration of the string which is unilaterally constrained at one of the points of string termination. Dynamic motion of the plucked ideal string against the termination condition (TC) with three different profile geometries are simulated and obtained results are examined. In addition, a method for quantifying the effect of minute geometric imperfections of the mostly flat bridge on the string vibration is provided.

Although the cases examined here are for bridges with mostly flat profile geometries, the obtained conclusions hold

to some degree for cases where the bridge profile geometries are more versatile, cf. [2].

Compared to the previously published work, we propose a new and relatively simple approach for modeling the non-linear bridge-string interaction and consequently the dynamic motion of the entire string. In this work the proposed model is demonstrated using physical parameters that are obtained from a Chinese lute biwa, thus presenting an applied approach.

2. IDEAL STRING

For analyzing the phenomenon of interest, it is sufficient to describe the dynamic motion of the string using the *ideal* string. Phenomena such as losses or dispersion are discarded. We consider the wave equation for the linear and lossless flexible string:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

with $u(0, t) = u(L, t) = 0$, where L and $u(x, t)$ are the speaking length and the displacement of the string, respectively. In (1) the value $c = \sqrt{T/\mu}$ is the speed of the traveling waves on the string, where T is the tension and μ is the linear mass density of the string [16].

String parameters for all the calculations in the current paper are the same as used by Taguti in [8]. Taguti investigated a biwa string with the following parameters: string length $L = 0.8$ m; linear mass density $\mu = 0.375$ g/m; string tension $T = 38.4$ N. From here it follows that the speed of the traveling waves along the string is $c = 320$ m/s and the fundamental frequency of such a string is $f_0 = 200$ Hz.

3. STRING EXCITATION

The string plucking condition can be introduced as follows. We assume that at the point $x = l = 3/4L$ the emerging traveling wave is of the form

$$u(l, t) = \begin{cases} A \left(\frac{t}{t_0} \right)^2 e^{2(1-t/t_0)}, & \text{if } 0 \leq t \leq t_0, \\ A, & \text{if } t > t_0. \end{cases} \quad (2)$$

In (2) $A = 1$ cm is the amplitude of the outgoing traveling wave and the duration of the excitation is $t = t_0 = 4$ ms. Selection of the plucking condition (2) ensures that the plucking force acting on the string point $x = l$ ceases if $t \geq t_0$ (time derivative of (2) is proportional to the plucking force) [2].

It can be shown that (1) may be satisfied by superposition of nondispersive traveling waves $u_r(t - x/c)$ and $u_l(t + x/c)$ moving in either directions along the string emerging from the plucking point $x = l$. At this point $u_r(l, t) = u_l(l, t) = u(l, t)$. These two waves u_r and u_l are simply a translation of the plucking condition (2) from the point $x = l$ to other segments of the string [16].

In the case of ideal rigid string termination where no TC is present, the boundary value $u(0, t) = u(L, t) = 0$ is satisfied if the wave $u_r(t - x/c)$ propagating to the right at the

point $x = L$ creates the wave $u_l(t + x/c) = -u_r(t - x/c)$ moving to the left and the wave $u_l(t + x/c)$ propagating to the left at the point $x = 0$ creates the wave $u_r(t - x/c) = -u_l(t + x/c)$ moving to the right. This procedure can be interpreted as equivalent to the digital waveguide approach [17, 18, 19].

It follows that for the current model the string displacement $u(x, t)$ at any point x of the string and for all time instants t is simply the resulting sum of waveforms u_r and u_l moving in both directions

$$u(x, t) = u_r \left(t - \frac{x}{c} \right) + u_l \left(t + \frac{x}{c} \right). \quad (3)$$

The method for modeling the bridge-string interaction is explained in Sections 5 and 6.

4. BRIDGE GEOMETRY

Slightly curved bridges of the lutes mentioned in Section 1 are usually located at the far end of the neck. Similarly the geometric contact condition (TC) is located at the termination point of the string. Figure 2 shows the traveling waves u_r and u_l , string displacement $u(x, t)$, and the location of the rigid termination (bridge) relative to the string.

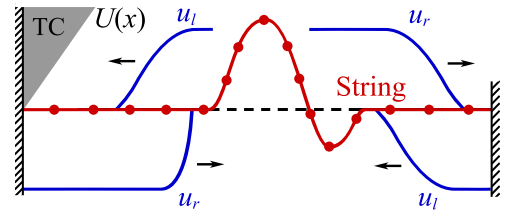


Figure 2. Scheme of the string displacement model. The traveling waves u_r and u_l (solid lines), and the forms of the string (solid lines marked by circles). Position of the TC relative to the string is shown using gray formation.

4.1 Case 1: Linear bridge with a sharp edge

The function $U(x)$ that describes the profile of a flat bridge is calculated as follows

$$U(x) = \begin{cases} kx, & \text{if } x \leq x_c, \\ \infty, & \text{if } x > x_c, \end{cases} \quad (4)$$

where $k = \tan \theta = 0.008$ is the slope of the linear function where $\theta \approx 0.008$ rad $\approx 0.46^\circ$. Value $x_c = 15$ mm marks the coordinate of the truncation of the linear function.

4.2 Case 2: Linear bridge with a curved edge

The profile of a bridge with a curved edge is calculated as follows

$$U(x) = \begin{cases} kx, & \text{if } x \leq x_b, \\ \frac{1}{2R}(x - x_b)^2, & \text{if } x > x_b, \end{cases} \quad (5)$$

where the parameter k has the same value and meaning as in the previous case. Parameter $R = 10$ mm is the curvature radius of the corresponding parabolic function

$f(x) = (2R)^{-1}x^2$ at its minimum. Coordinate $x_b = 10$ mm marks the transition between linear and parabolic parts of the geometry.

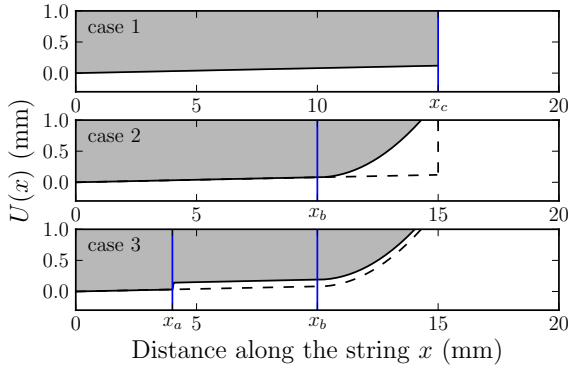


Figure 3. Termination condition geometry for the cases under study. Solid vertical lines mark the positions of the characteristic values x_a , x_b and x_c . Case 1: linear bridge with a sharp edge. Case 2: linear bridge with a curved edge, the dashed line shows the profile of the case 1 for comparison. Case 3: bridge with a small geometric imperfection, the dashed line shows the case 2 for comparison.

4.3 Case 3: Bridge with a geometric imperfection

The bridge in this case is similar to the previous case with the exception of an addition of small imperfection in the form of discontinuity in the linear part of the TC in (5). The bridge profile geometry for this case can be expressed in the following form

$$U(x) = \begin{cases} kx, & \text{if } x \leq x_a, \\ kx + y, & \text{if } x_a < x \leq x_b, \\ \frac{1}{2R}(x - x_b)^2 + K, & \text{if } x > x_b, \end{cases} \quad (6)$$

where the parameters k and R have the same value and meaning as in the previous cases. Parameter $y = 0.11$ mm raises the value of linear function in the interval $x = (x_a, x_b]$ where $x_a = 4$ mm and $x_b = 10$ mm. Constant $K = kx_b + y$ is presented in order to preserve continuity of the form in vicinity of the point $x = x_b$.

The TC geometries presented in (4) - (6) are shown in Fig. 3.

5. BRIDGE-STRING INTERACTION MODEL

In order to model the bridge-string interaction we assume that the reflecting wave $u_r(t - x/c)$ moving to the right appears only at the point $x = x^*$, where the amplitude of the string deflection $u(x^*, t) \geq U(x^*)$. The position of this point x^* is determined by the TC geometry $U(x)$ in the following way. Since the termination is rigid we must have $u(x^*, t) = U(x^*)$, and this condition results in the appearance (addition) of reflected wave

$$u_r\left(t - \frac{x^*}{c}\right) = U(x^*) - u_l\left(t + \frac{x^*}{c}\right), \quad (7)$$

where the waves u_l and u_r correspond to any waves that have reflected on earlier time moments and are currently located at $x = x^*$. The proposed method ensures that the amplitude of the string deflection, which is determined by (3), will never exceed the value $U(x)$.

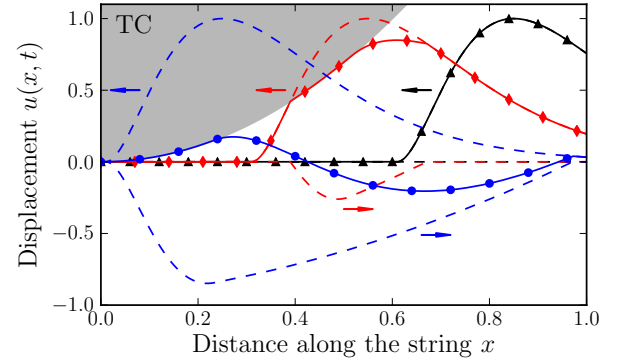


Figure 4. Reflection of the first wave from the termination. The traveling waves u_r and u_l (dashed lines), and the forms of the string (solid lines marked by signs) shown for consequent nondimensional moments of times $t_1 = 0.4$ (triangle); $t_2 = 0.7$ (diamond); $t_3 = 1.0$ (circle).

In Fig. 4 we demonstrate the form of the string in vicinity of the geometric termination during the reflection of the first wave $u_l(t + x/c)$ only. Using the procedure described above, the string deflection as a function of the nondimensional distance along the string is computed for three consequent normalized nondimensional ($c = 1$) moments of time. At the moment $t = t_1$ the form of the string (solid line marked by triangles) is determined only by the traveling wave u_l . At the next moment $t = t_2$ the small segment of the string is in contact with the surface of the termination, and the reflected wave $u_r(x, t_2)$ has appeared (dashed line). The corresponding form of the string deflection is shown by solid line marked with diamonds. At the moment $t = t_3$ the string is in contact with the surface of the termination on the segment closer to the string edge ($x = 0$). The form of the string at this moment is shown by solid line marked with circles, and the reflected wave $u_r(x, t_3)$ is also shown by the dashed line. Thus at some moments the string wraps or clings to the termination, and during that time the form of the string on some segment simply repeats the form of the termination.

6. NUMERICAL IMPLEMENTATION

The bridge-string interaction model and the ideal string vibration are implemented numerically by using discrete t - x space with the time mesh Δt and the space mesh Δx . Values for the Δt and Δx are selected so that

$$c \frac{\Delta t}{\Delta x} = 1, \quad (8)$$

where $c = \sqrt{T/\mu}$. Selection of the step-sizes Δx and Δt according to (8) ensures maximum accuracy of the result for any given resolution of the computational grid. Thus,

the transmission of the traveling waves u_r and u_l with respect to the points of the discrete t - x space are

$$u_r(x_n, t_m) = u_r(x_{n-1}, t_{m-1}), \quad (9)$$

$$u_l(x_n, t_m) = u_l(x_{n+1}, t_{m-1}), \quad (10)$$

where the index $n = 0, \dots, N$ corresponds to the discrete space points and the index $m = 0, \dots, M$ corresponds to the discrete time points. Values of the corresponding coordinates x and t in (9) and (10) can be calculated as $x = x_n = n\Delta x$ and $t = t_m = m\Delta t$, respectively.

In order to satisfy the boundary condition at the right side of the string, namely $u(L, t) = 0$, the mechanism presented in Sec. 3 is used. For every successive time moment t_m

$$u_l(x_N, t_m) = -u_r(x_N, t_{m-1}), \quad (11)$$

where $x_N = N\Delta x = L$.

The effect of the geometric TC on the string vibration can be implemented numerically as follows. According to Sec. 5 the traveling wave u_r only appears in the vicinity of the bridge at the discrete point $x_n = x_n^*$ where the amplitude of the string deflection $u(x_n^*, t) \geq U(x_n^*)$. Thus, for every successive time moment t_m and for all x_n^*

$$u_r(x_n^*, t_m) = u_r(x_n^*, t_m) - \Delta u, \quad (12)$$

where $\Delta u = u(x_n^*, t_m) - U(x_n^*)$. Expression of the form (12) is more suitable for the iterative numerical scheme used to generate the result compared to the expression (7) shown in Sec. 5. Finally, when the aforementioned operations are conducted the final form of the string's displacement with respect to the discrete computational grid takes the form

$$u(x_n, t_m) = u_r(x_n, t_m) + u_l(x_n, t_m). \quad (13)$$

Numerical parameters selected to calculate the presented results are: $\Delta x = 0.985$ mm, $\Delta t = 3.077$ μ s, number of spatial points $N = 816$, including spatial points dedicated for the bridge $N_{TC} = 25$, number of the time points $M = 130000$, from here it follows that the temporal sampling rate is 325 kHz. The relevant part of the computer code written using the Python programming language is available for examination at the supplementary web page of this article [20].

7. RESULTS AND DISCUSSION

Figure 5 shows the time series of the string deflection $u(l, t)$ computed at the plucking point $x = l$. Visual inspection of the string deflection $u(l, t)$ reveals that for all the presented cases the strings vibrate in two distinct regimes. The strong influence of the bridge on the string's motion is noticeable for a certain period of time, and its duration depends on the bridge geometry. During this time prolonging for $t = t_{np}$ the string vibrates in *nonperiodic regime*. One can clearly see that after the moment $t = t_{np}$ the behavior of the deflection $u(l, t)$ becomes seemingly highly periodic. Closer examination reveals that the string's displacement is actually still slightly changing and therefore is not absolutely

periodic (string continues to interact with the bridge) but the change is small and can be neglected. This regime is called here the *periodic vibration regime*. It must be noted that this almost periodic vibration regime is possible only when the bridge profile is mostly flat and the string is considered ideal and lossless.

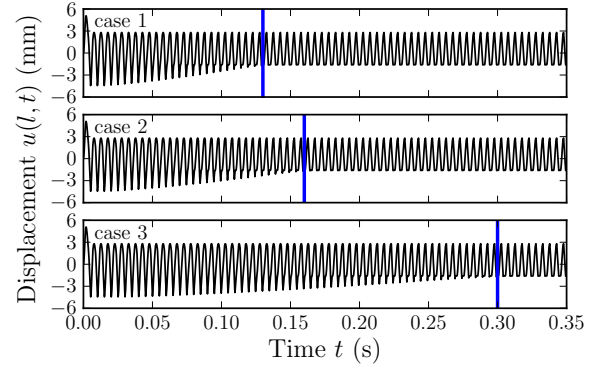


Figure 5. Time series of the string deflection $u(l, t)$ for the cases 1, 2, and 3. Nonperiodic and periodic vibration regimes are separated by vertical lines corresponding to the time moment $t = t_{np}$.

Table 1 shows the corresponding durations of the nonperiodic vibration regimes t_{np} for the cases under study. In addition, the corresponding number of string deflection $u(l, t)$ periods P_{np} are shown.

	t_{np} (s)	P_{np}
Case 1	0.13	26
Case 2	0.16	32
Case 3	0.30	60

Table 1. The duration and the number of the string deflection periods of the nonperiodic vibration regime.

The transitions between the nonperiodic and periodic regimes presented in Fig. 5 are also visible in the spectrograms presented in Fig. 6. All spectrograms are calculated using the Hanning window of the size 45 ms and the overlap value 55% of the window size. The animations of the simulated string vibration terminated against the bridges with profile geometries described in (4) – (6), are available for download on the supplementary web page of the current article [20].

7.1 Case 1: Linear bridge with a sharp edge

Spectrogram of the signal related to the case 1 is shown in Fig. 6a. Dashed vertical line corresponds to the duration of the nonperiodic vibration regime t_{np} of the string. It can be seen that during the nonperiodic vibration regime the energy of the lower vibration modes is being transferred to the higher modes. This effect of spectral widening can be noticed when comparing Figs. 6a and 7. Figure 7 shows the spectrogram of the corresponding linear case where no amplitude limiting TC is applied. Transfer of the energy

from lower to higher vibration modes is a sign of nonlinear behavior resulting from the interaction of the vibrating string and the bridge. This phenomenon is similar to the slapped bass effect [6] and the nonlinear limitation of the string amplitude by the damper in the part-pedaling effect in the grand piano [21, 22]. In the periodic vibration regime ($t > t_{np}$) the spectrum remains constant which is an expected result (*cf.* Fig. 5).

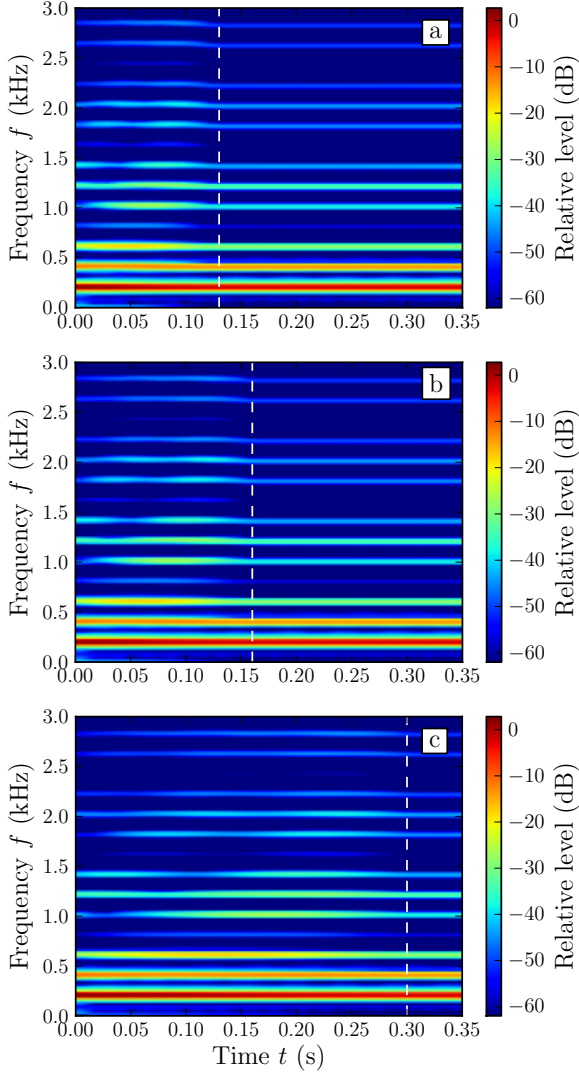


Figure 6. Spectrogram of the signal $u(l, t)$ for the cases: a) case 1, b) case 2 and c) case 3. The transition between nonperiodic and periodic vibration regimes at $t = t_{np}$ is shown using dashed line.

7.2 Case 2: Linear bridge with a curved edge

The spectrogram corresponding to the case 2 is shown in Fig. 6b. As can be seen the result in this case is similar to the case 1 with the exception of the 30 ms longer nonperiodic vibration regime.

7.3 Case 3: Bridge with a geometric imperfection

Figure 6c shows the spectrogram for the case 3. Now the nonperiodic vibration regime is 300 ms long, which is al-

most two times longer compared to the case 2. Again, the energy transfer from lower to higher modes is visible during the nonperiodic vibration regime.

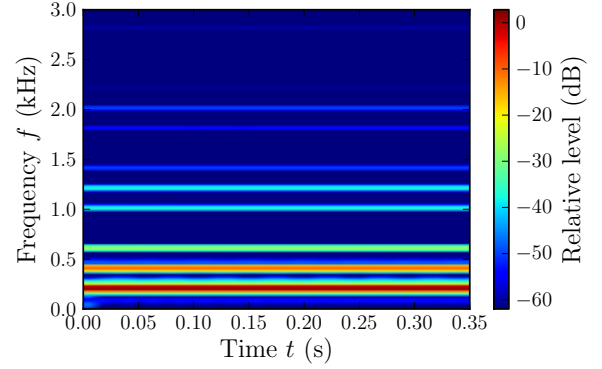


Figure 7. Spectrogram of the signal $u(l, t)$ for the linear case (no TC).

Relatively long nonperiodic vibration regime in connection with the properties of nonlinear dynamic systems can make playing such an instrument challenging. The timbre of the instrument can be very strongly influenced by the selection of the plucking point and the plucking manner, which results in uneven timbre behaviour. This effect is present for example in the sitar, and it makes the learning to play the sitar more complicated compared to the similar Western instruments.

Figure 8 shows four periods of the string deflection $u(l, t)$ during the periodic vibration regime, where the interaction of the string with the bridge is minimal. Figure 8 presents a comparison of all nonlinear cases 1 – 3 to the corresponding linear case. Nonlinear cases are rendered almost identical for $t > t_{np}$. This result is explained by the fact that the selected contact condition profiles defined by (4) - (6) have linear sections near to the string termination point ($x = 0$). With the progression of time this linear section of the bridge *trims* the effects of the other (non-linear) sections of the geometry, thus eventually rendering the periodic string vibrations for all nonlinear cases almost identical.

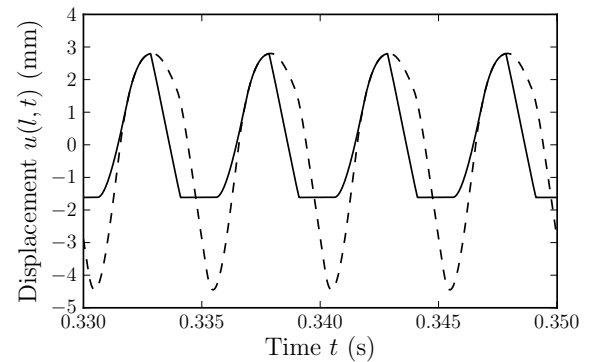


Figure 8. Four periods of the string displacement $u(l, t)$ for the nonlinear cases 1, 2 and 3 are shown using solid line (all identical). Corresponding linear case (no TC) is shown using dashed line.

In addition to the aforementioned results it was noted that a small glide and shift of the fundamental frequency f_0 (and consequently the frequencies of all the other modes, because $f_i = i f_0$ where i is the mode number) of the otherwise harmonic vibration is present. This effect appears for all presented cases and only during the nonperiodic vibration regime after which the frequency slides back to normal (i.e. $f_0 = 200$ Hz). Emergence of this effect is explained by the effective shortening of the speaking length of the string due to the spatial extent of the bridge and the interaction of the string with the bridge.

7.4 String vibration spectrum in the periodic vibration regime

After the period of nonperiodic vibration regime has passed, the string enters the periodic vibration regime. The spectrum of the string vibrations for any time instant of interest is computed using Fourier analysis. If

$$u(x, t) = \sum_i (A_i \cos \omega_i t + B_i \sin \omega_i t) \sin \left(\frac{i\pi x}{L} \right), \quad (14)$$

with normal-mode angular frequencies $\omega_i = i\omega_0$, where $\omega_0 = 2\pi f_0$ and i is the mode number, then

$$A_i = \frac{2}{L} \int_0^L u(x, t) \sin \left(\frac{i\pi x}{L} \right) dx, \quad (15)$$

$$B_i = \frac{2}{L\omega_i} \int_0^L v(x, t) \sin \left(\frac{i\pi x}{L} \right) dx, \quad (16)$$

where $v(x, t)$ is the velocity of the string. It follows that the string mode energy E_i of the i th mode is defined by

$$E_i = \frac{M\omega_i^2}{4} (A_i^2 + B_i^2), \quad (17)$$

where $M = \mu L$ is the total mass of the string. And the mode energy level is defined as

$$EL_i = 10 \log_{10} \left(\frac{E_i}{E_0} \right). \quad (18)$$

Fourier analysis using (18) shows that the spectra of cases 1, 2 and 3 are almost identical for $t > t_{np}$ (cf. Fig. 8). As mentioned earlier this result is explained by the fact that the selected contact condition profiles defined by (4) - (6) have linear sections near to the string termination point ($x = 0$). Figure 9 shows the comparison of the spectrum of the linear case (no TC) with those of the nonlinear cases 1, 2 and 3. The spectrum of the linear case is shown for the time interval $t = (t_0, \infty)$ and the nonlinear cases are shown for the time interval $t = (t_{np}, \infty)$.

Results from spectrogram analysis shown in Fig. 6 are confirmed here by calculations made using (18) and the resulting spectrum is shown in Fig. 9.

Widening of the spectra compared to the linear case and the transfer of energy from lower to higher vibration modes is visible. Relative level of some higher modes grow up to 25 dB. This means that resulting tone of the musical instrument that is equipped with the rigid, slightly curved bridge

which influences the string vibration is completely different from that of an instrument having a regular bridge.

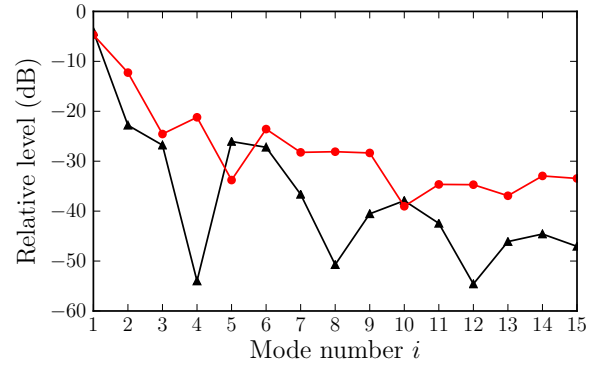


Figure 9. Stationary spectrum of the string vibrations for $t > t_{np}$. Spectra corresponding to the nonlinear cases 1, 2 and 3 are shown using circles (all identical). Linear case (no TC) is shown using triangles.

8. CONCLUSIONS

This article introduced a model that simulates the vibration of an ideal string terminated against a bridge with an arbitrary geometry. Additionally, a method for modeling the effect of minute geometric imperfections of the bridge geometry on the string vibration was presented. It was shown that the lossless string vibrates in two distinct vibration regimes. In the beginning the string starts to interact in a nonlinear fashion with the bridge, and the resulting string motion is nonperiodic. After some time of nonperiodic vibration, the string vibration settles in a *almost* periodic regime, where the dynamic motion of the string is repetitious in time.

The duration of the nonperiodic vibration regime depends on the geometry of the termination. It was concluded that minor imperfection of the bridge profile geometry elongate the duration of the nonperiodic vibration regime and produce noticeable changes in the evolution of the timber in the nonperiodic regime of vibration. The resulting spectrum in the periodic regime is identical for all nonlinear cases studied. Comparison of the resulting spectra in the periodic vibration regime of the linear and nonlinear cases showed that the interaction of the string with the rigid bridge widens the spectrum by transferring energy from lower to higher vibration modes.

Acknowledgments

This work was conducted in February - April 2013, when Dmitri Kartofelev was visiting Aalto University. His visit was supported by the Kristjan Jaak Scholarship program from the Archimedes Foundation, Estonia.

This research was partly supported by the European Union through the European Regional Development Fund and by the Estonian Ministry of Education and Research (SF 0140 077s08).

The work of Heidi-Maria Lehtonen was supported by the Finnish Cultural Foundation.

The authors are grateful for the useful comments and suggestions from Dr. Balázs Bank.

9. REFERENCES

- [1] Wikipedia the free encyclopedia: “Plucked string instrument,” http://en.wikipedia.org/wiki/Plucked_string_instrument, 2013
- [2] A. Stulov and D. Kartofelev, “Vibration of strings with nonlinear supports,” in press, *Applied Acoustics*, 2013.
- [3] C. P. Vyasarayani, S. Birkett and J. McPhee, “Modeling the dynamics of a vibrating string with a finite distributed unilateral constraint: Application to the sitar,” *J. Acoust. Soc. Am.*, vol. 125, no. 6, pp. 3673–3682, 2009.
- [4] C. V. Raman, “On some Indian stringed instruments,” in *Proc. Indian Assoc. Cultiv. Sci.* vol. 7, pp. 29–33, 1921.
- [5] A. Krishnaswamy and J. O. Smith, “Methods for simulating string collisions with rigid spatial obstacles,” in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, October 19–22 New Paltz, NY, USA, pp. 233–236, 2003.
- [6] E. Rank and G. Kubin, “A waveguide model for slap-bass synthesis,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 21–24, Munich, Germany, pp. 443–446, 1997.
- [7] T. Taguti, “Numerical simulation of vibrating string subject to sawari mechanism,” in *Proc. 19th Int. Congr. on Acoustics*, September 2–7, Madrid, Spain, pp. 1–6, 2007.
- [8] T. Taguti, “Dynamics of simple string subject to unilateral constraint: A model analysis of sawari mechanism,” *Acoustical Science and Technology*, vol. 29, pp. 203–214, 2008.
- [9] R. E. L. Turner and P. Villaggio, “A note on the motion of a string on a unilaterally reacting foundation,” *SIAM Journal on Applied Mathematics*, vol. 49, no. 4, pp. 1223–1230, 1989.
- [10] S. Siddiq, “A physical model of the nonlinear sitar string,” *Archives of Acoustics*, vol. 37, no. 1, pp. 73–79, 2012.
- [11] R. Burridge, J. Kappraff and C. Morshedi, “The sitar string, a vibrating string with a one-sided inelastic constraint,” *SIAM Journal on Applied Mathematics*, vol. 42, no. 6, pp. 1231–1251, 1982.
- [12] S. M. Han and M. A. Groesenbaugh, “Non-linear free vibration of a cable against a straight obstacle,” *Journal of Sound and Vibration*, vol. 273, pp. 337–361, 2004.
- [13] H. Cabannes, “Presentation of software for movies of vibrating strings with obstacles,” *Appl. Math. Lett.*, vol. 10, no. 5, pp. 79–84, 1997.
- [14] C. Valette, “The mechanics of vibrating strings,” in *Mechanics of Musical Instruments*, edited by J. Kergomard and G. Weinreich, Springer-Verlag, Vienna, pp. 115–183, 1995.
- [15] G. Evangelista and F. Eckerholm, “Player–instrument interaction models for digital waveguide synthesis of guitar: touch and collisions,” *IEEE Transactions on audio, speech, and language processing*, vol. 18, no. 4, pp. 822–832, 2010.
- [16] N. H. Fletcher and T. D. Rossing, *The physics of musical instruments*, Springer, 1998.
- [17] J. O. Smith, “Physical Modeling using Digital Waveguides,” *Computer Music Journal*, vol. 16, no. 4, pp. 74–91, 1992.
- [18] V. Välimäki, J. Pakarinen, C. Erkut, and M. Karjalainen, “Discrete-time modelling of musical instruments,” *Reports on Progress in Physics*, vol. 69, no. 1, pp. 1–78, 2006.
- [19] J. O. Smith, *Physical audio signal processing for virtual musical instruments and audio effects*, <https://ccrma.stanford.edu/~jos/pasp/>, W3K Publishing, 2010.
- [20] Supplementary web page of the current article: “Animations of the string vibration terminated against a bridges with arbitrary geometry’s: Cases 1 – 3,” <http://www.cs.ioc.ee/~dima/stringarbitrary.html>
- [21] H.-M. Lehtonen, A. Askenfelt and V. Välimäki, “Analysis of the part-pedaling effect in the piano,” *J. Acoust. Soc. Am.* vol. 126, no. 2, pp. EL49–EL54, 2009.
- [22] A. Stulov, V. Välimäki and H.-M. Lehtonen, “Modeling of the part-pedaling effect in the piano,” in: *Acoustics 2012*, April 23–27, Nantes, France, pp. 1223–1228, 2012.

Coupled Modes and Time-Domain Simulations of a Twelve-String Guitar with a Movable Bridge

Miguel Marques $\star\diamond$, José Antunes \star , Vincent Debut \star

\star Applied Dynamics Laboratory, Campus Tecnológico e Nuclear, Instituto Superior Técnico
Universidade Técnica de Lisboa, Estrada Nacional 10, 2695-066 Bobadela, Portugal
and

\diamond Mechanical Engineering Department, Instituto Superior Técnico
Universidade Técnica de Lisboa, Av Rovisco Pais 1, 1049-001 Lisboa, Portugal

miguel.e.marques@gmail.com, jantunes@ctn.ist.utl.pt, vincentdebut@ctn.ist.utl.pt

ABSTRACT

Coupling between the different vibrating sub-systems of a musical instrument is an important feature in music acoustics. It is the reason why instruments of similar families have such different and characteristic sounds. In this work, we propose a model for a twelve strings (six pairs) guitar, such that the strings are coupled with the instrument body through the moving bridge, which is the relevant component for energy transmission from the strings to the guitar body and back. In this preliminary study, the guitar body is modelled as a simple plate, strings being assumed to display planar-only vertical motions. However, the coupled equations thus obtained can be readily extended to cope with real guitar body modes and orbital string motions. After obtaining the coupled modes of the instrument, we illustrate the instrument time-domain coupled dynamics, by considering the characteristic modal frequencies typical of a Portuguese guitar. In particular we show how, when only one string alone is plucked, energy is transmitted to all other strings, causing sympathetic vibrations, which contribute to give this guitar its own characteristic sound identity.

1. INTRODUCTION

The acoustic guitar has been the subject of many studies, both on its theoretical description and on the experimental side. When modelling real guitars, the bridge role should not be simplified as a simple enforced pinned boundary condition. Realistic instrument models impose the need for incorporating the body dynamics, as well as the coupling between strings and the body at the moving bridge [4]. Furthermore, through adequate modelling, we believe that the bridge position may be optimized in order to achieve the most satisfying radiative acoustic response.

Motivated by the wish to extend our understanding of the acoustics of the Portuguese guitar, we set ourselves to

model the dynamics of an archtop twelve strings acoustic guitar. For such purpose, we built a conceptual model for a set of twelve (six pairs of) strings coupling, at some point of the strings (where the bridge would be located), to a body, through the bridge, which will be modelled as a spring with a very large stiffness. This work is intended to be mostly of methodological nature. In this paper, the body will be simplified as a thin wood plate. Other coupling techniques have already been proposed in the literature of musical acoustics (e.g. [1–3]). Specifically for guitars, different models, which do not take into account the vibrations of the *dead side* of the string, have already been proposed [4–7], but to our knowledge, the sympathetic vibrations of a guitar with a movable bridge, displaying a *dead side* of the string after the bridge, has never been addressed nor modelled. A model of coupling between one string and the violin body, which has the same features of our model, has been proposed by some of the present authors, in order to address the problem of the *wolf note* and the nonlinear behaviour of the string/bow interaction [8].

The Portuguese guitar (*Cithara lusitanica*) is a pear-shaped instrument with twelve metal strings (six courses), descendant from the renaissance european cittern. This instrument is widely used in Portuguese traditional music, mainly in Fado, and more recently also started to play a considerable role among urban Portuguese musicians. Unlike most common guitars, this guitar has a bent soundboard (arched top) with a bridge somewhat similar, although smaller in size, to the bridge of a violin, a neck typically with 22 fixed metal frets and it is tuned by a fan-shaped tuning mechanism, consisting in twelve screws, acting as pegs, mounted with small gliding pins where the strings are attached to adjust its tension. It has the typical tuning of the European cittern tradition, and has kept an old plucking technique, described in sixteenth century music books. The first courses are composed by plain steel strings and tuned in unison, and the remaining are combinations of a plain steel string and an overspun copper on steel string tuned one octave bellow. There are two different models of the Portuguese guitar: the Lisbon guitar and the Coimbra guitar, named after two Portuguese cities where the two most important Fado styles emerged. They differ in some de-

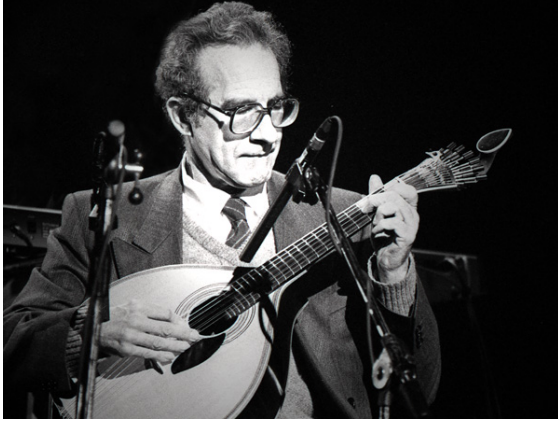


Figure 1. The master musician Carlos Paredes, playing the Coimbra Portuguese Guitar. Photo credits: Egidio Santos.

tails, such as the body measurements, the string length (the Coimbra guitar has a larger arm), and the Lisbon guitar is tuned (each string) a whole tone above the Coimbra guitar. Strings used in these two guitars are therefore different in size and linear mass. More about the Portuguese guitar can be read in [9, 10] and about experimental vibrational measurements of this instrument in [11]. In figure 1, a Coimbra model of the Portuguese guitar is shown as an example.

2. GUITAR MODEL

Throughout this section, we will give a full description of the computational model proposed in this paper. For simplicity, in this conceptual model, we will replace the guitar body by a thin, isotropic, rectangular soundboard plate, fixed at the ends. For simplicity, we will neglect any energy transfer occurring through the nut or the fingerboard, and through the tailpiece; this aspects have been addressed by several authors (e.g. [4, 8]). Therefore, our fully coupled system will consist in a rectangular plate with fixed ends, and twelve strings, attached to the plate trough a set of springs, displayed in six pairs. For simplicity, we will only consider string vibrations in the direction normal to the soundboard plate, although it would be relatively straightforward to implement the other direction as well. We will decompose the strings and the plate displacements respectively as

$$\begin{aligned} Y_s(x_s, t) &= \sum_n \psi_n^s(x_s) q_n^s(t), \\ Z_b(x_b, y_b, t) &= \sum_{j,l} \psi_{j,l}^b(x_b, y_b) q_{j,l}^b(t), \end{aligned} \quad (1)$$

with the script $s = 1, \dots, S$, for a total of S strings; where Y_s is the displacement for each string, Z_b is the displacement of the body, ψ stand for the modeshapes and q stand for the modal amplitudes. In this modal formulation, the plate dynamics is governed by

$$m_{jl} \ddot{q}_{jl}^b(t) + 2m_{jl} \omega_{jl} \zeta_{jl} \dot{q}_{jl}^b(t) + m_{jl} \omega_{jl}^2 q_{jl}^b(t) = F_{jl}^b(t), \quad (2)$$

where $j = 1, \dots, J$ is related to the number of nodal lines along the xx direction, $l = 1, \dots, L$ is the number of nodal

lines along the yy direction, m_{jl} are the modal masses, ω_{jl} are the modal angular frequencies, ζ_{jl} are the modal damping coefficients, and F_{jl}^b are the generalized forces on the plate. To simplify the formulation, these equations can be organised in order to have a single modal index g , with a total of $G = J \times L$ modes, by ordering the frequencies in an ascendent form. The modal masses of the body are $m_g = \rho \frac{L_x L_y}{4}$, with ρ being the plate surface density, and L_x, L_y the plate dimensions. Each string will have its own set of modal dynamical equations, given as

$$m_n \ddot{q}_n^s(t) + 2m_n \omega_n \zeta_n \dot{q}_n^s(t) + m_n \omega_n^2 q_n^s(t) = F_n^s(t), \quad (3)$$

where $n = 1, \dots, N_s$ are the modes of each string s . The modal masses of each string are $m_n = \mu \frac{L_s}{2}$, with μ being the linear density of the string, and L_s the size of the string (throughout this work, we will assume that all strings have the same size, which is the case for the Portuguese guitar). Furthermore, the string and the body will have modeshapes respectively given by

$$\begin{aligned} \psi_n^s(x_s) &= \sin\left(\frac{n\pi x_s}{L_s}\right), \\ \psi_{jl}^b(x_b, y_b) &= \sin\left(\frac{j\pi x_b}{L_x}\right) \sin\left(\frac{l\pi y_b}{L_y}\right), \end{aligned} \quad (4)$$

and eigenfrequencies (assuming no inharmonicity) respectively given by

$$\begin{aligned} \omega_n^s &= n\omega_0^s, \\ \omega_{jl}^b &= h_b \sqrt{\frac{E_y}{12\rho(1-v_p^2)}} \left[\left(\frac{j\pi}{L_x}\right)^2 + \left(\frac{l\pi}{L_y}\right)^2 \right], \end{aligned} \quad (5)$$

where ω_0^s is the s string fundamental frequency, h_b is the plate height, E_y is the Young bulk modulus, and v_p is the Poisson coefficient. Throughout the paper, we will refer to ω_{jl}^b as ω_g^b , and to $\psi_{jl}^b(x_b, y_b)$ as $\psi_g^b(\mathbf{r})$, assuming a proper indexation mapping $g \rightarrow (j, l)$, according to what has been previously stated. Given a general force $F(\mathbf{x}, \mathbf{t})$, acting on a surface S of a vibrating system with modeshapes $\psi_m(\mathbf{x})$, the corresponding forces on the modal space will be

$$F_m(\mathbf{t}) = \int_S F(\mathbf{x}, \mathbf{t}) \psi_m(\mathbf{x}) dS. \quad (6)$$

Let us define $F_c^s(x_c, t)$ and $F_c^b(\mathbf{r}_c^s, t)$ as being the force exerted on the bridge, as seen, respectively, from each string s and from the body at each coupling point \mathbf{r}_c^s (we are assuming that all strings will meet the bridge at the same distance from the nut, x_c). We will consider that the bridge his thin enough such that it can be modelled as a spring with very large stiffness. Therefore, these forces are

$$\begin{aligned} F_c^s(x_c, t) &= K_c [\delta(x_s - x_c) Y_s(x_c, t) - \delta(\mathbf{r} - \mathbf{r}_c^s) Z_b(\mathbf{r}_c^s, t)] \\ &\quad + C_c [\delta(x_s - x_c) \dot{Y}_s(x_c, t) - \delta(\mathbf{r} - \mathbf{r}_c^s) \dot{Z}_b(\mathbf{r}_c^s, t)] \\ &= \delta(x_s - x_c) \left\{ \sum_{m=1}^{N_s} \psi_m^s(x_c) [K_c q_m^s(t) + C_c \dot{q}_m^s(t)] \right\} \\ &\quad - \delta(\mathbf{r} - \mathbf{r}_c^s) \left\{ \sum_{m=1}^G \psi_m^b(\mathbf{r}_c^s) [K_c q_m^b(t) + C_c \dot{q}_m^b(t)] \right\} \end{aligned} \quad (7)$$

where K_c , C_c are, respectively, the (very large) stiffness constant and the damping constant of the bridge (C_c play a role forcing the dissipation of high frequency oscillations at the bridge constraint, because these are of numerical nature), and

$$F_c^b(\mathbf{r}_c^s, t) = -F_c^s(x_c, t), \forall s \in [1, S] \quad (8)$$

The modal projections forces (7) are given by

$$\begin{aligned} F_n^s(t) &= \sum_{m=1}^N [K_c q_m^s(t) + C_c \dot{q}_m^s(t)] \psi_m^s(x_c) \psi_n^s(x_c) \\ &\quad - \sum_{m=1}^G [K_c q_m^b(t) + C_c \dot{q}_m^b(t)] \psi_m^b(\mathbf{r}_c^s) \psi_n^s(x_c), \\ F_g^b(t) &= \sum_{s=1}^S \sum_{m=1}^G [K_c q_m^b(t) + C_c \dot{q}_m^b(t)] \psi_m^b(\mathbf{r}_c^s) \psi_g^b(\mathbf{r}_c^s) \\ &\quad - \sum_{s=1}^S \sum_{m=1}^{N_s} [K_c q_m^s(t) + C_c \dot{q}_m^s(t)] \psi_m^s(x_c) \psi_g^b(\mathbf{r}_c^s). \end{aligned} \quad (9)$$

We can rewrite these equations as the linear system of ODEs

$$[\mathcal{M}]\{\ddot{\mathcal{Q}}(t)\} + [\mathcal{C}]\{\dot{\mathcal{Q}}(t)\} + [\mathcal{K}]\{\mathcal{Q}(t)\} = \{\mathcal{F}(t)\}, \quad (10)$$

with the matrices and vectors built from the model parameters from all strings and the body

$$\begin{aligned} [\mathcal{M}] &= \text{diag}(m_1^s, \dots, m_N^s, m_1^b, \dots, m_G^b), \\ [\mathcal{C}] &= 2 \cdot \text{diag}(m_1^s \omega_1^s \zeta_1^s, \dots, m_N^s \omega_N^s \zeta_N^s, m_1^b \omega_1^b \zeta_1^b, \dots, m_G^b \omega_G^b \zeta_G^b), \\ [\mathcal{K}] &= \text{diag}(m_1^s \omega_1^{s^2}, \dots, m_N^s \omega_N^{s^2}, m_1^b \omega_1^{b^2}, \dots, m_G^b \omega_G^{b^2}), \\ \{\mathcal{Q}(t)\} &= \{q_1^s(t), \dots, q_N^s(t), q_1^b(t), \dots, q_G^b(t)\}^T, \\ \{\mathcal{F}(t)\} &= \{F_1^s(t), \dots, F_N^s(t), F_1^b(t), \dots, F_G^b(t)\}^T. \end{aligned} \quad (11)$$

Notice that, each string will have its own, independent, modal family N_s , and therefore, the dimension of the matrices and vectors in (12) will be $D = G + \sum_s N_s$.

2.1 Frequency Domain Analysis

Considering the eigensolutions $q_k(t) = \bar{q}_k e^{\bar{\lambda}_k t}$, we can transform the system (10) in an eigenvalue problem, becoming

$$\left(\begin{bmatrix} [0] & [\mathbf{I}] \\ -[\mathcal{M}]^{-1}[\bar{K}] & -[\mathcal{M}]^{-1}[\bar{C}] \end{bmatrix} - \bar{\lambda}_k [\mathbf{I}] \right) \begin{Bmatrix} \bar{q}_k \\ \bar{\lambda}_k \bar{q}_k \end{Bmatrix} = \{0\}, \quad (12)$$

where $[\bar{K}]$ and $[\bar{C}]$ are the effective stiffness and damping matrices, given by $[\bar{K}] = [\mathcal{K}] + K_c[\Phi_c]$ and $[\bar{C}] = [\mathcal{C}] + C_c[\Phi_c]$, and $[\Phi_c]$ is the coupling matrix,

$$[\Phi_c] = \begin{bmatrix} [\Phi_c^{s1}] & [0] & [0] & \dots & [\Phi_c^{s1-b}] \\ [0] & [\Phi_c^{s2}] & [0] & \dots & [\Phi_c^{s2-b}] \\ [0] & [0] & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ [\Phi_c^{b-s1}] & [\Phi_c^{b-s2}] & \dots & \dots & [\Phi_c^b] \end{bmatrix}, \quad (13)$$

given the matrices

$$\begin{aligned} [\Phi_c^s] &= \begin{Bmatrix} \psi_1^s(x_c) \\ \vdots \\ \psi_n^s(x_c) \end{Bmatrix} \cdot \{ \psi_1^s(x_c) \dots \psi_n^s(x_c) \}, \\ [\Phi_c^{s-b}] &= - \begin{Bmatrix} \psi_1^s(x_c) \\ \vdots \\ \psi_n^s(x_c) \end{Bmatrix} \cdot \{ \psi_1^b(\mathbf{r}_c^s) \dots \psi_g^b(\mathbf{r}_c^s) \}, \\ [\Phi_c^{b-s}] &= - \begin{Bmatrix} \psi_1^b(\mathbf{r}_c^s) \\ \vdots \\ \psi_g^b(\mathbf{r}_c^s) \end{Bmatrix} \cdot \{ \psi_1^s(x_c) \dots \psi_n^s(x_c) \}, \\ [\Phi_c^b] &= \sum_{s=1}^S \begin{Bmatrix} \psi_1^b(\mathbf{r}_c^s) \\ \vdots \\ \psi_g^b(\mathbf{r}_c^s) \end{Bmatrix} \cdot \{ \psi_1^b(\mathbf{r}_c^s) \dots \psi_g^b(\mathbf{r}_c^s) \}. \end{aligned} \quad (14)$$

The eigenvalues of (12) are related to the modal frequencies and modal damping coefficients as

$$\bar{\lambda}_k = -\bar{\omega}_k \bar{\zeta}_k \pm i \bar{\omega}_k \sqrt{1 - \bar{\zeta}_k^2}. \quad (15)$$

From this relationship, we can deduce the guitar modal damped frequencies, modal damping coefficients and mode-shapes to be

$$\bar{\omega}_d k = \Im(\bar{\lambda}_k), \quad (16)$$

$$\bar{\zeta}_k = -\frac{\Re(\bar{\lambda}_k)}{|\bar{\lambda}_k|}, \quad (17)$$

and

$$\begin{aligned} \bar{\psi}_k(\mathbf{x}) &= \sum_{s=1}^S \sum_{m=1}^{N_s} \bar{q}_m^k \Theta((s-1)L_s < \mathbf{x} < sL_s) \psi_m^s(x) \\ &\quad + \sum_{m=1}^G \bar{q}_m^k \Theta(\mathbf{x} > SL_s) \psi_m^b(\mathbf{r}), \end{aligned} \quad (18)$$

where $\Theta(\alpha)$ is an Heaviside-like step function, such that it is equal to one when the argument α is true, and zero otherwise (notice that no subsystems are being summed), \bar{q}_m^k is the m^{th} term of the k^{th} eigenvector, and \mathbf{x} is defined as

$$\mathbf{x} = \begin{cases} x_s & \text{if } \Theta((s-1)L_s < \mathbf{x} < sL_s) = 1 \\ \mathbf{r} & \text{if } \Theta(\mathbf{x} > SL_s) = 1 \end{cases} \quad (19)$$

In short, the coupled system eigenfunctions (18) correspond to the modes of the S coupled strings and the modes of the coupled soundboard, while (19) corresponds to the coordinate of each string when dealing with the coupled strings modes, and corresponds to the coordinates of the soundboard when dealing with the coupled soundboard modes. As for the original decoupled modes, we normalized (18) such that the maximum amplitude of its absolute value is one.

2.2 Time Domain Analysis

2.2.1 Initial Conditions

We will assume that the initially, all strings are static, having zero initial displacements as well as zero initial velocities, and there will rather be an initial excitation force acting on one or more strings (*i.e.* a finger or a nail plucking the string, or a pick striking the string) during a short initial time interval. We will model the excitation force in such a way that its modal projection (over the excited string modal space) will be

$$F_n^f(t) = \left\{ K_f [Z_f(t) - Y_s(x_f, t)] + C_f [\dot{Z}_f(t) - \dot{Y}_s(x_d, t)] \right\} \psi_n^s(x_f), \quad (20)$$

where K_f , C_f are respectively the finger/nail/pick stiffness constant (here taken to be very large) and damping constant, $Z_f(t)$ is the finger/nail/pick displacement, and $Y_s(x_f, t)$ is the string displacement at the point where it is plucked/struck. This displacements are, respectively,

$$Z_f(t) = \dot{Z}_f t = \frac{Z_{max}}{t_{max}} t|_{t_{min}}^{t_{max}}, \quad (21)$$

and

$$Y_s(x_f, t) = \sum_{m=1}^{N_s} q_m^s(t) \psi_m^s(x_f). \quad (22)$$

This choice allow us to easily perform simulations in which different strings are plucked/struck at different times.

2.2.2 Time-step integrating procedure

We start by reducing the ODEs system (10) to a system of first order ODEs, based on the unconstrained subsystem modal responses

$$\begin{Bmatrix} \dot{Q}(t) \\ \ddot{Q}(t) \end{Bmatrix} = \begin{bmatrix} [0] & [\mathbf{I}] \\ -[\mathcal{M}]^{-1}[\bar{K}] & -[\mathcal{M}]^{-1}[\bar{C}] \end{bmatrix} \begin{Bmatrix} Q(t) \\ \dot{Q}(t) \end{Bmatrix} + \begin{Bmatrix} 0 \\ \mathcal{F}_f(t) \end{Bmatrix}, \quad (23)$$

where $\mathcal{F}_f(t)$ is a vector which includes all the terms of $F_n^f(t)$ in the correspondent excited modes, and is zero otherwise. We will refer to (23) as

$$\dot{p}(t) = \mathbf{A} p(t) + F_e(t). \quad (24)$$

The analytical solution of (23) is

$$p(t) = p(t_0)e^{\mathbf{A}(t-t_0)} + e^{\mathbf{A}t} \int_{t_0}^t e^{-\mathbf{A}\tau} F_e(\tau) d\tau. \quad (25)$$

Assuming a very short time step $\Delta t = t - t_0 \rightarrow 0$, it becomes possible to approximate $F_e(t)$ as being constant during each Δt . We can then discretize equation (25), obtaining the numerical solution

$$p^{t_{i+1}} = p^{t_i} e^{\mathbf{A}\Delta t} + \mathbf{A}^{-1} (e^{\mathbf{A}\Delta t} - \mathbf{I}) F_e^{t_i}, \quad (26)$$

where \mathbf{I} is the identity matrix. Notice that at each step t_i , one must recompute the vector $F_e^{t_{i-1}}$, using the results $p^{t_{i-1}}$, according to what as previously stated in (20). This method is quite stable (for the considered external forces

$F_e(t)$) if given an accurate calculation of the state transition matrix $e^{\mathbf{A}\Delta t}$. As to accuracy, this method provide highly accurate results up to the time step bound $\Delta t \leq T_{min}/10$, where T_{min} is the smallest period of the system (the period of the highest considered mode).

3. RESULTS

Throughout this section, we will consider simulations using the typical values of a Lisbon Portuguese guitar. All strings have a length of 44 cm from the nut to the bridge, and 17.5 cm from the bridge to the stop tailpiece (total length is 61.5 cm). Table 1 contains the strings notes, the corresponding frequencies according to the standard tuning of the Lisbon Portuguese guitar, and the corresponding linear masses. Notice that the frequencies shown in table 1 represent the fundamental frequency at which the active part of the string, *i.e.* the length between the nut and the bridge, should be vibrating; the actual frequency at which the full-length string is tuned will be given by multiplying the presented value by the factor $\frac{44}{44+17.5}$. Based on the average results obtained in experimental identifications, all strings will have equal damping coefficients $\zeta_n^s = 0.05\% \forall n \in N_s$. Each string s will have a total number of degrees of freedom N_s such that the frequency of the highest mode will be $f_{N_s} \sim 10kHz$. As for the body, we

strings pair notes	1 st	2 nd	3 rd	4 th	5 th	6 th
frequency (Hz)	493.88 493.88	440 440	329.63 329.63	493.88 246.94	440 220	293.66 146.83
linear density (10 ⁻⁴ kg/m)	3.78	3.94	6.20	3.78	3.94	11.30
number of modes used	28	32	42	28	32	27
	28	32	42	57	64	95

Table 1. Notes and corresponding frequencies of the Lisbon Portuguese guitar standard tuning.

will assume the soundboard plate to be squared with 30 cm by side, surface density $\rho = 0.5 \text{ kg/m}^2$, $\zeta_g^b = 1\% \forall g \in G$, and we adjust the parameters E_y , v_p and h_b such that the first body frequency will be $f_1^b = 275 \text{ Hz}$ (which is one of the values measured in [11] for a Portuguese guitar body); also, we have considered 36 modes, such that the maximum body frequency is $\sim 10kHz$. The total number of modes considered for the system is therefore 543; with this choice, we have verified that the frequencies of the coupled system are convergent in a range from 0Hz to $\sim 8kHz$. The strings of each pair will be separated by a distance of 4mm, and each pair will be separated by a distance of 8mm (these are the typical values chosen by portuguese guitar luthiers). The stiffness and damping constants of the coupling spring, and of the finger exciting force will respectively be $K_c = K_f = 10^7$, and $C_c = C_f = 5$.

3.1 Frequency Domain Results

In Figure 2, we show the modeshape of the first fully coupled system mode, which is dominated by the 11th string, the lowest tuned frequency of the system. Except for the

11th string, all other strings will have qualitatively equivalent vibrations to that of the 1st string in this system mode, as detailed in the figure. In Figure 3, we show the mode-

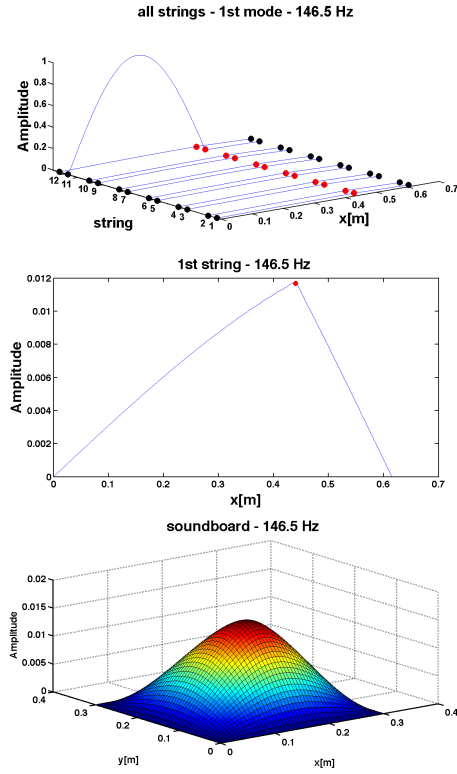


Figure 2. First mode of the coupled system. The red dots represent the bridge position. Upper plot, string modal responses; medium plot, detail of the 1st string modal response; lower plot, detail of the body modal response.

shape of the fifteenth system mode, which is dominated by the 1st string. In this mode, we observe significant vibrations in all strings tuned to *b4* or *b3*. We find that the first pair of strings has a rather different phase than the *b4* string of the fourth pair, and on the other hand, the *b3* string of the fourth pair, which frequency corresponds to the second mode of the *b4* strings, share the same phase as the first pair, and has an amplitude considerably smaller than that of the *b4* strings. The eleventh string (which is the *d3* string of the sixth pair), is interestingly vibrating in its fifth mode with a very low amplitude; the fourth mode of this string has a frequency of about 525 Hz, which is almost 30 Hz higher than the fifteenth system mode. We also notice that there are small discrepancies between the frequency of the system modes and the frequency of the dominating strings. This suggests that the coupling of the different subsystems, as well as the *dead* side of the strings, adds some inharmonicity to the system, which would account for the audible differences between different musical instruments.

3.2 Time Domain Results

We performed time domain simulations of the situations in which: (i) the musician will only excite the first (*b4*) string, (ii) the musician will only excite the eleventh (*d3*) string. Although typically twelve-string guitar players would pluck

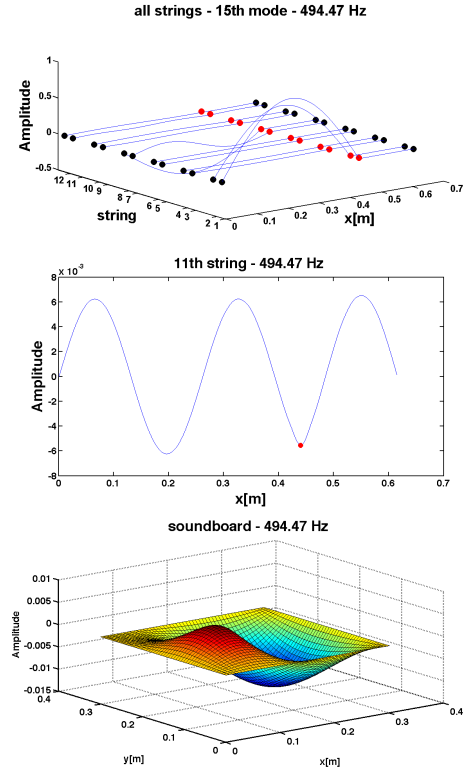


Figure 3. Fifteenth mode of the coupled system. The red dots represent the bridge position. Upper plot, string modal responses; medium plot, detail of the 11th string modal response; lower plot, detail of the body modal response.

a string pair rather than a single string, we will focus in analysing the most simple possible scenarios. The string excitation lasts for 0.01 s, and the simulation will last for 3 s; however, we find that in the first scenario, almost all energy of the system has been damped after 1.5 seconds, which is remarkably similar to what we typically hear when plucking a portuguese guitar string. In Figures 5 and 6, we show the plots of the energy evolution in each string, in the body, and the total energy of the system, respectively for the scenarios (i) and (ii). The energy is computed by the formula

$$E = \sum_{k=1}^D \left(\frac{1}{2} m_k \dot{q}_k^2(t) + \frac{1}{2} m_k \omega_k^2 q_k^2(t) \right), \quad (27)$$

for the total energy, and by equivalently summing the relevant modes of each string or the body, for the subsystems energy. In Figure 4, we show the average energy of each subsystem, relative to the average total energy of the coupled system.

As expected, in the first scenario, the first string is the subsystem which has the most significant amount of energy, but interestingly, its energy will practically vanish after the first 0.6 seconds, while the coupled system still has a fair amount of energy. Shortly after 0.2 seconds have passed, the energy of the second, seventh and eighth strings (all the strings tuned to the same fundamental frequency *b4*, and the string tuned to the second harmonic *b3*) will be comparable in magnitude with the energy of the first string,

while the total energy of the system will have a larger magnitude from this moment on. Particularly the second string, at the time in which it achieves its maximum energy (when close to 0.3 seconds), has almost the same energy as the excited string, which is not surprising given that they are tuned to the same frequency and they are very closely located at the bridge. The body displays an energy profile which reveals that there is energy being transferred back and forth between itself and the twelve strings. All the strings of the system will display some amount of energy, even if they have been tuned in frequencies rather lower than that of the excited strings. This will give this guitar its own distinguishable sound.

In the second scenario, there is not a very efficient transfer of energy from the excited string to the other subsystems (comparing with the first scenario), and the only subsystems which will receive a relevant amount of energy are the soundboard, the *d4* string (which is within the same pair, and tuned to the second harmonic of the excited string), and the strings of the "neighbour" pair (*a3*, *a4*). Based in our observations, the most efficient transfer of energy occurs when exciting strings with more than one subsystem tuned to a multiple frequency, due to the fact that this subsystems will experience sympathetic vibrations.

4. CONCLUSIONS

We have developed a conceptual model to accurately perform, in both frequency and time domain, analysis of twelve-string guitars as a fully coupled system. The formulation considered is sufficiently versatile to be also applicable to model any other plucked string instrument, regardless of the number of strings or the geometry in which they are disposed on the instrument body. In this preliminary analysis, we have found results which corroborate that the body, through the bridge, will account for a significant part of the energy transmission across the multiple subsystems. We also show the relevance of the "dead side" of the string (*i.e.* the continuation of the string after the bridge) for the wave propagation across the string. We stress that this results might be a significant contribution for works on the optimisation of guitar characteristics and radiation. Future work will include the use of real body modes, provided by experimental modal identifications, as well as more realistic string properties.

Acknowledgments

The authors warmly thank Professors Pedro Serrão, António Relógio Ribeiro and Octávio Inácio for interesting discussions on modelling issues, as well as Pedro Caldeira Cabral for very important discussions and valuable information on guitar building and performance issues. This work was supported by FCT - Fundação para Ciência e Tecnologia - Portugal under the project PTDC/FIS/103306/2008.

5. REFERENCES

[1] Weinreich, "Coupled Piano Strings", in *J. Acoust. Soc. Am* Vol. 62, 1977

[2] J. Carrou, F. Gautier, N. Dauchez, and J. Gilbert, "Modelling of Sympathetic String Vibrations", in *Acta Acoustica united With Acustica* Vol. 91, 2005, pp. 277–288.

[3] J. Carrou, F. Gautier, and R. Badeau, "Sympathetic String Modes in the Condert Harp", in *Acta Acoustica united With Acustica* Vol. 95, 2009, pp. 744–752.

[4] Woodhouse, "On the Synthesis of Guitar Plucks", in *Acta Acoustica United With Acustica* Vol. 90, 2004, pp. 928–944.

[5] G. Derveaux, A. Chaigne, P. Joly, and E. Bécache, "Time-domain simulation of a guitar: Model and method", in *J. Acoust. Soc. Am.* Vol. 114, No.6, 2003, pp. 3368–3383.

[6] E. Bécache, G. Derveaux, A. Chaigne, and P. Joly, "Numerical Simulation of a Guitar", in *Computers Structures* Vol. 83, 2005, pp. 107–126.

[7] A. Nackaerts, B. Moor, and R. Lauwereins, "Coupled string guitar models", in *Proceedings of the WSES International Conference on Acoustics and Music: Theory and Applications 2001*, Skiathos, 2001, p.119.

[8] O. Inácio, J. Antunes, and M.C.M. Wright, "Computational modelling of string-body interaction for the violin family and simulation of wolf notes", in *J. Of Sound And Vibration* Vol. 310, 2008, pp. 260–286.

[9] P. C. Cabral, *A Guitarra Portuguesa*. Ediclube, 1999.

[10] P. C. Cabral, *The History of The Guitarra Portuguesa*, (chapter in *The Lute In Europe 2*). Menziken, 2011.

[11] O. Inácio, F. Santiago, and P. C. Cabral, "The portuguese guitar acoustics: Part 1 - vibroacoustic measurements", in *Proc. of the 4th ibero-American Congress Acústica*, Guimarães, 2004, pp. 14–17.

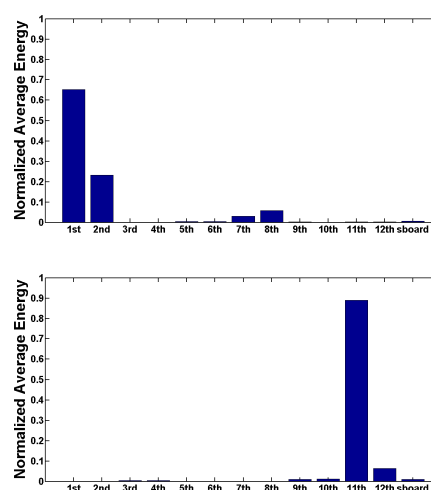


Figure 4. Average Energy of each subsystem; upper plot: excitation of the first (b4) string, lower plot: excitation of the eleventh (d3) string.

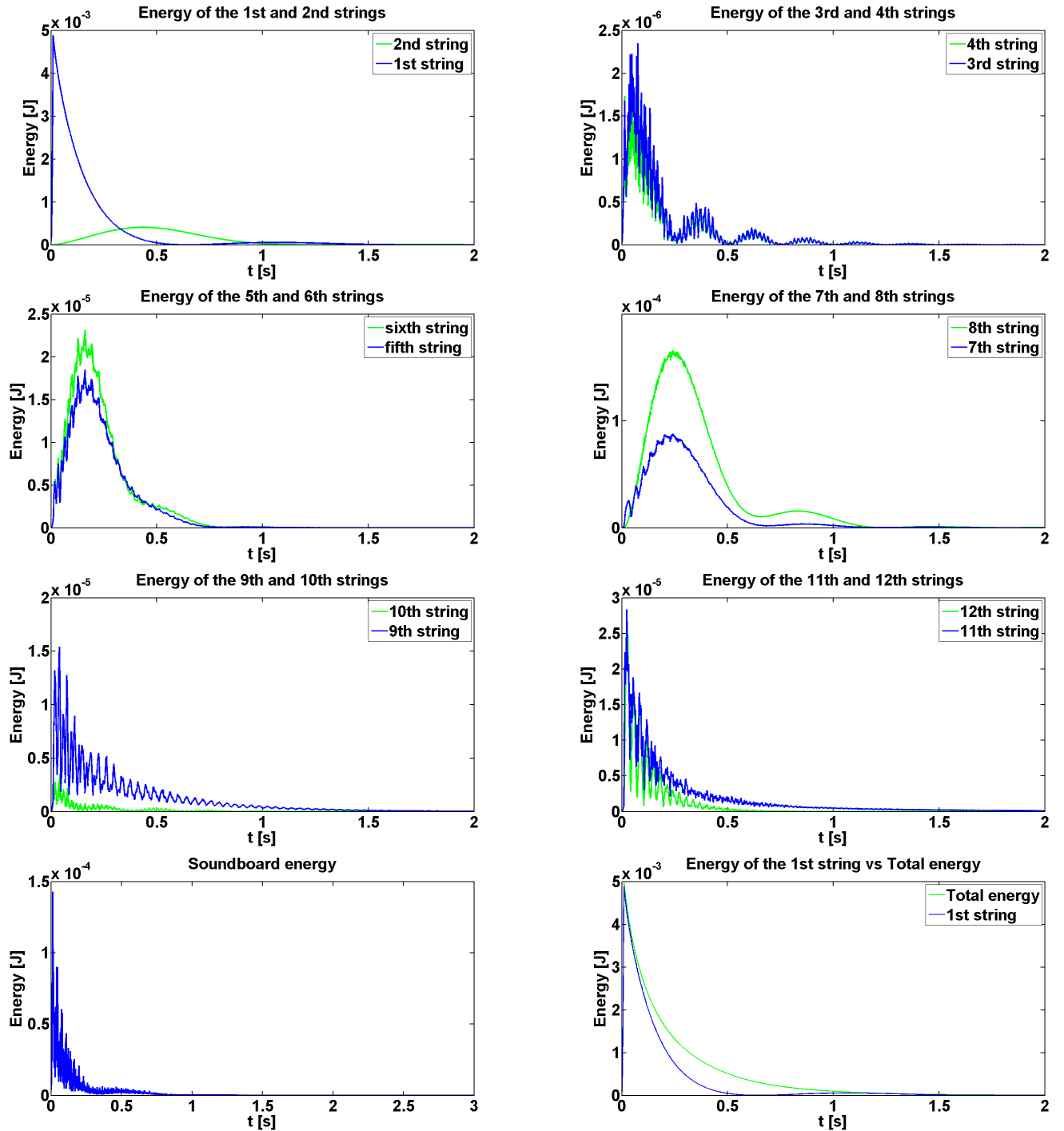


Figure 5. Energy plots for an excitation of the first (b4) string.

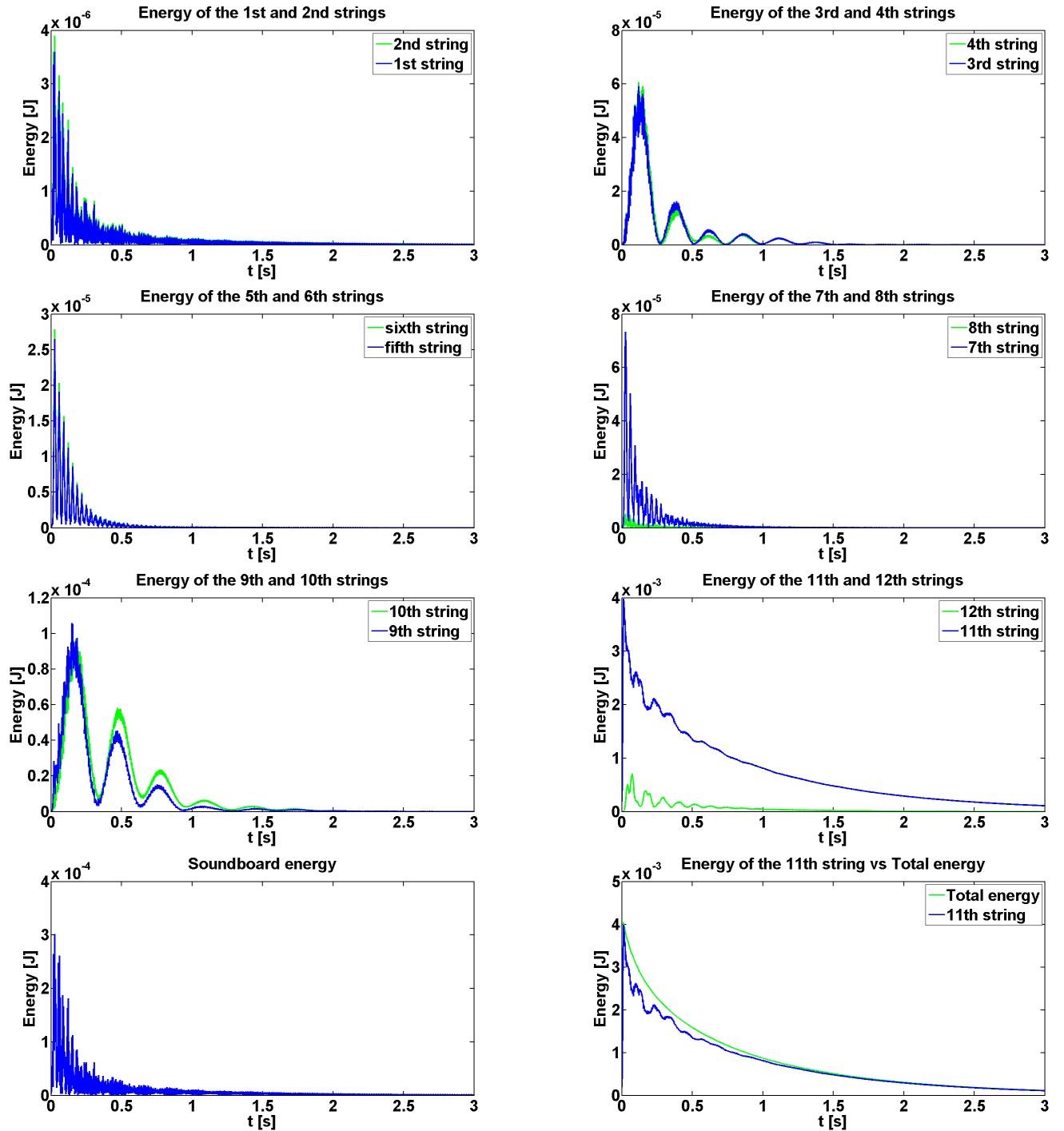


Figure 6. Energy plots for an excitation of the eleventh (d3) string.

DISTRIBUTED PIANO SOUNDBOARD MODELING WITH COMMON-POLE PARALLEL FILTERS

Stefano Zambon

Viscount International S.p.A., Mondaino (RN), Italy

s.zambon@viscount.it

ABSTRACT

The soundboard plays a major role in defining the distinctive spectral and temporal characteristics of piano tones. Within the context of physics-based sound synthesis, it is customary to model the radiation effects of the soundboard as a common post-processing block for all the notes. In this paper, a computationally efficient technique is proposed for the simultaneous computation of multiple responses, corresponding to distributed excitation positions along the bridge. The method employs an approximation of measured impulse responses with several sets of parallel second-order resonators sharing the same poles, followed by a common FIR part. Details about experimental setup, parameter estimation and computational cost are covered and sound examples are provided.

1. INTRODUCTION

Physics-based methods approach the problem of synthesizing the sounds produced by musical instruments from a *source* point of view, trying to model each of the functional parts of the acoustic instrument with a dedicated computational block. In the case of the piano [1], a common macro-architecture consists in three blocks having the roles of *exciter* (the hammer), *resonator* (the string) and *radiator* (the soundboard), connected in series as in Fig. 1.

In real pianos, the soundboard has two main roles in characterizing the sound. First, it provides a terminating impedance to all the strings, thus modifying the frequencies and decay times of the partials of each string and enabling the exchange of energy between different strings (i.e. the *sympathetic resonance* phenomenon). However, it is common in real-time physics-based synthesis algorithms to neglect the impedance properties of the soundboard as their effects can be more easily included in the string model [1, 2]. As its other important function, the soundboard acts as a radiator for the strings, amplifying the vibration to louder sound pressure levels, smoothing the attack part of the sound and giving a distinctive color to piano tones.

From the physical point of view, the soundboard is a moderately thin (around 0.5-1cm in thickness) wooden plate reinforced with several *ribs* in order to maintain the necessary stress needed to support the tension of all the strings.

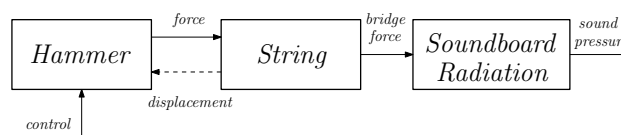


Figure 1. General architecture of a physics-based piano synthesizer.

Generally the amplitude of vibration is small enough that linear models can be considered without significant reductions in quality. Both experimental measurements [3–5] and offline numerical simulations [6, 7] have shown how the force-pressure transfer functions of piano soundboards vary significantly for different excitation positions, i.e., it changes for every note of the instrument due to the variability of normal mode shapes.

Nevertheless, in the majority of physics-based sound synthesis methods, soundboard radiation is modeled as a single filter for all the strings, without considering this differences along the keyboard or using a simple equalization approach [8]. Soundboard response variability for each note can be obtained when *commuted synthesis* is employed [9], i.e. when the response is factored together with the hammer excitation and stored as a lookup table. A variation of this technique consists in using short, truncated excitation signals to model the early part of the soundboard impulse response and a reverb-like algorithm to model the decaying tail, as it is done for the harpsichord in [10]. Nevertheless, factoring the soundboard impulse response into the excitation signal is possible only if the rest of the system is linear and time invariant, which is quite a strong assumption in the case of piano strings, and in addition runtime changes of the parameters becomes more difficult to achieve.

In this paper a method for efficiently modeling the soundboard force-pressure transfer function at different bridge positions is proposed¹, following an approximation of the low-frequency range with a parallel set of second order resonators that share the same poles for different excitation positions. The method has been implemented in a physics-based digital piano recently introduced into the market [12] following a joint research between Viscount International SpA and the Universities of Verona and Parma [2].

The algorithm needs a set of target impulse responses that can be either obtained from offline numerical simulations

¹ The actual DSP structure described here is also part of a filed PCT patent [11], which did not cover measurement and parameters estimation details.

or by direct measurement of real piano soundboards, which are described in Sec. 2. The DSP structure of the proposed method is explained in Sec. 3, while Sec. 4 propose a robust estimation procedure for the parameters involved. Results and computational cost issues are discussed in Sec. 5. Finally, Sec. 6 concludes the paper.

2. SOUNDBOARD IR MEASUREMENTS



Figure 2. Microphone setup during the recording sessions of Seiler mod. D (top) and Yamaha YUS1 (bottom).

Three different pianos were measured during various recording sessions that took place at the Viscout International SpA semi-anechoic room in Mondaino (RN), Italy between July and September 2012. The instruments chosen were one concert grand piano (*Steinway mod. D*), a smaller piano with a distinctive tone (*Seiler mod. 242 D*) and an entry-level upright (*Yamaha YUS1*). Sound absorbing material was placed below and around the instruments in order to minimize the reflections from the floor. Pictures of the setup can be seen in Fig. 2.

The main objective of the recording sessions was to obtain the transfer functions between the transversal force at the bridge and the acoustic pressure near the instrument. We thus hit the metal pins at which strings are attached with a 30 grams hammer equipped with a PCB force transducer. The microphones used include a pair of AKG 414 in A/B configuration (used primarily for near-field recording), a Schoeps ORTF MSTC 4 U stereo microphone (used to capture the far-field response) and a custom-made wooden sphere with two Schoeps MK4 capsules placed in the pianist's position. Piano strings were carefully damped by the use of pressure-sensitive adhesive and other sources of noises such as pedal bars were eventually removed or dampened.

We developed an interactive tool in Python to analyze and process the measured data (Fig. 3). In this way, the sound designer can visualize various aspects of the impulse response he is working with, while being able to apply post-processing effects and load the result in a real-time for piano physical model. The software automatically segments a multichannel recording corresponding to a set of hits for a single soundboard position, then computes the impulse response by the use of regularized frequency-domain deconvolution techniques [13].

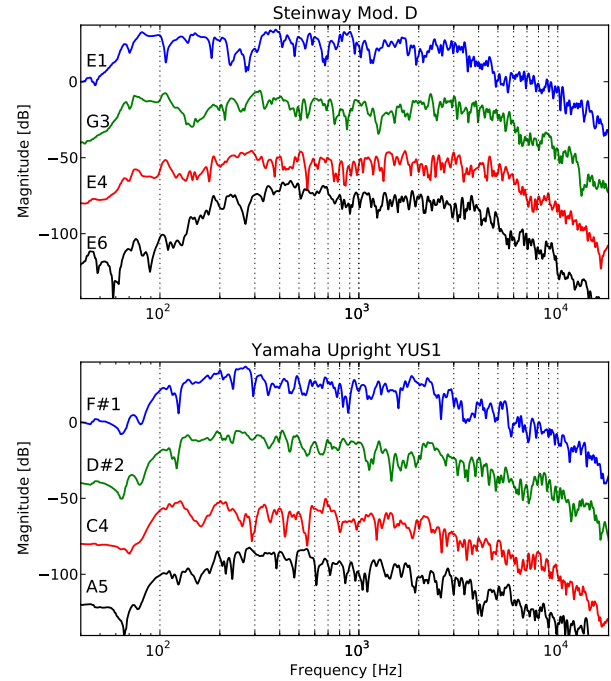


Figure 4. Measured force-pressure transfer functions for two different pianos at four bridge positions.

Deconvolution is a noise-sensitive process, especially in the high-frequency range due to the lowpass characteristic of the hammer force. As a simple and effective way to reduce noise, we employ a *frequency-dependent windowing* of the impulse response exploiting the fact that high-frequency modes of the soundboard decay much faster than low-frequency ones. In order to do so, the measured discrete-time impulse response $h(n)$ is split into low-frequency $h_{LP}(n)$ and high-frequency $h_{HP}(n)$ with a pair of complementary filters by zero-phase forward-backward filtering. The noise-reduced impulse response is then simply computed as

$$h_{NR}(n) = w_{LF}(n) h_{LP}(n) + w_{HF}(n) h_{HP}(n), \quad (1)$$

where $w_{LF}(n)$ is a low-frequency window having length equal to $h(n)$ and $w_{HF}(n)$ is a shorter high-frequency window whose length is set as a parameter. Filter and window types can also be chosen interactively, a typical choice for the latter being half or quarter-length Hanning windows in order to preserve the attack transients.

As the final outcome of this process, a large database of

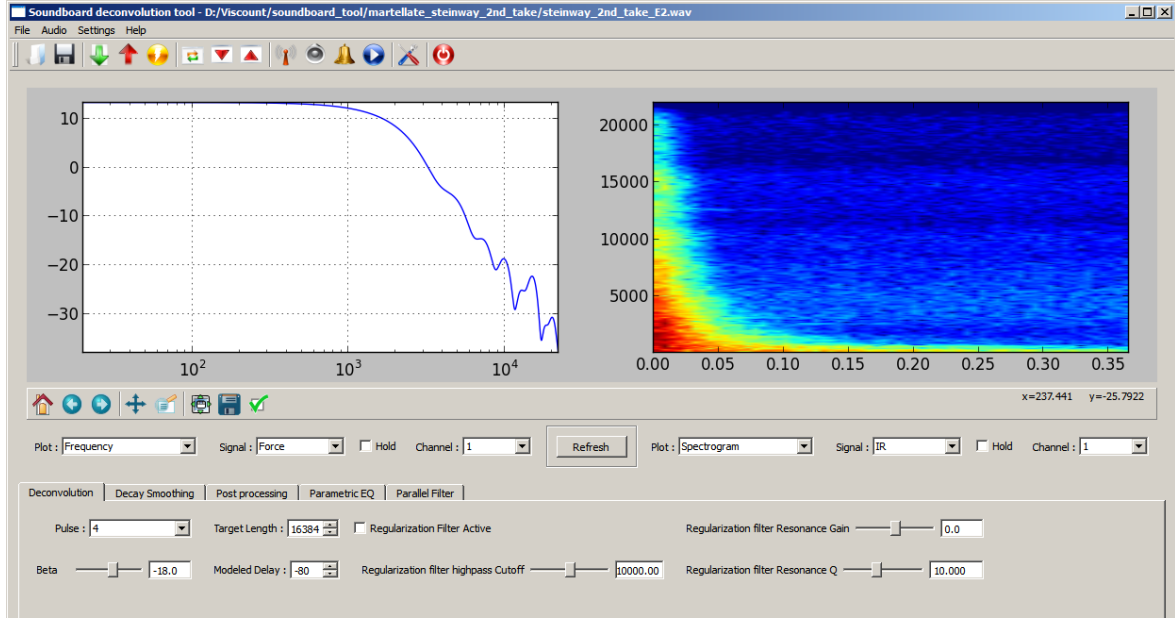


Figure 3. Screenshot of the interactive tool developed and used to analyze and process soundboard impulse responses.

impulse responses corresponding to various hitting positions on the bridge and different microphones or recording positions was obtained. With our setup, we found out that a length of 16384 samples at 44100 Hz was enough to cover the full decay to the noise floor level. The data obtained was in line with previous measurements presented in literature [3–5].

Impulse responses taken at different soundboard positions show both localized *micro-scale* differences, which are caused by the variability of the modal excitation patterns along the soundboard and *macro-scale differences*, particular in the low-frequency region where modes are less dense. Fig. 4 shows an example of such macro-scale variations: in the treble range the highpass character of the responses is usually more pronounced, while in the middle-range section of the keyboard (third to fourth octave) the frequencies in the range of 200-400Hz are often attenuated. Impulse responses can also be very different under the temporal profile², where generally in the treble range they present a smoother attack and a wider stereo image.

3. FIXED-POLE PARALLEL FILTERS

From the measured impulse responses we have obtained a set of FIR filters which can be implemented in real-time using e.g. FFT block-based convolution. However, when low latency is required and thus short buffer lengths (≤ 64 samples) have to be used, the computational load can be demanding if multiple positions and/or output channels are needed. Even though non-uniform partitioned convolution algorithms exist to significantly reduce the cost [14], their implementation is far from easy in a DSP environment where predictable computation times have to be ensured.

² See bottom plots in Fig. 7 or the companion webpage for some examples.

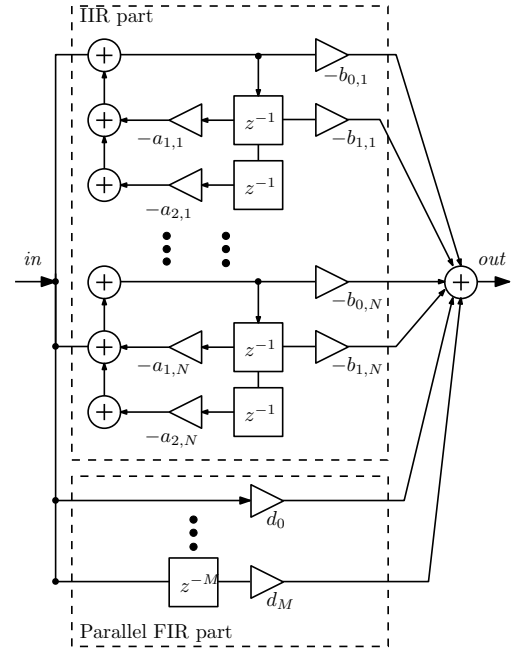


Figure 5. Block diagram for the system of Eq. (2) showing the parallel resonator block and the separate FIR part.

The idea behind fixed-pole parallel filters for instrument body modeling [15, 16] is to approximate a target impulse response with a IIR system composed by a set of N parallel second-order resonators and an additional M -th order FIR system connected in parallel:

$$H(z) = \sum_{n=1}^N \frac{b_{0,n} + b_{1,n}z^{-1}}{1 + a_{1,n}z^{-1} + a_{2,n}z^{-2}} + \sum_{m=0}^M d_m z^{-m}, \quad (2)$$

whose block diagram corresponding to Direct-Form 1 implementation can be seen in Fig 5. The poles of the system are set beforehand either by choosing a logarithmic fre-

quency distribution spanning the whole frequency range, trying to mimic the human ear resolution, or can be extracted using system identification techniques on the warped target response [17]. Once the poles and, consequently, the coefficients $a_{1,n}$, $a_{2,n}$ are fixed, the computation of the zeroes of the system in order to minimize the error between the modeled and target response becomes a *linear in-parameter* problem and can thus be solved using e.g. traditional least-squares fitting methods. The additional FIR part can help in the case of largely non-minimum phase responses, as those that are typically obtained from piano soundboard measurements.

3.1 Common-Pole Soundboard Modeling

Our primary goal in designing a filtering structure for soundboard radiation was being able to model the differences along the keyboard while maintaining high-quality results (comparable to full FIR filtering) and low computational costs. The key ideas behind the approximations are the following:

1. Time-domain characteristics have to be reproduced faithfully at least in the attack portion of the sound, in order to give the necessary "smoothing" to the forces synthesized by the physics-based string model.
2. We do not need to model the soundboard frequency response for each note precisely. Instead, we can chose a limited set of P (2-8) responses that capture the macro behaviour of the soundboard in different keyboard regions. Eventually, single-note frequency deviations can be easily included in the string model when modal synthesis is used [2].
3. The significant perceptual differences between transfer functions taken at different bridge positions are in the low-frequency range, especially below 4-5kHz. This happens partly because the ear is obviously more sensitive to differences in that range and partly because high frequency modes decay faster and are more densely distributed. Moreover, measurement noise becomes significant in the high-frequency range and thus impulse responses are more reliable for low frequencies.
4. The frequencies and decay times of soundboard normal modes do not vary with excitation position. The differences are found in modal amplitudes, which are determined by the modal shapes at each resonant frequency. Therefore, using a common-pole description for the chosen soundboard regions and varying only the zeroes between them seems to be a natural choice for lower the computational complexity [18].

The computational structure resulting from these choices is depicted in Fig. 6. Force signals coming from string synthesis modules are grouped in P separated splits and for each of them a stereo signal is derived using a simple panning model made up of gains and delays. Although not much physical in nature, this is a simple and effective way

to differentiate the soundboard model inside a specific keyboard region. Left and right signals are then processed by a separate filtering structure, designed to match a particular microphone channel among the ones available.

The first part of the system can be interpreted as P separate IIR systems having order N . Following the signal flow in Fig. 6, first the N zeroes of the system are computed separately for each region, then the signals for each mode are summed and passed through a common all-pole system made up of parallel second-order sections. Finally, a medium-length (512-1024 samples) FIR filter is applied at the end of the chain to help match the temporal behaviour in the attack portion of the sound.

The main difference between the DSP architecture in Fig. 6 and the original formulation of the parallel-filter structure in Fig. 5 is the serial connection of the FIR filtering block to the rest of the system instead of a parallel one. There is actually a parallel FIR part in the proposed scheme, consisting only of a single gain g_p but it is just used to compensate eventual gain differences between the responses. Serial connection works better in this case since the FIR part is also factored between the responses and can be efficiently implemented by a separate FFT convolution algorithm. Moreover, the linear systems needed to solve the least-squares problem coming from (2) can be badly conditioned with piano soundboard impulse responses when few poles and a long FIR filter are needed.

4. PARAMETER ESTIMATION

The goal of the estimation procedure is deriving the parameters of the algorithm shown in Fig. 6 starting from a set of P target impulse responses $h_{n,p}$ with $1 \leq p \leq P$. Since we are employing a completely independent system for each of the output channels, we are not making explicit references to the channel in the notation.

The first parameters that need to be chosen are the number of the poles N and the length M of the FIR part, with typical values being in the range of 24-96 for N and 512-1024 for M . Then, the poles of the system have to be set by either analyzing the set of target responses or by imposing a predetermined distribution. While there are methods in literature that estimate the best common-poles for a given set of impulse responses [19], we chose instead to use a parametric poles distribution in order to have a more robust procedure. When working in a highly-interactive environment, in fact, having non-optimal but predictable results can be a better choice than more sophisticated fitting algorithms that need extra parameters or that can behave badly in particular cases.

Thus, the poles are distributed logarithmically between a minimum and maximum frequency f_{\min} , f_{\max} , with f_{\min} typically set around the first relevant resonance of the soundboard (60-80 Hz) and f_{\max} controls the bandwidth over which we want to differentiate the multiple responses (usually between 3 and 6 kHz). The logarithmic distribution is a common choice in this context [15, 16] since it mimics the human ear frequency resolution. However, the modal density of soundboard impulse responses grows with frequency, therefore having closely spaced poles in the low-

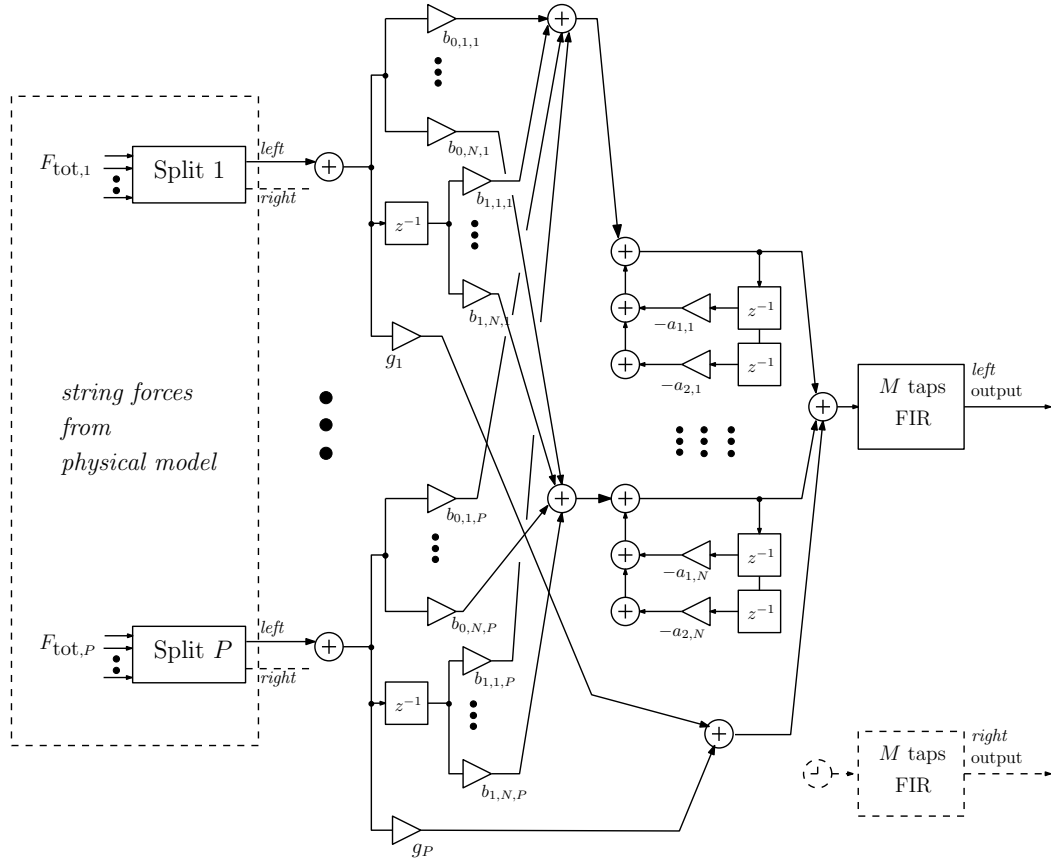


Figure 6. Block diagram for position-dependent soundboard modeling consisting in a common-pole parallel filter section followed by a common FIR filter for each channel.

frequency region may not be an optimal solution. We therefore offered as an alternative to the sound designer a mixed linear/logarithmic distribution, where N_{lin} poles are spaced linearly up to a frequency f_{lin} (around 200-500 Hz) and logarithmically above this limit, taking care to have a monotonically increasing distribution of frequency increases. More precisely, the frequency of the n -th pole f_n is determined by

$$f_n = \begin{cases} f_{\text{min}} + \Delta f n, & \text{if } 0 \leq n < N_{\text{lin}} \\ f_{\text{lin}} + \frac{\Delta f}{f_L - 1} (f_L^{n - N_{\text{lin}}} - 1) & \text{if } N_{\text{lin}} \leq n < N \end{cases} \quad (3)$$

where

$$\Delta f = f_{\text{lin}} - \frac{f_{\text{lin}} - f_{\text{min}}}{N_{\text{lin}} - 1}$$

and the logarithmic frequency base f_L is found as the positive root of the polynomial

$$\Delta f f^{N - N_{\text{lin}} - 1} + (f_{\text{lin}} - f_{\text{max}}) f + f_{\text{max}} - f_{\text{min}} - \Delta f.$$

Examples of this distribution can be seen by looking at the vertical markers in the top plots of Fig. 7. Pole radii are set either at constant Q , specified as a parameter, or in a way that the magnitude responses of the filter crosses each other at the -3dB point [20]. Clearly, the remaining N poles are distributed in order to form complex-conjugated pairs.

The common FIR filter h_{FIR} also needs to be set beforehand and is thus not part of the parameter estimation procedure. Natural choices in this case are, e.g., taking the first M samples (using a window function) of one of the P target responses, $h_{n,\bar{p}}$, which becomes the “main” response between the set, i.e. the one which will be approximated better than the others. Otherwise, the time average of all the windowed target responses can be used leading to a better overall fit in the frequency domain. However, this method is useful only if the parallel filter order is high enough ($N \geq 64$) to ensure proper time modeling of the attack part, otherwise we will get P different responses none of which may show the characteristic attack of piano tones.

When a single preferred response having index \bar{p} is chosen, a useful preprocessing step consists in modifying the target responses in order that they all share the same high-frequency content above the frequency limit f_{max} of the resonators, giving in this way an “easier” target to the procedure. More specifically, referring to the Zeta-transforms of the target responses, we can compute the modified targets $H'_{t,p}(z)$ as

$$H'_{t,p}(z) = H_{\text{LP}}(z) H_{t,p}(z) + H_{\text{HP}}(z) H_{t,\bar{p}}(z), \quad (4)$$

where $H_{\text{LP}}(z)$, $H_{\text{HP}}(z)$ are a pair of complementary low and highpass filters having cutoff frequency set to f_{max} .

At this point, for each response index p we need to estimate the set of free parameters

$$\mathbf{q}_p = [b_{0,1,p}, b_{1,1,p}, \dots, b_{0,N,p}, b_{1,N,p}, g_p]$$

that determine the zeroes of the IIR part whose transfer function will be denoted as $H_{\text{IIR},p}(z)$. Writing the expression for the resulting response $H_p(z)$ we have:

$$H_p(z) = H_{\text{IIR},p}(z) H_{\text{FIR}}(z) = \sum_{n=1}^N \frac{b_{0,n,p} + b_{1,n,p}z^{-1}}{1 + a_{1,n,p}z^{-1} + a_{2,n,p}z^{-2}} H_{\text{FIR}}(z) + g_p H_{\text{FIR}}(z). \quad (5)$$

Therefore, optimizing the set of parameters \mathbf{q}_p can be seen as a system identification problem where we desire to minimize the error between the target impulse response $h'_{t,p}(n)$ and the output of the IIR system in (5) when the signal $h_{\text{FIR}}(n)$ is used as input. Since the problem is linear in its parameters, traditional fitting method such as least-squares regression can be used in the same way as it is done, e.g., in [16] for room-equalization system identification problems. In this way, an overdetermined linear system is derived from the specifications as

$$\mathbf{h}_{t,p} \simeq \mathbf{M}_p \mathbf{q}_p, \quad (6)$$

where \mathbf{M}_p is the modeling matrix whose rows consist of the result of filtering the input signal $h_{\text{FIR}}(n)$ by the response of each parallel section, plus P rows for the parallel gain coefficients g_p . Minimizing the squared error between the computed and target output we can obtain the optimal set of parameters $\mathbf{q}_{\text{opt},p}$ as

$$\mathbf{q}_{\text{opt},p} = (\mathbf{M}_p^H \mathbf{M}_p)^{-1} \mathbf{M}_p^H \mathbf{h}_{\text{FIR}}, \quad (7)$$

where the operator H stands for the complex conjugation operation.

Finally, the estimation procedure is iterated 5-10 times updating the phase of the target response at each step, implementing a *magnitude-priority* design method described in [21]. This is a compromise between using a magnitude-only specification for the target filter and full complex design, which generally produces a better magnitude fit without sacrificing too much the temporal characteristics of the result.

Luckily, the procedure is robust enough so that typically there is no need to change the estimation parameters for a given set of soundboard recordings, so the sound designer can focus his attention on other aspects more relevant from the perceptual point of view.

5. DISCUSSION

Fig. 7 shows the results for $P = 4$ responses taken from those measured with the Seiler mod. 242 D piano. We can see that magnitude of the frequency response is very close to the target in the specified frequency range, which can be increased when more resonators are available. Time-domain fitting for the responses different than the "main" target is accurate only when total filter order N and specified bandwidth f_{max} are high enough to capture the differences. However, in this case even with 32 filters all the responses maintain a piano-like characteristic and, although

there may be audible differences when compared to the target, we are still capturing most of its macro-scale features. For sound examples and more comparison plots, see the companion web page of this article³.

From the computational point of view, the proposed algorithm is much faster than the original set of FIRs implemented by uniform block-based FFT convolution. Referring to the structure of Fig. 6, the number of Multiplies per Output Sample (MPOS) is

$$\text{MPOS} = \frac{(2N + 1)P + 2N + 4(B + 1)\lceil \frac{M}{B} \rceil + 2\text{MPOS}_{\text{FFT}}}{B},$$

where B is the audio buffer size which is generally no greater than 128 samples at 44100Hz for low-latency synthesis and MPOS_{FFT} is the cost for a FFT operation (which is in the order of $\mathcal{O}(B \log B)$). In our example, with $B = 64$, $P = 4$, $N = 64$, $M = 768$, the speedup compared to the full set of 16384 taps convolutions via FFT is in the order of 50x and scales almost linearly with the number of soundboard responses P .

6. CONCLUSIONS

The force-pressure transfer functions of three different pianos have been measured at various bridge and microphone positions. From their analysis, an efficient filtering technique that is being able to capture the most important sound characteristics has been derived. The method enables the efficient simultaneous computation of multiple transfer functions, each one related to a different spatial region of the soundboard, by employing a low-frequency approximation with common-pole parallel second-order filters. The estimation procedure is robust and can be easily applied to a set of target impulse responses by manipulating few, intuitive parameters. Although the discussion has been restricted to piano soundboard modeling, the method can also be applied for other instruments having a similar radiation block or as an alternative to full convolution in the case of room simulation when different source positions have to be considered.

Acknowledgments

Measurement equipment, plus assistance during the recording sessions, have been provided by Angelo Farina's acoustic research group at the University of Parma, Federico Fontana of the University of Udine and David Monacchi of the Conservatory of Pesaro. The author wish to thanks Eugenio Giordani for the work during the recording sessions and the subsequent feedback on the developed method. A special acknowledgment goes to Bálažs Bank for the precious suggestions during the investigation phase and the detailed comments on this manuscript.

7. REFERENCES

- [1] B. Bank, F. Avanzini, G. Borin, G. De Poli, F. Fontana, and D. Rocchesso, "Physically informed signal processing methods for piano sound synthesis: a research overview," *EURASIP Journal on Applied Signal Processing*, vol. 2003, pp. 941–952, 2003.
- [2] B. Bank, S. Zamboni, and F. Fontana, "A modal-based real-time piano synthesizer," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 4, pp. 809–821, 2010.

³ <http://www.physispiano.com/smac2013/index.html>

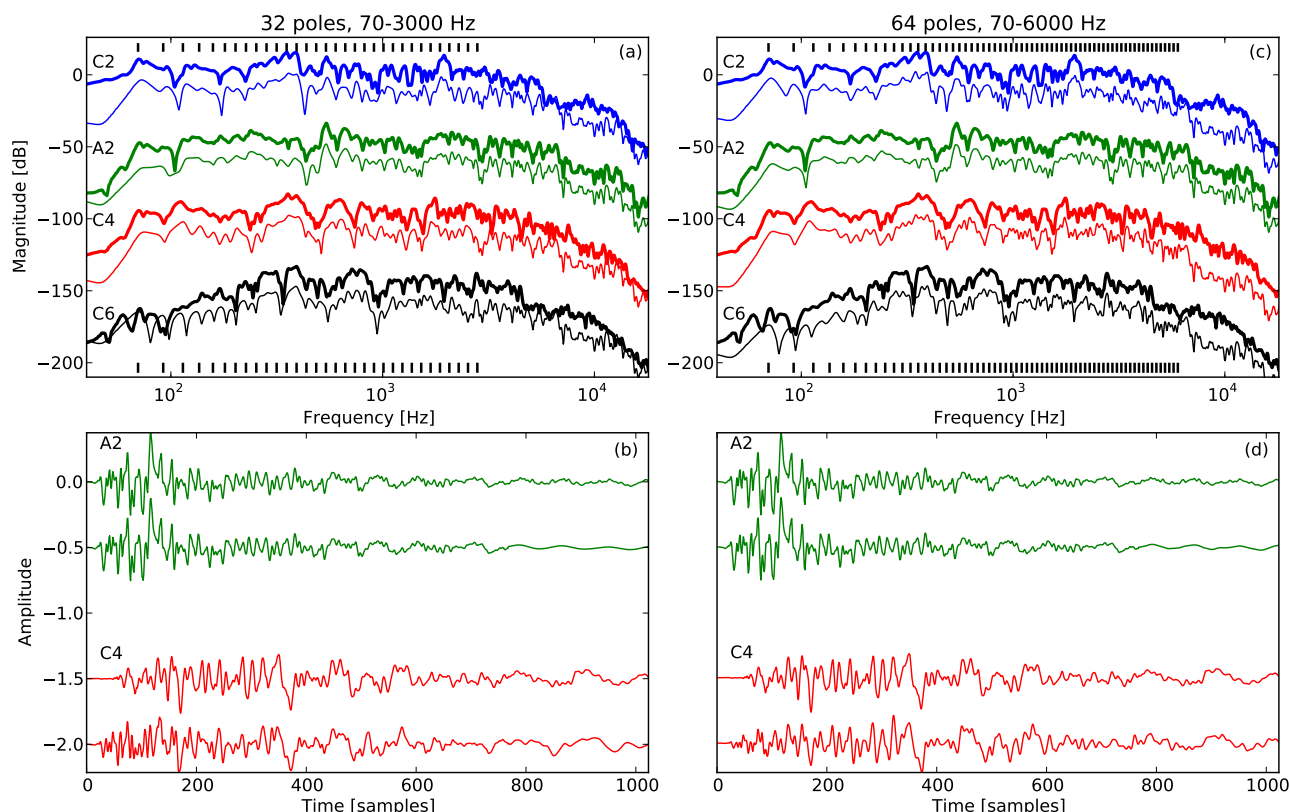


Figure 7. Filter design example for a set of 4 responses of the Seiler piano, with 32 resonators up to 3kHz (a, b) and 64 resonators up to 6kHz(c, d). M is set to 768 samples in both cases. The top plot shows the magnitude response of the target (16384 samples, plotted as thick lines) along with the modeled responses (thin lines). Bottom plots show the target and modeled time responses for the "main" response index ($\bar{p} = 1$) and one of the other targets. Pole frequencies are visualized in top plots with a set of vertical markers. Curves have been offset for clarity.

- [3] K. Wogram, "Acoustical research on pianos: vibrational characteristics of the soundboard," *Das Musikinstrument*, vol. 24, pp. 694–702, 1980.
- [4] H. Suzuki, "Vibration and sound radiation of a piano soundboard," *The Journal of the Acoustical Society of America*, vol. 80, 1986.
- [5] N. Giordano, "Sound production by a vibrating piano soundboard: Experiment," *The Journal of the Acoustical Society of America*, vol. 104, 1998.
- [6] —, "Simple model of a piano soundboard," *The Journal of the Acoustical Society of America*, vol. 102, p. 1159, 1997.
- [7] J. Berthaut, M. Ichchou, and L. Jezequel, "Piano soundboard: structural behavior, numerical and experimental study in the modal range," *Applied Acoustics*, vol. 64, no. 11, pp. 1113–1136, 2003.
- [8] B. Bank, "Physics-based sound synthesis of the piano," Master's thesis, Budapest Univ. of Technol. and Economics, Budapest, Hungary, 2000.
- [9] S. A. Van Duyne and J. O. Smith, "Developments for the commuted piano," in *Proceedings of the 1995 International Computer Music Conference, Banff*. ICMA, 1995, pp. 335–343.
- [10] V. Välimäki, H. Penttinen, J. Knif, M. Laurson, and C. Erkut, "Sound synthesis of the harpsichord using a computationally efficient physical model," *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 934–948, 2004.
- [11] S. Zambon, E. Giordani, B. Bank, and F. Fontana, "A system to reproduce the sound of a stringed instrument," Filed PCT international patent, March 2013.
- [12] Viscount International, "Physis Piano," <http://www.physispiano.com>, 2013.
- [13] O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *Speech and Audio Processing, IEEE Transactions on*, vol. 6, no. 2, pp. 189–194, 1998.
- [14] E. Battenberg and R. Avizienis, "Implementing real-time partitioned convolution algorithms on conventional operating systems," in *Proceedings of the 14th International Conference on Digital Audio Effects*. Paris, France, 2011.
- [15] B. Bank, "Direct design of parallel second-order filters for instrument body modeling," in *Proc. International Computer Music Conference (ICMC 2007)*, Copenhagen, Denmark, 2007, pp. 458–465.
- [16] —, "Perceptually motivated audio equalization using fixed-pole parallel second-order filters," *IEEE Signal Processing Letters*, vol. 15, pp. 477–480, 2008.
- [17] B. Bank and G. Ramos, "Improved pole positioning for parallel filters based on spectral smoothing and multiband warping," *IEEE Signal Processing Letters*, vol. 18, no. 5, pp. 299–302, 2011.
- [18] B. Bank, "Modeling the variation of soundboard response by shaping filters," Internal technical report for Viscount International SpA, 2010.
- [19] Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 320–328, 1994.
- [20] B. Bank, "Audio equalization with fixed-pole parallel filters: An efficient alternative to complex smoothing," *J. Audio Eng. Soc.*, vol. 61, no. 1/2, pp. 39–49, 2013.
- [21] —, "Magnitude-priority filter design for audio applications," in *132th Audio Engineering Society Convention*, 2012.

Computing virtual acoustics using the 3D finite difference time domain method and Kepler architecture GPUs

Craig J. Webb

Acoustics group / EPCC, University of Edinburgh
C.J.Webb-2@sms.ed.ac.uk

ABSTRACT

The computation of virtual acoustics for physical modelling synthesis using the finite difference time domain is a computationally expensive process, especially at audio rates such as 44.1kHz. However, the high level of data-independence is well suited to parallel architectures such as those provided by graphics processing units. This paper describes the use of the latest Nvidia Kepler cards to accelerate the computation of three-dimensional schemes. The CUDA language and hardware architecture allow many possible approaches to computing even a basic model. Various techniques are considered, such as full tiling, iteration slicing, and the use of shared memory. A standard simulation was used to measure the performance of these different approaches. Benchmark times were compared for the latest Nvidia Tesla K20 GPU against the previous generation cards. Results show the continuing maturity of the hardware, especially in terms of data caching, which allows basic code designs to perform as well as more complex shared memory versions.

1. INTRODUCTION

Virtual acoustics can be approached by direct numerical simulation of the three-dimensional wave equation. This can be used for auralizations, creating a model of a virtual environment, or for physical modelling synthesis by embedding instruments into the space. The finite difference time domain (FDTD) method is an efficient technique for computing such simulations [1]. However, for 3D systems at audio sample rates such an approach is still extremely computationally expensive [2], but can benefit from parallel computing using graphics processing units (GPUs) [3].

This paper examines the use of the latest Nvidia Kepler architecture GPUs to accelerate the computation of the standard FDTD discretisation of the 3D wave equation. The CUDA language and GPU hardware allow many different approaches to this particular computation, each of which produce the same output but with varying efficiency. Finding the optimal solution is a matter of experimentation and tuning of different implementation methods.

This paper examines six different approaches to a standard simulation model, ranging from 2D to 3D threading

and paying close attention to the use of shared memory. These solutions are benchmarked on the latest Nvidia Tesla K20 card, as well as the previous generation Fermi Tesla card for comparison.

2. FINITE DIFFERENCE SCHEME

Virtual acoustic simulations using FDTD are based on the 3D wave equation, which in second order form is given by:

$$\frac{\partial^2 \Psi}{\partial t^2} = c^2 \nabla^2 \Psi \quad (1)$$

Here Ψ is the target acoustical field quantity, c is the wave speed in air, ∇^2 is the 3D Laplacian. The standard FDTD discretisation [4] leads to the following update equation for interior grid points:

$$w_{l,m,p}^{n+1} = (2 - 6\lambda^2)w_{l,m,p}^n + \lambda^2 S_{l,m,p}^n - w_{l,m,p}^{n-1} \quad (2)$$

where $w_{l,m,p}$ is the discrete acoustic field, $\lambda = \frac{cT}{X}$, and

$$S_{l,m,p}^n = w_{l+1,m,p}^n + w_{l-1,m,p}^n + w_{l,m+1,p}^n + w_{l,m-1,p}^n + w_{l,m,p+1}^n + w_{l,m,p-1}^n \quad (3)$$

The stability condition for the scheme can be derived from von Neumann analysis [4], such that for a given time step T the grid spacing X must satisfy:

$$X \geq \sqrt{3c^2 T^2} \quad (4)$$

At the Courant limit where $\lambda = 1/\sqrt{3}$, the update equation reduces to:

$$w_{l,m,p}^{n+1} = \frac{1}{3} S_{l,m,p}^n - w_{l,m,p}^{n-1} \quad (5)$$

Fixed boundary conditions were used for testing. The update equation for a given grid point uses the six nearest neighbours from one time step ago, and the centre point from two time steps ago, as shown in figure 1.

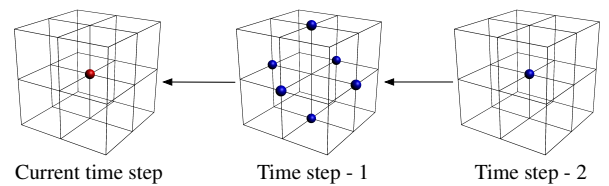


Figure 1. Grid points used for the update equation.

3. GPU COMPUTING USING CUDA

Nvidia's Kepler architecture is the 3rd generation of GPU cards that can be used specifically for general purpose computing using the CUDA language. The original compute 1.x cards were followed in 2010 by the Fermi architecture 2.x cards. With each generation, the hardware has changed significantly, along with the development of features in CUDA. Whilst the 3D FDTD wave equation computation is trivial to parallelize at each time step of the simulation, the downside is that it is clearly memory bandwidth limited. The compute-to-memory access ratio for each update is very low (generally less than two), and so achieving optimal efficiency on the GPU depends solely on the movement of data around the system.

GPU cards provide multiple memory types that can be programmed directly, as well as optimisations such as directing the use of cache lines [5]. These, together with the multiple options for threading the given data set, provide a wide scope of differing approaches to any given problem. Six different methods are considered here for the 3D FDTD scheme.

4. IMPLEMENTATION METHODS

The implementation design for the scheme consists of setup code, followed by a loop over of the time iterations of the simulation. Within this time loop the CUDA kernel threads are launched that update the state of the system, followed by processing the input and output, as shown in figure 2. Only two data grids are required for this basic scheme, as

Figure 2. Time loop with kernel launches.

```

1  for (n=0;n<NF;n++)
2  {
3      UpDate<<<dimGrid,dimBlock>>>>(u_d,u_l_d,L2);
4      // perform I/O
5      inout<<<I,I>>>>(u_d,out_d,ins,n);
6      // update pointers
7      dummy_ptr = u_l_d;
8      u_l_d = u_d;
9      u_d = dummy_ptr;
10 }
```

the values from two time steps ago can be read from memory before overwriting the new state values.

The first consideration in terms of the threading design is whether to issue enough threads to update the entire 3D state, or to issue threads that cover a 2D slice and then use an iteration within the kernel over the remaining dimension. The latter approach allows greater flexibility in terms of data reuse, whilst limiting the number of threads in use. The kernel code should of course be designed to maximise *memory coalescing* with regard to the decomposition of the 3D data into linear memory.

Both approaches can utilise shared memory, which allows blocks of threads to store and use data in a collaborative manner. The main issue with using shared memory for finite difference schemes is the complication of always needing to access neighbouring grid points when at the edges of a thread block. Two different approaches are considered here, for each of the thread designs.

4.1 3D Tiling

The first method is the 3D tiling approach, where the entire 3D data set is covered by individual threads. So if the data set consists of one million grid points, then one million threads are issued to update the state. Threads are launched in groups known as *blocks*, and multiple blocks then make up the thread *grid*. Each of these objects can be one, two or three dimensional. Experimentation showed that a 32 x 4 x 2 thread block is most efficient here. The kernel code is shown in figure 3. Note that the neighbouring data

Figure 3. 3D Tiling kernel.

```

1  --global-- void UpDate(double *u,double *u_l,double L2)
2  {
3      // get X,Y,Z from thread and block Id's
4      int X = blockIdx.x * Bx + threadIdx.x;
5      int Y = blockIdx.y * By + threadIdx.y;
6      int Z = blockIdx.z * Bz + threadIdx.z + 1;
7
8      // Test that not at halo, Z block excludes Z halo
9      if ( (X>0) && (X<(Nx-1)) && (Y>0) && (Y<(Ny-1)) ){
10
11         // get linear position
12         int cp = Z*area+(Y*Nx+X);
13
14         u[cp] = L2*(u_l[cp-1]+u_l[cp+1]+u_l[cp-Nx]
15                 +u_l[cp+Nx]+u_l[cp-area]+u_l[cp+area])-u[cp];
16     }
17 }
```

points are accessed using shifts, and so only one calculation of the linearly decomposed position is required. Bx, By and Bz define the size of the thread block, with Nx, Ny and Nz defining the size of the 3D data grid, and 'area' is defined as Nx*Ny. This kernel is the simplest possible arrangement, reading data directly from global memory. The next step is to attempt to minimise data movement by using shared memory.

4.2 3D Tiling with shared memory

In order to implement a shared memory version of the above kernel, the main issue is how to deal with data access at the edges of the thread block. A 2D shared memory array is used, and in this version it will be the same size as the 2D thread block that is employed here. A block size of 32 x 8 was found to be most efficient. Figure 4 shows the new kernel code.

Each thread loads one element of data into the shared memory array (at line 16), followed by a thread synchronisation. Having tested that the current position is not a boundary point, the sum of the neighbouring grid points is computed. This requires four conditional statements, which pick up data from global memory if the position is it the edge of a thread block, otherwise it is read from the shared memory array. The final line reads the remaining Z-dimension neighbours, and writes the updated value to global memory.

The overall effect is to reduce the reads from global memory from six to two, when a given thread is not at the edge of a block, which is a significant reduction in memory access.

Figure 4. 3D Tiling with shared memory kernel.

```

1  __global__ void UpDate(double *u, double *u1, double L2)
2  {
3      __shared__ double uS1[BxS][ByS];
4
5      int tdx = threadIdx.x;
6      int tdy = threadIdx.y;
7
8      int X = blockIdx.x * BxS + tdx;
9      int Y = blockIdx.y * ByS + tdy;
10     int Z = blockIdx.z + 1;
11
12     int cp = Z*area+(Y*Nx+X);
13     double sum = 0.0;
14
15     // Load shared
16     uS1[tdx][tdy] = u1[cp];
17     __syncthreads();
18
19     // Test that not at halo, Z block excludes Z halo
20     if ( (X>0) && (X<(Nx-1)) && (Y>0) && (Y<(Ny-1)) ) {
21
22         if (tdx==0) sum+= u1[cp-1];
23         else sum+= uS1[tdx-1][tdy];
24
25         if (tdx==BxS-1) sum+= u1[cp+1];
26         else sum+= uS1[tdx+1][tdy];
27
28         if (tdy==0) sum+= u1[cp-Nx];
29         else sum+= uS1[tdx][tdy-1];
30
31         if (tdy==ByS-1) sum+= u1[cp+Nx];
32         else sum+= uS1[tdx][tdy+1];
33
34         u[cp] = L2*(sum+u1[cp-area]+u1[cp+area]) - u[cp];
35     }
36 }

```

4.3 3D Tiling with extended shared memory

A second approach to dealing with the problem of picking up data at the edges of a thread block is to use a shared memory array which is larger than the thread block and contains this edge data. Provided that the data is loaded correctly, a shared memory array of size $[Bx+2][By+2]$ will contain all necessary data for the X and Y neighbour points. This ‘extended’ approach requires a more complicated arrangement for loading the shared memory data, but results in a much cleaner update from the neighbouring values as shown in figure 5.

Instead of four conditional statements around the summing point of the code, there are now four conditionals used in loading the shared memory array. If a thread is at the edge of a block it loads its own point plus one extra edge point. This loads the entire extended shared memory array, and so the update at line 36 reads all X and Y neighbours directly from the array.

Whilst this approach leads to the same reduction in global memory reads as the non-extended version, it does have a major effect on efficiency, as detailed in section six.

4.4 2D Slicing

The first three methods above have all focussed on a 3D tiling approach using the maximum amount of threading. The next three methods will employ a different strategy, namely using a 2D slicing approach. Instead of issuing threads to cover the entire data set, only enough threads are issued to cover a 2D slice of the data, for example a Z slice of size X by Y. Each thread then iterates over the

Figure 5. 3D Tiling with extended shared memory kernel.

```

1  __global__ void UpDate(double *u, double *u1, double L2)
2  {
3      __shared__ double uS1[BxS+2][ByS+2];
4
5      int tdx = threadIdx.x;
6      int tdy = threadIdx.y;
7
8      int X = blockIdx.x * BxS + tdx;
9      int Y = blockIdx.y * ByS + tdy;
10     int Z = blockIdx.z + 1;
11
12     // get linear position
13     int cp = Z*area+(Y*Nx+X);
14
15     // Load shared
16     tdx++; tdy++;
17     uS1[tdx][tdy] = u1[cp];
18
19     if ( (tdy==1) && !(Y==0) ) {
20         uS1[tdx][tdy-1] = u1[cp-Nx];
21     }
22     if ( (tdy==ByS) && !(Y==(Ny-1)) ) {
23         uS1[tdx][tdy+1] = u1[cp+Nx];
24     }
25     if ( (tdx==1) && !(X==0) ) {
26         uS1[tdx-1][tdy] = u1[cp-1];
27     }
28     if ( (tdx==BxS) && !(X==(Nx-1)) ) {
29         uS1[tdx+1][tdy] = u1[cp+1];
30     }
31     __syncthreads();
32
33     // Test that not at halo, Z block excludes Z halo
34     if ( (X>0) && (X<(Nx-1)) && (Y>0) && (Y<(Ny-1)) ) {
35
36         u[cp] = L2*(uS1[tdx-1][tdy]+uS1[tdx+1][tdy]
37                  +uS1[tdx][tdy-1]+uS1[tdx][tdy+1]
38                  +u1[cp-area]+u1[cp+area]) - u[cp];
39     }
40 }

```

remaining dimension (i.e. the Z dimension), updating all the grid points along that column. This of course requires a loop inside the kernel code itself.

To some extent, this approach is counter-intuitive. The main goal of parallel programming is to perform as much work as possible in parallel. However, there are advantages in using this approach, mainly in that it is easier to re-use data. In particular, it is possible to remove one of the two global memory reads in the Z dimension, as the thread is iterating over that data. By combining this with an X by Y shared memory tile, it is possible to reduce the global memory reads down to one single access per iteration over the Z dimension.

Prior to the Fermi architecture GPU cards, this technique was the most efficient approach to use with 3D data sets due to the lack of effective caching on the early cards [6]. However, this changed significantly with the compute 2.x architecture of Fermi, and made the 3D tiling approach useful. As caching levels continue to improve, it also leads to improved performance of direct global memory accessing, as will be demonstrated.

The kernel code for the basic global memory version of the 2D slicing is shown in figure 6. This closely resembles the kernel for the original 3D tiling method, but with the addition of a FOR loop over the interior of the Z dimension. A 64×8 thread block size was found to be the most efficient.

Figure 6. 2D Slicing kernel.

```

1  __global__ void UpDate(double *u, double *u1, double L2)
2  {
3      // get X,Y,Z from thread and block Id's
4      int X = blockIdx.x * BxL + threadIdx.x;
5      int Y = blockIdx.y * ByL + threadIdx.y;
6      int Z, cp;
7
8      // Test that not at halo
9      if ( (X>0) && (X<(Nx-1)) && (Y>0) && (Y<(Ny-1)) ){
10
11          for (Z=1; Z<(Nz-1); Z++){
12
13              // get linear position
14              cp = Z*area+(Y*Nx+X);
15
16              u[cp]=L2*(u1[cp-1]+u1[cp+1]+u1[cp-Nx]+u1[cp+Nx]
17                  +u1[cp-area]+u1[cp+area]) - u[cp];
18          }
19      }
20  }

```

4.5 2D Slicing with shared memory

As with the 3D tiling methods, the shared memory implementation has to account for the pickup of neighbouring data at the edges of the thread blocks. The same approaches are used. Firstly a method that uses a block size shared memory array with conditional statements around the summing point. Secondly, using an extended size shared memory array, with conditional statements used at the loading point of the code. The kernel code for these methods is shown in figures 7 and 8.

4.6 Cache optimisations

Aside from the design of the kernel code itself and experimenting with the size of the thread block, further speedups can be obtained by optimising the cache usage. On the compute 2.x Fermi cards, using the *cudaFuncSetCacheConfig()* command to prefer the L1 cache produces efficiency gains for all the above kernels, some by as much as 15%.

An additional feature of the Kepler architecture is the ability to use a read-only data cache which is separate from the standard L1 and L2 cache [5]. The above kernels can make use of this when accessing data from the six neighbour points from one time step ago. In the parameters of the kernel declaration, the data pointer is declared as:

```
const double * __restrict__ u1
```

This functions correctly even though the data pointers are swapped around at each iteration in time. Unlike the L1 cache configuration, this feature does not always provide efficiency gains. Some kernels, such as the first 3D tiling method, benefited from using this cache, whilst others did not.

5. TESTING PROCEDURE

A standard test simulation was used to benchmark both the latest Tesla K20 card, as well as the previous generation Tesla C2050. Each of the six kernel methods was tested on both systems. The simulation models a $3.4 \times 4.0 \times 2.8 = 38\text{m}^3$ space. At a sample rate of 44.1kHz this requires data grids of size: $256 \times 296 \times 212 = 16,064,512$ points.

A raised cosine impulse was used as the input to the model, injected as a soft source at a given grid point. The simulation was computed for 44,100 samples, and using double precision floating-point arithmetic. All codes were compiled using CUDA version 5.0, and for compute architectures 2.0 or 3.5 as appropriate for the card. For reference, table 1 shows the key features of both the K20 and C2050 graphics cards.

Description	C2050	K20
Compute capability	2.0	3.5
CUDA cores	448	2,496
Clock speed	1.15 GHz	706 MHz
Memory bandwidth	144 GB/sec	208 GB/sec
Peak double precision	515 Gflops	1.17 Tflops

Table 1. GPU card specifications.

The difference in the hardware architectures is clear, as the Kepler card has five times as many core processors as the Fermi card, but running at a lower clock rate. Whilst the peak double precision performance is twice as high, the memory bandwidth is only 44% greater.

6. RESULTS

Tables 2 and 3 show the resulting computation times for each of the six kernels methods, firstly for the C2050 card, and then the K20. The timing points used were defined directly before the main time iteration loop, and directly after the loop, having performed a *cudaDeviceSynchronize*. The ratio figure is the percentage time relative to the 3D tiling base case.

Kernel Method	Time (s)	Ratio
3D Tiling	227.1	-
3D Tiling shared	340.8	150.1%
3D Tiling ext shared	272.7	120.1%
2D Slicing	300.8	132.5%
2D Slicing shared	334.3	147.2%
2D Slicing ext shared	227.6	100.2%

Table 2. Computation times for C2050 Fermi card.

Kernel Method	Time (s)	Ratio
3D Tiling	156.9	-
3D Tiling shared	218.7	139.4%
3D Tiling ext shared	198.1	126.3%
2D Slicing	183.3	116.8%
2D Slicing shared	164.6	104.9%
2D Slicing ext shared	150.8	96.1%

Table 3. Computation times for K20 Kepler card.

Table 4 shows a comparison of the timing data for each card, and the relative speedup achieved by the K20 card over the C2050 card.

Method	C2050(s)	K20(s)	Speedup
3D Tiling	227.1	156.9	x1.45
3D Tiling shared	340.8	218.7	x1.56
3D Tiling ext shared	272.7	198.1	x1.38
2D Slicing	300.8	183.3	x1.64
2D Slicing shared	334.3	164.6	x2.03
2D Slicing ext shared	227.6	150.8	x1.51

Table 4. Computation times (seconds) and speedup for K20 card compared to C2050 card.

Starting with the C2050 Fermi card data, the most notable result is that the basic 3D tiling kernel is just as efficient as the best shared memory method, the extended 2D slicing. Despite the major deduction in data reading from global memory, no performance benefit is seen. Indeed, all of the other shared memory kernels delivered worse results, some by up to 50%. The extended shared memory approach outperforms the standard shared memory version in both the 3D and 2D cases.

For the K20 Kepler card, the basic 3D tiling kernel rates second in terms of efficiency, but is only slightly behind the extended 2D slicing by some 4%. The method of using shared memory in the 2D case shows only minor variation. Comparing the two cards, the headline result is a x1.5 speedup for the fastest method running on the K20 card, with a computation time of 150.8 seconds.

This also improves on previously reported test data [7], in which a headline time of 184 seconds was shown using a version of a 2D slicing kernel with shared memory, and running an identical simulation but using the Nvidia GeForce GTX 480 card. Whilst the GeForce cards are designed for the gaming market and do not have same level of double precision support, they do have comparable, or in some cases better, memory bandwidth. For example, the GTX 480 has a bandwidth of 177.4 GB/sec which is greater than the C2050 Fermi card.

7. LARGE-SCALE ROOM MODELS

The K20 graphics card has 5Gb of global memory available, allowing large-scale room models to be simulated. The model detailed above can be extended from a grid size of 16 million up to 310 million points in each of the two grids. At 44.1kHz this gives a volume of 756m³, and a computation time of 47 minutes per second of output at double precision.

The inclusion of useable boundary conditions, and effects such as viscosity, increase the computation time. The latter effect also requires the use of three data grids rather than the two used here, and so reduces the maximum available volume to around 500m³. The use of single precision floating-point arithmetic doubles the maximum simulation space, or produces efficiency gains, although this can lead to stability issues for the boundary conditions when running at the Courant limit.

8. CONCLUSIONS

Six different approaches to the kernel design for the basic 3D FDTD scheme were optimised and benchmarked on the Tesla K20 card. What was once considered a ‘naive’ approach of simply reading directly from global memory now produces efficient kernels, both in the case of the Fermi and Kepler architectures. Making use of cache optimisations allows these basic codes to perform as well as the more complex shared memory versions.

Ultimately, these forms of finite difference schemes are always memory bandwidth limited. The ability to compute double precision floating-point arithmetic, and the amount of parallelisation, is secondary to the speed with which data can be moved around. Further comparisons can be made by benchmarking the latest GeForce Kepler cards, such as the GTX Titan, which has even greater memory bandwidth than the K20 tested here. The simultaneous use of multiple GPU cards is an effective approach to achieving scalable efficiency gains.

Acknowledgments

This work is supported by the European Research Council under Grant StG-2011-279068-NESS.

9. REFERENCES

- [1] L. Savioja, “Real-time 3D finite-difference time-domain simulations of low and mid-frequency room acoustics,” in *Proceedings of 13th Int. Conf on Digital Audio Effects*. Austria, Sept, 2010.
- [2] C. Webb and S. Bilbao, “Computing room acoustics with CUDA - 3D FDTD schemes with boundary losses and viscosity,” in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011.
- [3] L. Savioja, D. Manocha, and M. Lin, “Use of GPUs in room acoustic modeling and auralization,” in *Proc. Int. Symposium on Room Acoustics*, Melbourne, Australia, Aug 2010.
- [4] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*. Pacific Grove, California: Wadsworth and Brooks/Cole Advanced Books and Software, 1989.
- [5] Nvidia, “Cuda C programming guide,” *CUDA toolkit documentation*. [Online]. [Cited: 24th Mar 2013.] <http://docs.nvidia.com/cuda/>, 2012.
- [6] P. Micikevicius, “3D finite difference computation on GPUs using CUDA,” in *Proceedings of 2nd Workshop on General Purpose Processing on Graphics Processing Units*, New York, NY, USA, 2009, pp. 79–84.
- [7] J. Lopez, D. Carnicero, N. Ferrando, and J. Escolano, “Parallelization of the finite-difference time-domain method for room acoustics modelling based on CUDA,” *Mathematical and Computer Modelling*, vol. 57, no. 78, pp. 1822 – 1831, 2012.

Figure 7. 2D Slicing with shared memory kernel.

```

1  --global-- void UpDate(double *u, double *ul, double L2)
2  {
3      __shared__ double uS1[BxL][ByL];
4
5      int tdx = threadIdx.x;
6      int tdy = threadIdx.y;
7
8      // Get 3D position
9      int X = blockIdx.x * BxL + tdx;
10     int Y = blockIdx.y * ByL + tdy;
11     int Z, cp;
12
13     // Initial variables
14     double ulcpm = 0.0;
15     double ulcp = ul[area+(Y*Nx+X)];
16     double ulcpp, sum;
17
18     for (Z=1; Z<(Nz-1); Z++){
19
20         // Get linear position
21         cp = Z*area+(Y*Nx+X);
22         ulcpp = ul[cp+area];
23         // load shared
24         uS1[tdx][tdy] = ulcp;
25         __syncthreads();
26
27         if ( (X>0) && (X<(Nx-1)) && (Y>0) && (Y<(Ny-1)) ){
28
29             sum = 0.0;
30             if (tdx==0) sum+= ul[cp-1];
31             else sum+= uS1[tdx-1][tdy];
32
33             if (tdx==BxL-1) sum+= ul[cp+1];
34             else sum+= uS1[tdx+1][tdy];
35
36             if (tdy==0) sum+= ul[cp-Nx];
37             else sum+= uS1[tdx][tdy-1];
38
39             if (tdy==ByL-1) sum+= ul[cp+Nx];
40             else sum+= uS1[tdx][tdy+1];
41
42             u[cp] = L2*(sum+ulcpm+ulcpp) - u[cp];
43
44             ulcpm = ulcp;
45             ulcp = ulcpp;
46             __syncthreads();
47         }
48     }
49 }

```

Figure 8. 2D Slicing with extended shared memory kernel.

```

1  --global-- void UpDate(double *u, double *ul, double L2)
2  {
3      __shared__ double uS1[BxL+2][ByL+2];
4
5      int tdx = threadIdx.x;
6      int tdy = threadIdx.y;
7
8      int X = blockIdx.x * BxL + tdx;
9      int Y = blockIdx.y * ByL + tdy;
10
11     int Z, cp;
12     double ulcpm = 0.0;
13     double ulcp = ul[area+(Y*Nx+X)];
14     double ulcpp;
15     tdx++; tdy++;
16
17     for (Z=1; Z<(Nz-1); Z++){
18
19         cp = Z*area+(Y*Nx+X);
20         ulcpp = ul[cp+area];
21         // load shared
22         uS1[tdx][tdy] = ulcp;
23
24         if ( (tdy==1) && !(Y==0) ){
25             uS1[tdx][tdy-1] = ul[cp-Nx];
26         }
27         if ( (tdy==ByL) && !(Y==(Ny-1)) ){
28             uS1[tdx][tdy+1] = ul[cp+Nx];
29         }
30         if ( (tdx==1) && !(X==0) ){
31             uS1[tdx-1][tdy] = ul[cp-1];
32         }
33         if ( (tdx==BxL) && !(X==(Nx-1)) ){
34             uS1[tdx+1][tdy] = ul[cp+1];
35         }
36         __syncthreads();
37
38         if ( (X>0) && (X<(Nx-1)) && (Y>0) && (Y<(Ny-1)) ){
39
40             u[cp] = L2*(uS1[tdx-1][tdy]+uS1[tdx+1][tdy]
41                     +uS1[tdx][tdy-1]+uS1[tdx][tdy+1]
42                     +ulcpm+ulcpp) - u[cp];
43
44             ulcpm = ulcp;
45             ulcp = ulcpp;
46             __syncthreads();
47         }
48     }
49 }

```


Music acoustics education

Music Acoustics Education at the Erich Thienhaus Institute in Detmold

Malte Kob

Erich Thienhaus Institute — University of Music Detmold

kob@hfm-detmold.de

ABSTRACT

The education in music acoustics has been one of the main objectives of the founder of the first institute for Tonmeister education, Erich Thienhaus. Since 1949 the Erich Thienhaus Institute (ETI) has offered a unique education that is dedicated to the technical and musical aspects of music production. Within the Bologna process the education at ETI has been structured into a bachelor of music, three flavours of master programs and one Ph.D. course program. Two education programs are dedicated to music acoustics: one master of science and the Ph.D. program. The presentation presents the structure and content of these programs and potential links to related course programs inside and outside Europe.

1. INTRODUCTION

The University of Music Detmold with its Erich Thienhaus Institute (ETI) for Tonmeister education offers a broad spectrum of education and research in music acoustics that combines artistic and scientific education on an equally high level. Musicians and lecturers join their expertise in their fields and offer a broad variety of courses and teaching. The variety of equipment at ETI is remarkable: up-to-date recording and sound processing equipment as well as acoustic measuring equipment is available and further developed in Detmold. The institute features an electronics lab, a reverberation chamber – that is also used in its original meaning –, an anechoic chamber, and studios for performing music and sound reproduction including two wave field synthesis systems.

Education in music acoustics has been one of the main objectives of the founders of the first institute for Tonmeister education, Erich Thienhaus. Being a physicist and musician Erich Thienhaus established the first education that was dedicated to connect music, the art of recording and acoustics.

In the course of the education of music acoustics in Detmold significant contributions were achieved by docents from the laboratory of music acoustics of the PTB Braunschweig. Best known to sound engineers and Tonmeisters is Jürgen Meyer, who taught music acoustics from 1968 to 2004 in Detmold in his unique way of conveying scientific observations and investigation results to the mind and ears

of musicians. His book “Acoustics and the Performance of Music — Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instruments Makers” [1] is a convincing example of this bi-lingual approach, but also any of his presentations that often are enriched by recorded or live music demonstrations. For Tonmeisters the graphic representation of radiation patterns of orchestra instruments is a valuable source of knowledge for the placement of spot microphones and the reference for directivity research of music instruments.

At PTB many more colleagues actively worked in the field of music acoustics, among them Klaus Wogram, who worked intensively on the acoustics of brass instruments, and Ingolf Bork, who graduated as Tonmeister in Detmold.

Another colleague who contributed to music acoustics research and Tonmeister education in Detmold is Thomas Görne. His book “Tontechnik” [2] has been established among Tonmeisters as an introductory book to all issues related to basics of engineering sciences, including music acoustics and technical aspects of recording.

With the winter semester 2008/2009 the former diploma program for Tonmeister education was transformed into a bachelor of music program with optional subsequent master course programs. This profound change had and still has great impact on music acoustics education in Detmold and for music acoustics in Germany. Details of the course programs and the chances of this new education approach are content of this article.

2. BACHELOR OF MUSIC PROGRAM: TONMEISTER EDUCATION

The bachelor of music “music transmission” is the basic education for Tonmeister education in Detmold. The B.Mu. program aims to develop and steer the musical, scientific and technical talents of students in order to prepare them for the tasks entailed in being artistic recording directors of both art and popular music productions and audio-visual media productions, including PA systems.

It consists of three major sets of course modules and has a scheduled length of 8 semesters:

- musical: theoretical & practical music lessons
- scientific: mathematics, electrical engineering, digital technology, acoustics and other subjects tailored to the course
- practical: imparting of basic knowledge and gaining of practical experience with studio equipment for all types of recordings and productions

Detailed information on the bachelor program is given on the web page of the course program [3].



Figure 1. Study of music score in the recording producer program

The Bachelor program is extended by three master programs with different flavors. Two of them are masters of music and taught in german: “recording producer / Musikregie” and “Sound director / Klangregie”.

The M.Mu. recording producer program continues the artistic aspects of the bachelor education and complements it with a two-semester course program.

The M.Mu. sound director program is scheduled for four semesters. It addresses the problem that in contemporary music many pieces need a sound direction, that is, they have to be performed with electro-acoustic means, live electronics and / or computer-based control. Thereby it should be self-evident that the technical equipment is operated by a musically competent professional who contributes creative ideas using his score knowledge beside technical planning. The master ‘sound director’ wants to provide the musical and technical expertise needed for planning and design of projects in contemporary music. This includes composition-based sound reinforcement, creating and playing of live-electronics and the control and programming of the required processes. The goal is the realization of a concert with works of contemporary music using electro-acoustic means.

The third is a master of science that is mostly taught in english. This course program is presented in detail in the next section.

3. MASTER OF SCIENCE PROGRAM: MUSIC ACOUSTICS

The master of science program is the first scientific master program at a music school in Germany. The course program is scheduled for four semesters and offers a look behind the scene: How exactly is the sound of music created? What is the physical principle of music sound generation, and tries to answer the question what exactly musicians are doing when they produce music? Questions such as: “How does a guitar body vibrate in slow motion, and: how can instrument makers control this? Where does a bassoon radiate its sound, and: what does this implicate for miking it?” will be addressed in the course program on various levels.

The M.Sc. program is intended for bachelor of music / Tonmeister students or students with a bachelor of science degree in a technical discipline. Since music acoustics is an interdisciplinary field of research with focus on the acoustic aspects of music performance and perception, the following topics are addressed in the program:

- Function of musical instruments
- Physics of sound generation
- Measurement technique for music instruments
- Interaction of musician, instrument, room and listener
- Perception of timbre and music
- Analysis of timbre and structure of music sounds
- Synthesis of musical sounds and physical modelling

The courses are organized as lectures, exercises, practical or individual work.

The aim of the course program is to offer to students with interest in music and acoustics a specialized program that builds up and consolidates scientific methodology, introduces fields of research at the intersections of music – instrument – musician – room – listener and prepares the students to actively participate in fundamental and applied research projects. Graduate students shall be prepared to continue own research within a subsequent Ph.D. program.

Consequently, apart from the courses, the students’ participation at various levels is offered in research and development cooperation projects.

Graduates of the program have job opportunities in many fields at the intersection of music and science e. g:

- Musical instruments making
- Acoustic consulting
- Sound design
- Research and development in complex dynamic systems

The modules are structured according to the intended development of skills of the master students. The order is defined in the study plan [4], the complete description of the courses is available in english and german [5]. The examination regulations are available online [6].

3.1 1st year modules

The modules in the first year intend to build a solid basis for scientific work in music acoustics. The following modules are scheduled:

1. Understanding scientific work
2. Fundamentals of music acoustics
3. Applied acoustics
4. Analysis and modelling of musical instruments
5. Practical music acoustics I

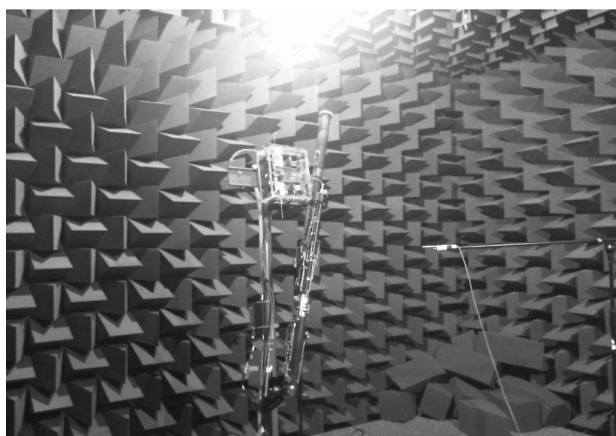


Figure 2. Directivity measurement of a bassoon in an anechoic chamber [7]

Elective course modules

In addition to the compulsory courses a number of 6 ETCS per semester can be chosen from a pool of courses that are offered for all master students at HfM Detmold.

6. The choice of courses can be made from these elective course modules:
 - chamber music
 - aural training
 - musicology
 - pedagogy
 - scientific project
 - freely designed module

Some of these modules require passing of a dedicated aptitude test before a student is admitted.

3.2 2nd year modules

The second year intends to foster the consolidation and application of the knowledge acquired until then. The elective course modules are continued, and these additional modules shall be followed:

7. Practical music acoustics II: within this module the theoretical and musical skills shall be applied to the construction and optimization of an own musical instrument. This might be acoustic or electronic, depending on the skills and interests of the student
8. Presentation of scientific work: the opportunity to present a scientific paper on a conference is offered.

Final project

9. Master thesis: the final examination of the M.Sc. program shall prove the student's ability to scientifically document research work. 23 ECTS are allocated for this final module.

The structure of the program encourages students to spend one semester at another institution which offers one or more modules that are scheduled during this semester. This option might be especially used in the third or fourth semester when the applied research is taught.

Erasmus exchange programs are established with various partners and will be used to support the mobility of students within this course program.

4. PH.D. COURSE PROGRAM

The successful graduation in the M.Sc. program prepares students with strong interest in scientific work for further studies in the frame of the music acoustics Ph.D. course program. The program is scheduled for six semesters, divided in two sections.

4.1 Orientation/education phase

The first section of the Ph.D. course program lasts two semesters and shall consolidate music acoustics knowledge and the aptitude in application of scientific methods. The section contains seminars and lectures as well as colloquia that can be followed in Detmold or – up to 16 ECTS – at another institution. A written exposé of the intended research topic – in the style of a licentiate in Scandinavia – completes this section and documents the status of the student's introduction into the field of research.

This phase is examined in an oral discussion. The successful examination grants access to the second phase of the Ph.D. program, the research phase.

4.2 Research phase

This phase contains the actual work on the Ph.D. topic, accompanied by two seminars and a colloquium that supports the student in her or his research.

The second phase takes four semesters and ends with writing of the Ph.D. thesis. The successful oral examination concludes the course program and will grant the degree "Ph.D." or "Dr.-Ing." in music acoustics.

As in the M.Sc. program the mobility of students for participation in summer/winter schools and temporal studies at partner universities that offer similar teaching programs is intended. Support for such activities and exchanges is

offered through training programs of the European Acoustical Society (EAA) and dedicated projects as detailed in the next section.

5. COOPERATION

The music acoustics group in Detmold collaborates with scientists inside and outside Germany, and performs research and applied projects in various fields of music acoustics.

A strong local cooperation exists with the centre of musicians' health. Current projects address the protection of orchestra musicians during rehearsals by protective screens. This project was driven by previous investigations of hearing capabilities in various groups of musicians that exhibit significant differences of hearing at high frequencies.

A recently established EU-funded project "BATWOMAN — Basic Acoustics Training- & Workprogram On Methodologies for Acoustics - Network" with 19 partner institutions will further increase the cooperation and mobility of education and training of students and researchers in music acoustics. The exchange of Ph.D. students among the

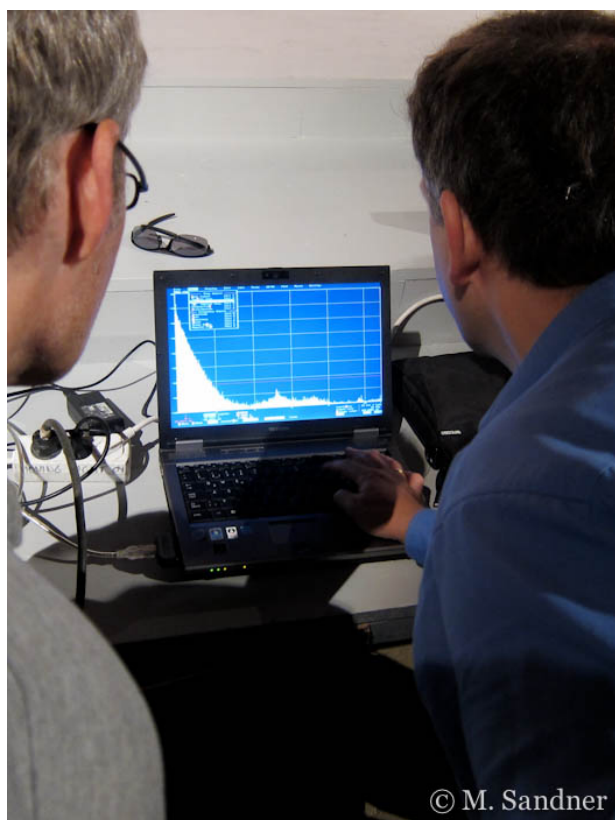


Figure 3. Measurement of room acoustic properties in Retretti, Finland

partners and the organization of training events at various locations will help to increase the impact of education and the application of methodologies of music acoustics within Europe.

6. CONCLUSIONS

The new structure of music acoustics education in Detmold complements the traditional high-level education of Tonmeister students with a scientific course program. The current master program is designed for bachelor students with Tonmeister, sound engineer or information science education who wish to qualify for master level. The Ph.D. program offers the opportunity to continue with education and project-related studies in the field of music acoustics.

The international character and course language of the courses shall help to bridge the music acoustics institutes inside and outside Europe. Students are encouraged to profit from exchange options between locations that enrich the studies with their individual focus in research and teaching. Erasmus and dedicated EU projects are used to increase mobility and networking of students and researchers in music acoustics.

Acknowledgments

The author wishes to thank the administration of the University of Music Detmold for the confidence and the endurance in putting up the first course concept for music acoustic education on university level at a music school in Germany. My colleagues and the first students of the course programs are acknowledged for their patience and support of the concept.

7. REFERENCES

- [1] J. Meyer, *Acoustics and the Performance of Music – Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instruments Makers*. Springer, 2009.
- [2] T. Görne, *Tontechnik*, 3rd ed. Carl Hanser Verlag GmbH, 2011.
- [3] M. Sandner, M. Kob, and M. Schubert. (2013) Course program description B.Mu. Music Transmission. ETI, HfM Detmold. [Online]. Available: <http://www.eti.hfm-detmold.de/studium-en/bachelor>
- [4] M. Kob. (2013) Study plan M.Sc. Music Acoustics. ETI, HfM Detmold. [Online]. Available: http://www.eti.hfm-detmold.de/studium/studiumpdf/studyplan_msc_musicacoustics_120423_v11e.pdf
- [5] —. (2013) Module description M.Sc. Music Acoustics. ETI, HfM Detmold. [Online]. Available: http://www.eti.hfm-detmold.de/studium/studiumpdf/module_msc_musicacoustic_120423_v6.pdf
- [6] —. (2013) Examination regulations M.Sc. Music Acoustics. ETI, HfM Detmold. [Online]. Available: http://www.eti.hfm-detmold.de/studium/studiumpdf/po_msc_musicacoustics_120423_e_v1.pdf
- [7] T. Grothe and M. Kob, "Investigation of bassoon directivity," in *Proceedings of SMAC 2013*, 2013.

THE MUSICAL ACOUSTICS RESEARCH LIBRARY (MARL): FULLY DIGITAL & ONLINE

Gary P. Scavone

Computational Acoustic Modeling Laboratory
CIRMMT, Music Technology
McGill University
Montreal, QC, Canada
gary@music.mcgill.ca

Jerry McBride

Music Library and Archive of Recorded Sound
Stanford University
Stanford, CA, USA
jlmcbri@stanford.edu

ABSTRACT

The Musical Acoustics Research Library (MARL) is a collection of research materials assembled by distinguished groups or individuals in the field of musical acoustics research. MARL was established at the Center for Computer Research in Music and Acoustics, Stanford University in the mid-1990s. A catalogue of the MARL contents was made available online and individual items were digitized and linked to the site upon request when resources allowed. In 2009, an agreement was reached between the various MARL representatives and the Stanford University Library for the transfer and digitization of the entire collection. The new MARL website is now officially online and its contents are freely available to the musical acoustics community in digital form.

1. BACKGROUND

The Musical Acoustics Research Library (MARL) is a collection of research materials from the Catgut Acoustical Society (CAS), Arthur Benade, John Coltman, and John Backus. The collection was established during the 1990s at the Center for Computer Research in Music and Acoustics (CCRMA), Stanford University [1].

MARL's origins date back to 1992, when CAS agreed to transfer to CCRMA an extensive set of files that had been compiled by Carleen Hutchins and other CAS members concerning the research of such people as Louis Condamine, Robert Fryxell, and John Shelleng, to name a few. CCRMA agreed to actively promote and find ways of making available the file contents. Gary Scavone, a PhD student at CCRMA at the time, volunteered to help coordinate these efforts, with faculty support by Max V. Mathews and the director of CCRMA, Chris Chafe. Scavone initiated discussions to augment the CAS files with the personal archives of Arthur Benade and John Backus, world leaders in the study of wind instrument acoustics. Available research files and some experimental equipment of John Backus were acquired in 1995 and an agreement was reached with Virginia Benade in 1997 to establish the Arthur

H. Benade Archive. In 1998, John Coltman agreed to contribute his personal research files to the collection. The Musical Acoustics Research Library was established to provide a single point of reference for these various archives.

CCRMA supported the creation of a World Wide Web site to help publicize and make available the MARL contents to members of the musical acoustics research community. While descriptions of the file contents were made available online, however, there was insufficient funding available to allow a mass digitization of all the files. Rather, individual files were scanned and made into PDF or DjVu documents as they were requested from the public.

2. DIGITIZATION PROJECT

In 2003, Gary Scavone took a faculty position at McGill University, Montreal, Canada. Arrangements were made to allow the scanning-on-demand modus operandi to continue, with Scavone being able to maintain the MARL website remotely. However, it became difficult to address inquiries regarding the file contents and there tended to be long delays in getting requested documents scanned. Thus, discussions were initiated in 2005 with the Stanford University Libraries (SUL), and in particular Jerry McBride, head of the Stanford Music Library, about transferring the MARL files to the SUL and undertaking a complete digitization of the files. From 2006, McBride spearheaded a number of initiatives to help achieve these goals. First, he succeeded in getting the Stanford Library administration to formally agree to the project. He then undertook discussions with and gained consent from the various MARL sub-collection representatives for the transfer of the files to the SUL. As well, he was able to gain financial support for the digitization project from a variety of sources, including the Stanford Initiative for the Arts, CCRMA, the Violin Society of America (VSA)¹, and the Stanford University Libraries.

The funding from the VSA was made possible by a fundraising effort begun in the 1990s by the CAS to support the MARL initiative. The CAS Forum Board agreed to transfer to Stanford these funds in support of the digitization effort.

Work on the MARL digitization project officially began in late 2009. The most time-consuming component of this

Copyright: ©2013 Gary P. Scavone et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ The Catgut Acoustical Society had officially merged with the VSA as the CAS Forum in 2004.

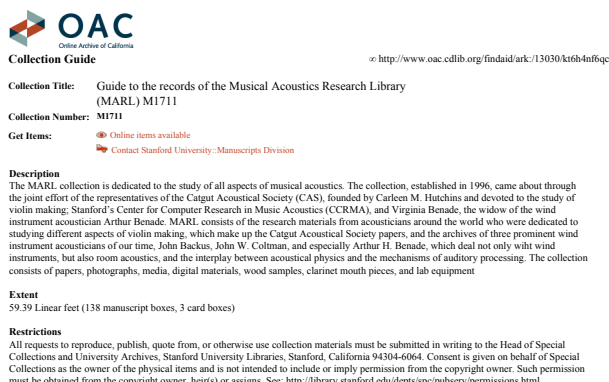


Figure 1. Online MARL finding aid.

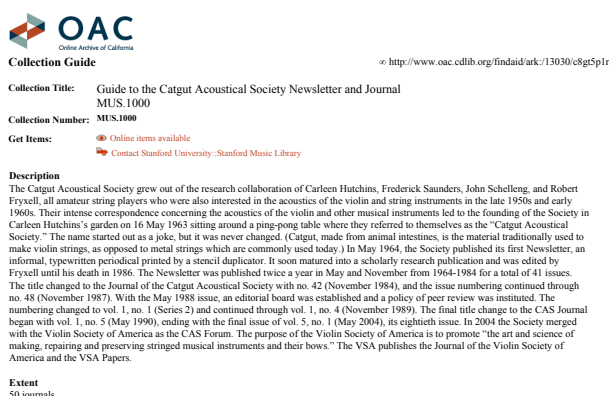


Figure 2. Online CAS Newsletter and Journal finding aid.

project was not the scanning itself but rather the task of attempting to contact authors of the various files to obtain copyright consent. The MARL finding aid² (see Figure 1) was posted in October 2011 (without digitized content). Digitized versions of most of the MARL files were added to the site in early April 2013. A finding aid with digitized versions of the complete set of the Catgut Acoustical Society Newsletters and Journals was made available in late April 2013³ (see Figure 2).

Some of the MARL files are not online because of copyright or author permission issues. Parties interested in a listing that lacks a digital link may contact Jerry McBride (jlmcbri@stanford.edu) to obtain further information.

3. FUTURE DEVELOPMENTS

The Musical Acoustics Research Library was established with the expectation that it would continue to grow over time via the addition of more sub-collections and Stanford is quite interested in adding to these archival collections of acoustical research materials. Digitization of modern archival collections is complex, and the musical acoustics research community might consider future discussions on

ways to assist libraries with securing copyright permissions for the documents and obtaining funding for cataloging and digitization to help ensure that research archives of our colleagues are preserved and made available to the public.

Carleen Hutchins passed away on 7 August 2009. Thirty-three boxes of research files from Carleen's estate were sent to Joe McNalley of the Hutchins Consort, which were subsequently gifted to Stanford in 2010. A finding aid⁴ for these materials has been posted, though funds do not currently exist to digitize the files.

Acknowledgments

Over the past twenty years, many people have contributed to MARL in various ways, all with the hope that its contents would one day be fully digitized and freely available. We gratefully acknowledge the substantial efforts of Virginia Benade, Joan Miller, Carleen Hutchins, John Coltman, Jay VandeKopple, Patricia Lamkie, Max Mathews, Chris Chafe, Julius Smith, Jonathan Berger, Mauk Hudig, and Fan Tao. We note with regret the passing of Carleen Hutchins (2009), John Coltman (2010), Joan Miller (2011) and Max Mathews (2011) ... their support and guidance was crucial to the success of this project.

Finally, we wish to acknowledge the organizations that helped fund the digitization project, including the Stanford Initiative for the Arts, CCRMA, the VSA-CAS, and the Stanford University Libraries.

4. REFERENCES

- [1] G. P. Scavone, "The musical acoustics research library," *The Catgut Acoustical Society Journal*, vol. 3, no. 6, 1998, (Series II).

²<http://www.oac.cdlib.org/findaid/ark:/13030/kt6h4nf6qc/>

³<http://www.oac.cdlib.org/findaid/ark:/13030/c8gt5pl1r/>

⁴<http://www.oac.cdlib.org/findaid/ark:/13030/kt067nf29d/>

Activities for a Course of Physics of Bowed Instruments

Jesús A. Torres

Academia de Ciencias (Acústica), Escuela de Laudería, Instituto Nacional de Bellas Artes y Literatura,
Hidalgo #20, Centro Histórico, Querétaro, México, CP 76000.
jesusalejandrott@yahoo.com.mx

ABSTRACT

Simple experiments, simulations and theoretical procedures are proposed as complement of an one-year course of acoustics, mainly focused in physics of the violin. Covered topics include string and body vibrations, the bridge, sound radiation, the wolf tone, and tonal quality. Majorly, explained procedures do not require expensive equipment or strong background in science. Therefore, the confluence of physics, music and fun allows interacting with scientific applications in musical instruments which could result a daunting experience by traditional methods.

1. INTRODUCTION

Undoubtedly, activities in a course of physics of stringed instruments for luthiers -mainly demonstrations- are a powerful (and maybe necessary) tool. Several suggested works covering specific related topics can be found, e.g. [1–4]; however, sorting them and proposing others in a systematic course could be useful.

Some years ago, Shonle [5] exposed ideas and experiences creating and imparting a course about physics of music. Also, he included interesting references supporting this kind of course; however, the classic book by Fletcher and Rossing [6] had not been published yet. Even though other remarkable books are nowadays available for the specific case of the violin (e.g. [7]), reference [6] also includes basic concepts of acoustics. Then, the latter emerges as an excellent option of a textbook to impart a course for undergraduate violin makers; although no school activities used to be suggested in the body.

In the *Escuela de Laudería*, the author imparts an one-year course of Acoustics deliberately focused in physics of bowed instrument. In this paper, some features and suggested activities are exposed for similar courses. Even self-taught persons desiring learning about physics of the violin can perform almost every activity here proposed, without strong scientific background as pre-requisite.

2. THE COURSE

Before starting the course, a preliminary reinforcement of some wave-related concepts at high-school level is desirable:

sine and cosine functions, frequency, period, wavelength; as well as sound speed, timber, and logarithmic scale. In fact, the admission process to the *Escuela de Laudería* includes an exam about these terms.

For the Acoustics course, scheduled for third year students, the Fletcher and Rossing book [6] is taken as textbook and hereafter, it only will be referred as “the book”. However, occasionally the Cremer book [7] is also mentioned as well as other references. It is strongly encouraged availability of the book while reading the present work and, if it is feasible, every referenced paper.

Taking account that students have not any other physics related course previously (although basics mathematics does), starting with physics directly applied to the violin could be seem inadequate. However, the worst enemy of the course has been explaining a concept under the promise that later it will be useful; instead, reaching new topics following the book and then explaining them has been, by far, more successful.

Note that almost every activity here proposed involves skills unknown by students. However, the aim is not obtaining a perfect result, but introducing the physical (and even musical) meaning of theoretical and experimental applied tools.

A pair of decades has passed from the first edition of the book and several, even hundreds of years of some of its referenced violin experiments. Nowadays, trying to replicate some of them in a classroom is feasible (and fun). Musical instruments, a computer with free software [8–13] and low-cost devices will emerge as the main tools; although analyzers and transducers are also desirable.

3. STRING VIBRATIONS

3.1 String Harmonics.

The first experiment of the chapter references to Mersenne hearing string overtones. A guitar is enough to obtain exactly his same result reported about that he heard at least four overtones. Several concepts of string vibrations are introduced for the class as natural frequencies, harmonic content linked to plucking point, and string modes (which a strobe lamp is useful). Also, frequency domain is explored recording the guitar sound through a microphone and processing the spectrum in a computer [9].

3.2 Motion of Bowed Strings.

Nowadays, the motion of a bowed string can be explored through several alternatives to the microscope used by Helmholtz,

e.g. using a high speed camera or as [2] proposed. However, we use an even simpler way. A rubber band of 13 mm thick is mounted between a polley and a motor as in [2], in order to be used as bow, but the Helmholtz motion is carefully obtained in elastic thread stretched using the hands. A strobe light reveals the expected motion.

Strongly recommended solving the first equation of the chapter *Bowed Instruments* in the book about the Helmholtz motion, proposed by Ramman: the teacher plots the spatial graph for one or two values of n and students for higher values, each one a different. Then, results of the whole classroom are mixed at the end by teacher. Solving this kind of equation may seem inadequate, but related concepts as sine waves and superposition will be unavoidable in the rest of the course and it is a good opportunity to handle them.

3.3 Bowing Speed.

The minimum bow speed to obtain a steady bowed motion in the string is obtained from a violin. The hair bow length is measured and divided by the time of action, playing as low as possible. The order of the referenced data of 0.04 m/s obtained by Askenfelt [14] can be reached.

Moreover, some bowed notes are recorded using a microphone in a computer [8]. From time analysis, students explore the time required to obtain a steady vibration of the bowed string. Again, reported values by Pickering [15] are feasible to reach (≈ 100 ms).

The wave concept of period is explained to detect the second harmonic in the recorded time signal; students must verify if this second harmonic appears before that the first as it is expected from the explanations in the Pickering work [15].

3.4 Stiffness and Longitudinal Motion.

Three sounds of a violin are recorded in a computer using a microphone [8]. In one of them, a sustained note of G3 is bowed to find some jittered periods in time analysis. Cent unit is explained to calculate the difference of periods between two different wave forms. Some wave forms can differ 30 cents as in the work by McIntyre and Woodhouse [16].

The other two sounds are not typically obtained using the bow being perpendicular to the string; instead, the bowing direction is parallel to string length. Therefore, longitudinal string waves will be driven; the goal will be detecting their frequencies for G string and D string, through spectrum analysis (≈ 1350 Hz and ≈ 2700 Hz respectively [17]).

4. VIOLIN BODY VIBRATIONS

4.1 Experimental and Simulated Mode Shapes.

Each student creates a rectangular membrane, inserting and removing two rods united by two threads in a soap solution. Shaking the membrane at very low frequencies, some mode shapes are naked eye detected. More details were

previously reported by the author [3]. The experiment allows an intuitive contact with concepts as mode shapes linked with modal frequencies; also, it results very fun.

Moreover, lower modes of the membrane are quickly simulated by students, through a Finite Element code. Several alternatives are available, although free software is enough complementing this activity; e.g. if LISA [13] is used, the process requires the next four steps:

1. In *Model*: choosing Transverse Vibration of Membrane in the 2D Modal Vibration options, increasing the number of modes to calculate, and creating an isotropic membrane as material.
2. Clicking *quadrilateral* command at right of the screen and customizing a refined mesh in *Elements*.
3. In *Constraints*: adding the zero displacement *property*, once that nodes in boundaries have been simultaneously selected.
4. In *File*: clicking *solve*.

4.2 The Concept of Response

Trying to store in the long-term memory of students about responses must involve output signal related with input signal use to result a challenge. Typically, they consider that an output signal could be enough analyzing the violin behavior. A direct contact with the importance of input/output relationship is suggested, through an evaluation of which open violin string sounds louder. Methodology is simple but it is recommended that students infer the experiment considering a sound meter as tool.

Each string is bowed and its sound pressure level is taken. For comparative purposes, set-up must imply the same position both sound meter and violin; but also, bowing force and speed for each string must be repeatable (trying to play as loud as possible); otherwise, measurements has no sense. D4 string resulted the loudest, surely by the well-known f-hole amplification.

4.3 Coupled Oscillators.

Analyzed system in 4.6 *Forced Vibration of a Two-Mass System Coupled* of the book is constructed using two different balls and two plastic springs: this coupled system is useful explaining the low frequency range of stringed instruments as guitars and violins. For the first vibrating system, the heavier ball is attached to the stiffer spring (if both springs are equal then its total length can be constrained); it will represent the top plate. The second system is the lighter ball and the other spring, representing the vibrating air. The natural frequency of each vertical decoupled system is calculated counting its cycles during some seconds, obtaining that $f_2 < f_1$.

The second system is attached to the first in the heavier ball. This coupled system now will exhibit two resonances, in the mode of lower frequency \bar{f}_1 both systems move in phase and in opposite phase in the mode of higher frequency \bar{f}_2 . Analogously to the violin (and the guitar), measurements shows that $\bar{f}_1 < \bar{f}_2 < \bar{f}_1 < \bar{f}_2$. Finally, a hand

drawn response of the coupled system is plotted by each student.

4.4 Violin as Helmholtz Resonator.

Considering a violin as a neckless Helmholtz resonator with a large face, its natural frequency can be estimated using

$$f_2 = \frac{c}{2\pi} \sqrt{\frac{1.85a}{V}}, \quad (1)$$

where c is the speed of sound (330 m/s), V is the volume of the enclosed air and a is the radius of the end correction of a flanged tube for the effective length of neck (see *Helmholtz Resonator* in the book); f_2 remains as the uncoupled air resonance explained in experiment 4.3 because this oversimplified model implies rigid walls rather coupled system. Students must estimate V and a for a violin, although f_2 can be smaller even than the well-known \bar{f}_1 , maybe because the face of the resonator surrounding the hole is not large.

4.5 Mode Shapes in a Violin

Using a free modal viewer [10], commercial data of vibrations in two violins [18] are explored: one is the “Titian” (Stradivari 1734) and the other is the “Plowden” (Guarneri 1735). For each violin, students must identify four relevant modes (also called signature modes): A_0 , T_1 , C_3 and C_4 . Especial attention is required comparing the phase between top and back plates. The label of each mode must be reasonably justified because several modes are available.

4.6 Mobility of a Guarnerius Violin

Comparing the point mobility of a violin of the classroom against a Guarneri is desirable, once that data of the latter is available in the book, from reference [19]. Even though it is possible that required instrumentation is available, the fact that mounting settings demand an experienced researcher must be considered.

Before discussing details of the Guarneri response in class, an interesting exercise is feasible. Each student creates a handmade response trying to copy the Guarneri response of the book. Then, each response is evaluated by the class following the features remarked by Moral and Jansson [19]. This exercise exhibits the hard task of imitating a famous violin response even before involving real instruments, transducers and related experience.

4.7 Free plates

Chladni method applied in free violin plates is well documented and related works focused for luthiers are available, e.g. [1]. With this in mind, only particular instrumentation used for the Chladni method in the course will be briefly exposed.

The wave generator of *Visual Analyzer* [8] provides a sinusoidal signal towards an amplified loudspeaker, which drives an antinodal point of the horizontal plate (also softly supported in two nodal points). Desired resonant frequencies, usually mode 2 and 5 [1], are extracted from peaks

of the corresponding sound spectrum of the tapping plate. The same software [8] is capable to calculate sound spectra in real time, through recordings obtained by a single microphone.

Also, tap tones must be heard by students trying to discern the obtained frequencies for Chladni patterns. Finally, a judgment about quality of each tested top and back plate must be done according to the double octave tuning criteria.

5. THE BRIDGE

Students must calculate the first resonant frequency of the bridge measured by Hacklinger [20], using his reported bridge stiffness and effective mass through $f = 1/2\pi\sqrt{k/m}$. Also, a similar theoretical estimation of the frequency decrease of that resonance must be calculated. Using the specified mute mass added in the bridge by Hacklinger, calculated frequencies are very similar to his experimental data included in the book.

An experiment is performed using a specially designed violin bridge, four times bigger, because the short dimensions of the typical bridge difficult its experimental analysis. The big bridge rests over its feet and its upper part is impulsively knocked with a finger nail. Meanwhile, a microphone records its near-field sound to process a spectrum, which amplitude peaks related with plotted mode shapes of the bridge in the book must be identified.

In order to detect a frequency change of the peak of the first rotational mode shape of the bridge, the experiment is modified in two ways. Firstly, to decrease the frequency, a mass is attached in a hip near to one kidney wing of the bridge; and finally, wedges are softly inserted in the bridge slits to increase the frequency.

6. SOUND RADIATION

6.1 Directivity Patterns

Frequency dependence of the sound directivity and the influence of the spatial distribution of the sources are explored. Desktop loudspeakers are supplied of sinusoidal signals by a built-in wave generator [8]. Changes in the sound field are deliberately obtained through two methods: one is that students hear both loudspeakers emitting the same sound while the frequency is iteratively changed; for the other method, the frequency is fixed while the spatial distribution of the sources is changed.

Later, a high frequency sinusoidal signal (beyond 3 kHz) drives a violin in its bridge, through a modified loudspeaker analogously to [4]. A sound-meter is used for radial scan in the plane of the bridge (guided by an attached thread) of the sound radiated by the violin; a multipolar directivity pattern is expected. If a sound-meter is no available, scan using an ear could be enough obtaining qualitative data. Students must plot the measured directivity pattern.

6.2 Phase in a Transfer Function

The first time that students will interact with phase in a transfer function is in Subsection *Multipole Expansion of*

the *Sound Field* of the book. Although it refers to radiativity plots, mechanical systems will be useful for introductory explanations.

After basic explanations about magnitude and phase graphs, the displacement as a function of frequency of a simulation using three different mass-spring systems [12] is plotted by students. Since the oscillating piston driving remains constant, the function can be considered as transfer response. A magnitude graph is obtained of the displacement of each system for different frequencies. The frequency span starts in lower frequencies than the lowest resonance, while its highest frequency must be greater than the highest resonance. Fifteen samples will be required. Also, corresponding phase plot must be included per each system, considering the displacement of the piston, related with the displacement of each analyzed mass. Then, six graphs must be obtained.

6.3 Radiation at Higher Frequencies

Exploring a numerical model of a structure radiating near to a critical frequency resulted very useful explaining this topic [3], although certainly, this kind of tools requires a lot of programming. Alternative activities can be also scheduled, as simply seeing any video of a shock wave in an airplane. Two activities are proposed at following.

Consider that wavelength in air decreases as the frequency increases, and that of a natural bending wave decreases as the square root of frequency [7]. Through simple relationships using constants of proportionality (try $\lambda_A = 330/f$, $\lambda_{Bx} = 5/\sqrt{f}$ and $\lambda_{By} = 2.5/\sqrt{f}$) plotting three graphs of wavelength vs. frequency: one including data of sound waves; while the other two containing data of parallel and perpendicular bending waves to a wood plate grain. Then, finding a mode shape expecting radiation at critical frequency exploring commercial data [18].

Finally, the terms of monopoles, dipoles and quadrupoles are exposed using a simulated ripples tank [11]. Once that students have seen how that simulation works, they must sort from one to six sources to obtain radiation patterns of the two-plate system simulated by Cremer [7] (briefly explained in the Fletcher and Rossing book). Special care must be payed with the relationship between the analogous bending waves and sound waves.

6.4 Bowed Violins

Explaining about the sawtooth nature of the force imparted by a bowed string towards the bridge. If an impact hammer and an analyzer is available, developing an experimental set-up to obtain the sawtooth signal is feasible. The impact hammer is used as bridge between the two terminations of a stretched string, as in [4]. Then, bowing the string will create the sawtooth signal in time domain provided by the force transducer. Also, exploring the frequency domain (only requiring a single change in the analyzer).

Long-Term Average Spectrum (LTAS) can be easily obtained of any violin recording from a computer. Although data will be not comparable with LTAS reported in the book, at least they are useful to detect prominent resonances of the used instrument (mainly the f-hole resonance).

7. THE WOLF TONE

Several players of bowed string instruments, but mainly cellists, have heard something about the wolf tone. However, students of the course have revealed that they have not a clear idea about what is exactly a wolf tone. Preliminary, two pendulums coupled by a spring between masses are used to explain the exchange of energy between body and string that causes the wolf tone; to hear it, some measurements in a cello are performed.

The spectrum of the sound radiated by a knocked cello is calculated. The frequency corresponding to the most prominent amplitude peak is searched, which typically must occur around 180 Hz. Surely, it will correspond to the resonance involved with the wolf tone. The cello is played at the located frequency. Meanwhile, the sound is measured and graphed in time domain. The result must be that a wolf tone is clearly perceived, i.e. a tone showing an undesirable oscillation of its amplitude in variations of ≈ 5 Hz. Then, plasticine is attached to the played string, between bridge and tail piece, trying to eliminate the wolf note at the located frequency.

8. TONAL QUALITY OF VIOLINS

Violin recordings are analyzed in frequency domain while songs are played. It results of interest because violins need not be bowed by machines in order to compare playing spectra, as Meyer found [21]. Students must try to identify, by ear and seeing the instantaneous spectrum, effects of the two lowest resonances of the violins played in recordings (viz. the air resonance and the first body resonance). Also, comparing the overall behavior of the instantaneous spectrum against Dunnwald data [22].

9. CONCLUSIONS

Activities have been proposed to follow the text of the chapter dedicated to Bowed Instruments, of the classical book *The Physics of Musical Instruments*. In fact, some of these activities imply obtaining results described by Fletcher and Rossing. Moreover, several procedures here explained can be easily extended for complete undergraduate projects even though they seem very simple. Although activities were designed for lutherie students, other profiles could be interested, mainly related with mechanics and specifically with acoustics and vibrations.

10. REFERENCES

- [1] C. Hutchins, "Plate tuning for the violin maker," *CASJ*, vol. 4, no. 1, pp. 52–60, 2000.
- [2] E. Broke Dale, "Classroom demonstration of the vibration of a bowed string," *Am. J. Phys.*, vol. 44, no. 11, pp. 1077–1079, 1976.
- [3] J. Torres, J. Villarreal, and R. Ramírez, "Experimental and simulated exploration of structural deflections and acoustic waves of guitar top plates," *Rev. Mex. Fis. E*, no. 58, pp. 1–6, 2012.

- [4] J. Torres and P. Rendon, "A simple method for synthesizing and producing guitar sounds," *Eur. J. Phys.*, no. 34, pp. 503–510, 2013.
- [5] J. Shonle, "Implementing a course on the physics of music," *Am. J. Phys.*, vol. 44, no. 3, pp. 240–243, 1976.
- [6] N. H. Fletcher and D. T. Rossing, *The physics of musical instruments*, 1st ed. New York: Springer-Verlag, 1991.
- [7] L. Cremer, *The Physics of the Violin*, 2nd ed. London: The MIT Press, 1984.
- [8] F. Distribution. (2012) Free audio editor (free software). [Online]. Available: www.free-audio-editor.com/index.htm
- [9] A. Accattatis. (2012) Visual analyzer (free software). [Online]. Available: www.sillanumsoft.org/prod01.htm
- [10] Polytec. (2012) Scan viewer software (free). [Online]. Available: www.polytec.com/us/products/vibration-sensors/vibrometer-software/scan-viewer-software/
- [11] P. Falstad. (2013) Math and physics applets. [Online]. Available: www.falstad.com/mathphysics.html
- [12] K. Perkins. (2013) Phet (educational java applets). [Online]. Available: phet.colorado.edu/
- [13] S. Holdings. (2012) Lisa (free fem software). [Online]. Available: www.lisa-fet.com
- [14] A. Askenfelt, "Measurement of bow motion and bow force in violin playing," *J. Acoust. Soc. Am.*, no. 80, pp. 1007–1015, 1986.
- [15] N. Pickering, "Transient response of certain violin strings," *CASJ*, no. 45, pp. 24–26, 1986.
- [16] M. McIntyre and J. Woodhouse, "The acoustics of stringed musical instruments," *Interdisc. Sc. Rev.*, no. 3, pp. 157–173, 1978.
- [17] A. R. Lee and M. P. Rafferty, "Longitudinal vibrations in violin strings," *J. Acoust. Soc. Am.*, vol. 73, no. 4, pp. 1361–1365, 1983.
- [18] S. Zygmuntowicz and G. Bissinger, "Strad3d - a unified inquiry (dvd)," *VSA Oberlin Acoustics*, 2009.
- [19] A. Moral and E. Jansson, "Input admittance, eigenmodes, and quality of violins," *Report STL-QPSR 2-3/1982*, pp. 60–75, 1983.
- [20] M. Hacklinger, "Violin timbre and bridge frequency response," *Acustica*, no. 39, pp. 323–330, 1978.
- [21] J. Meyer, "The tonal quality of violins," in *Proc. SMAC 83*. Stockholm: Royal Swedish Academy of Music, 1985.
- [22] H. Dunnwald, "Auswertung von geigenfrequenzgängen," in *Proc. 11th. ICA*, Paris, 1985.

TEACHING PHYSICS VIA THE WEB

USING MUSIC ACOUSTICS

Joe Wolfe

George Hatsidimitris

John Smith

John Tann

School of Physics, The University of New South Wales, Sydney

J.Wolfe@unsw.edu.au georgeh@unsw.edu.au John.Smith@unsw.edu.au

John.Tann@austmus.gov.au

ABSTRACT

The UNSW Music Acoustics site provides a learning experience for its users, but it has also provided one for its makers. This paper describes how it was made and some of what we learned in making it. It also describes a new, larger project, called Physclips, which is being made with a consistent philosophy, in the light of our experience. We describe these ideas as well as some of the principles from the educational literature.

1. INTRODUCTION

Any branch of science introduces its practitioners and students to principles with broader application. For music acoustics, the list is long and there are many favourites: wave-particle duality, Weber-Fechner response, evanescent waves, heterodyne production, coherence etc. Beyond that, it can be used to teach the practice of science: experimental design and observation, analysis, hypothesis formation and testing, modeling and more.

For teaching science, music acoustics has several advantages: first, many of its important phenomena (e.g. sound, vibration, harmony) are familiar and regarded as pleasant. Some (e.g. pressure waves) can be sensed and produced with no additional hardware, and sometimes (e.g. frequency ratios) with precision high enough to allow relatively subtle evaluation of models. This makes it easy and inexpensive to create excellent lab exercises. High quality audio interfaces for computers allow for higher precision and a range of analyses.

For these reasons, many university physics teachers, including the present authors, have developed a Science of Music course for the institution's general education program: a 'physics for poets' course.

Our experience with the UNSW course prompted us to consider writing a book. Up to the late twentieth century, texts on acoustics were often published without an accompanying record, CD or web site. In the 1990s the web offered a way of integrating text, sound and movies, as well as a new way of navigating among learning elements: this seemed like a more suitable medium.

Another pathway independently led us to use the web for education. In the mid 1990s, JS and JW made a career

switch and set up a research lab in music acoustics. In this field, research results are interesting not only to scientists in the area, but also to musicians especially and to the public generally. Further, music acoustics can be an effective way of introducing students to physics.

The Music Acoustics web site at UNSW [1] was launched in 1996 with both of these objectives: to explain the research of our lab to non-technical readers and to introduce interested users, perhaps especially school students, to aspects of physics.

Feedback from users told us that the site was valued, but also gave us guidance for improvements. Many of these came too late to be retrofitted to what was then already a large site.

More recently, we have used our experience to develop a new learning and teaching site, called Physclips [2]. In the future we intend this to grow so as to cover most of the topics of introductory physics, but the volumes completed to date are the areas most fundamental to acoustics: Mechanics, Sound and Waves. Unlike the Music Acoustics website, Physclips was planned and so has a coherent structure and a consistent philosophy.

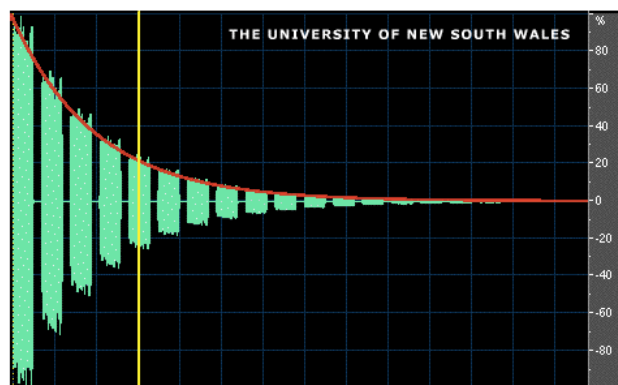
The first part of this paper gives examples from the Music Acoustics site and discusses its structure and purpose. The second part explains some of what we learned from making the site and from library research in multimedia presentation, as well as the ideas and principles that inform Physclips.

2. THE MUSIC ACOUSTICS SITE

The Music Acoustics Lab for which the site was developed began with three research projects. Two took advantage of new techniques to measure acoustic impedance over a wide dynamic range. The first involved measuring the acoustic response of the vocal tract at the lips, in parallel with the radiation impedance, which is a low impedance inertive shunt [3]. A second concerned flutes, which operate at impedance minima [4]. A third was an industrial project aimed at improving music perception by users of cochlear implants [5].

In each of these areas, we made web sites explaining our research progress for a non-specialist reader [1,6]. All three, however, involved subtleties and so led users naturally to further questions, which they sometimes sent to us. On the web, such further questions are usually answered by following links to further pages. In the twentieth century, before Wikipedia, we often saw it as our role to make these.

Auditory perception in particular was a source of questions, so we made further web pages to answer these. One, 'What is a decibel?', is still one of the most popular. It combines sound files and animated versions of their sound tracks, as suggested in Figure 1.



Broadband noise decreasing by 3 dB steps.

Figure 1. An early multimedia feature from Music Acoustics. This is one of several illustrations of the decibel scale: broad band noise pulses are reduced in amplitude by 3 dB each step. A cursor keeps pace with the plot of the sound track, illustrating the exponential decay.

The 'do-it-yourself' hearing test (Figure 2) also responded to these questions. On this page, an interactive table produces a large range of sound levels over most of the audible range, two to the octave. The user, who is advised to use headphones, can quickly make a set of equal loudness curves for each ear.

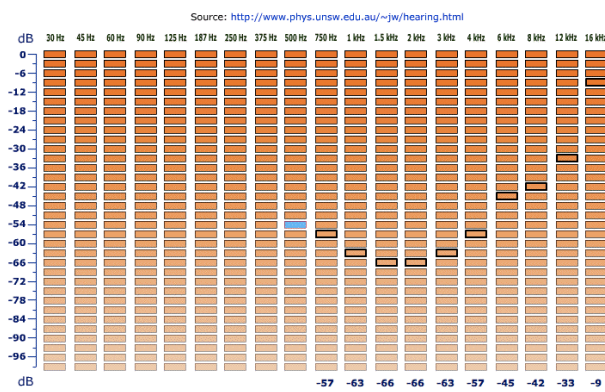


Figure 2. Also from Music Acoustics: an interactive table that allows users to determine their own equal loudness curves.

For each of the musical instruments we studied, as well as for the voice, we made web essays that introduced the acoustics of that instrument, linked to further pages on the questions raised. For the woodwind instruments, we made sites that included compendia containing sound files and impedance spectra for each of several dozen fingerings. For the flute this expanded into the 'Virtual Flute' [7], a web service that provides many thousands of machine-predicted alternative fingerings.

We made the FAQ in Music Acoustics, which is now rather large. Since then, new questions and issues have regularly required new pages. The result of this accumu-

lation is a large and growing site whose overall structure was never planned. Mostly, it comprises essays, which were illustrated with graphics, sound files and occasionally movies, interactive elements and animations.

3. PHYSCLIPS: A SITE WITH A PLAN

Mechanics, sound and waves are the physics topics upon which acoustics is based. They are also among the topics that students first meet in physics. We decided that we could contribute to learning and teaching introductory physics, at the level of senior high school or first year university, via the web.

Physclips is a large, integrated set of learning and teaching resources. It is funded by the Australian government's Office for Learning and Teaching and the UNSW School of Physics. Because of this, it has no advertising and is freely available to students, world-wide. It is also regularly used by teachers: all of its film clips and animations are available in compressed form so that they can be downloaded and included in teaching materials.

We started Physclips with the experience of the Music Acoustics site and the feedback we had received. We were joined by GH, a multimedia designer, who knew the literature on multimedia learning and teaching. We were starting from scratch, so we could plan ahead. In this section, we describe how we made it and the ideas behind it, in the hope that some of these ideas may be useful to others.

3.1 Experiments and theory

Physics is an experimental science: observation and experiment are paramount. Too often, students and sometimes even teachers miss this central point: the student may think that the lab exercises are to illustrate the theory, whereas the reality is completely the opposite: the entire structure of theoretical physics exists to explain the observations from the laboratory.

Ideally, one might hope to teach much of physics in a laboratory, but this is impractical for reasons of time and expense. In many cases, a movie of an experiment is a good substitute: this is what happened. Movies can be replayed, in real time or slow motion as necessary. Occasionally, we also use cartoons ('this shows qualitatively what we thought would happen') or animations ('this is what the simplified equations predict would happen').

Sometimes, the experiments we show require sophisticated equipment, and this is one of the common uses of Physclips: teachers, especially in schools, can use it to show experiments that they cannot perform in class. In many cases, however, we use familiar objects so that students can repeat or extend the experiments at home. In some cases, we give explicit advice on acquiring and setting up experimental gear.

Combinations of animations and movie clips can have advantages: an animation can show, superimposed on or alongside a movie, time-varying vectors or histograms and graphs. This can allow the student to share the point of view of the expert, whose 'mind's eye' 'sees' these

quantities. And of course sounds can be accompanied by spectrograms and oscillograms.

3.2 The structure and navigation

Why do so many students find physics difficult? One of the problems comes from the 'vertical' structure: higher levels depend on several levels below, and one needs to understand all levels. A typical chapter is long, and contains a number of complications and subtleties. Many students lose the view of the metaphorical forest while

dealing with many individual trees. A desire to help students overcome this difficulty inspired the structure of Physclips.

Each chapter of Physclips covers about the same material as a chapter in a typical introductory physics text. It begins, however, with a narrated multimedia tutorial, of typically five two-minute sections. These tutorials show movie clips of key experiments. Using these, plus graphics and animations as necessary, they develop the key ideas and derive the important results.

UNSW School of Physics Sydney, Australia

PHYSCLIPS
A multi level, multi-media resource

Volume 1: Mechanics
Volume II: Waves & Sound
Volume III: Light

PhysclipsWS > Sound > 4.5 Speed of sound

$2L - d = 58 \text{ m}$ $\Delta t = 17 \text{ ms}$

$$v = \frac{2L - d}{\Delta t} \approx 340 \text{ m.s}^{-1}$$

Clap 5
Clap 4
Clap 3
Clap 2
Clap 1

Time of flight method

values for speed

v_{water}	1.5 km.s^{-1}
v_{steel}	6 km.s^{-1}
v_{air}	0.34 km.s^{-1}

PHYSCLIPS
Waves and Sound

Introduction

1. Oscillations
2. Travelling waves I
3. Travelling waves II
4. Sound
5. The Doppler Effect
6. Quantifying sound
7. Interference, consonance
8. Standing waves
9. Human sound

play 4.1 Properties of sound
play 4.2 Longitudinal waves
play 4.3 Pressure and density
play 4.4 Physics of sound waves
play 4.5 Speed of sound

play pause

Links to related material

[Pitch, loudness and timbre](#)

Frequency and pitch. Amplitude and loudness. Timbre examples, with envelope and spectrum.

[Sound pressure and density: Transverse vs longitudinal waves](#)

Transverse vs longitudinal waves. $y(x)$ in a longitudinal wave. Density variations. A travelling longitudinal wave. Variations in pressure give rise to accelerations.

[Sound transmission](#)

Sound transmission through air. Bell jar experiment. If sound diffracts, why doesn't light?

[Speed of sound](#)

Time-of-flight measurements of the speed of sound. Clap-echo measurement. Clap-board (image vs sound)

Figure 3. From Physclips. At the end of the chapter about the physics of sound, we include a simple time-of-flight method for measuring the speed of sound using a recording of the direct sound of a hand-clap and its echo from a building. A portable computer or smart-phone is the only hardware needed. Below the scroll bar, a series of images shows another method using a video camera. Below the scroll bar are seen the logos and links to the first four of the support pages for this chapter. At right is the chapter outline for the volume Waves and Sound.

A ten-minute overview necessarily omits some subtleties and some of the longer mathematical derivations. To maintain rigour and to allow for interesting digressions, these overviews contain links that appear during and at the end of each section, as well as in the chapter splash

page: these link to broader and deeper discussion. Often the presentation pauses when the information content of the screen is high, and recommences with a 'click to continue' button. Support pages are also provided for neces-

sary mathematical tools: calculus, vectors, graphing and error analysis.

The Web makes it easy for users to navigate their own learning pathways, so we provide cross-links, site maps at different levels, and a search function.

Many of the users of the Music Acoustics and Physclips sites use them as a reference, searching for a particular topic or combination of key words. We sought therefore to facilitate searching on the multimedia overviews as well as on the support pages. Easy searching is also im-

portant for those using the multimedia overviews as lessons, and who wish to repeat sections. The Physclips scroll bar was designed for this purpose: when clicked, a set of key images and/or text and equations appears in order to simplify navigation to the desired section. Research suggests, however, that users often require instruction to take full advantage of this form of learner control [8].

UNSW School of Physics Sydney, Australia

PHYSCLIPS
A multi level, multi-media resource

Volume 1: Mechanics
Volume II: Waves & Sound
Volume III: Light

Search

funded by the Australian Office for Learning and Teaching

PhysclipsWS > Standing Waves > 8.5 Waves in pipes

$f_1 = \frac{v}{\lambda_1} = \frac{v}{2L}$
 $\sim \frac{340 \text{ ms}^{-1}}{2 \cdot 0.6 \text{ m}} = 280 \text{ Hz}$

— pressure
— displacement

λ_1

reflections in pipes open-open pipes open-closed pipes flarinet & clute

play
pause

PHYSCLIPS
Waves and Sound

Introduction
 1. Oscillations
 2. Travelling Waves I
 3. Travelling waves II
 4. Sound
 5. The Doppler Effect
 6. Quantifying sound
 7. Interference, consonance
 8. Standing waves
 ▶ play 8.1 Introduction
 ▶ play 8.2 Reflecting sine wave
 ▶ play 8.3 Waves on a string
 ▶ play 8.4 Strings & harmonics
 ▶ play 8.5 Waves in pipes
 ▶ play 8.6 More examples
 9. Human sound

Links to related material

[Travelling waves, superposition, reflection and transmission](#)

Wave pulses in a stretched string. Equations for a travelling wave. Linear media. Superposition. The limits of linearity. Reflections at fixed and free boundaries. Reflection and transmission at step changes in density.

[The travelling sine wave](#)

Describing the travelling sine wave. Comoving and fixed coordinates. $y = \sin(kx - \omega t)$. Three dimensional plots. Phases in a travelling wave.

[Waves in strings, reflections, standing waves and harmonics](#)

Plucked vs bowed strings. Mode diagrams and harmonics. Harmonic tuning on guitars. Touch fourths and natural harmonics.

[Physics of the sound wave. Wave equation. Acoustic impedance](#)

Displacement, compression and pressure. Newton's second law and acceleration. The wave equation for sound. Speed of sound. Acoustic impedance.

[Open vs closed pipes \(Flutes vs clarinets\)](#)

Reflection at open and closed ends. Mode diagrams for open-open and open-closed pipes. The harmonic series

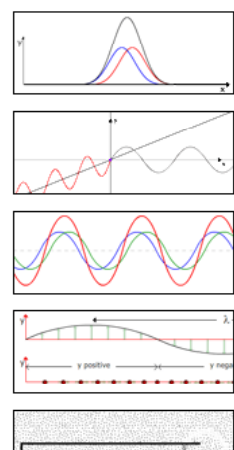


Figure 4. Standing waves in an open-open pipe and a closed-open pipe. To emphasise the importance of the boundary condition, the open-open pipe is a clarinet fitted with a flute head joint, while the closed-open pipe is a flute played with a clarinet mouthpiece. Below the main window are the scroll bar and the first of the support pages for this chapter.

4. EVIDENCE-BASED GUIDELINES

Education researchers have identified some potential problems in the use of animations and movies in teaching, and have suggested guidelines for their use in multimedia learning. [9,10]. Usually, Physclips is consistent with these. For example, one study recommends minimising the use of text in combination with moving visual material [9]: the narration reduces this problem.

Some of these guidelines would be regarded by experienced teachers as common sense. For instance, one recommended tactic is called segmentation [10], *i.e.* breaking the material up into digestible sections. Another tactic is signaling, *i.e.* drawing attention to important information. We do this in several different ways in Physclips, such as highlighting with colour, or by temporarily reducing other, competing information to grey or pale colours. Another recommendation is spatial contiguity, locating labels close to the objects labeled, rather than using a legend. We give examples of these evidence-based guidelines on the site [11].

5. USE AND THE FUTURE

The UNSW Music Acoustics site is typically accessed by a few thousand computers each day, each one downloading a dozen or more files, giving high hit rates. The usage statistics are somewhat higher for Physclips. The sites have won a number of national and international awards.

Since finishing the volume Waves and Sound, we have embarked on Light, and have posted several chapters. If funding continues, we intend to add additional volumes and to continue making this broad introduction to physics freely available on the web.

6. CONCLUSIONS

Hindsight makes for useful foresight: a planned educational website is better than one that accretes. Sometimes education research seems obvious in retrospect, but it is still worth consulting.

Acknowledgments

Physclips is supported by Australia's Office for Learning and Teaching. Physclips and Music Acoustics are supported by the School of Physics at the University of New South Wales.

7. REFERENCES

- [1] <http://www.phys.unsw.edu.au/music>
- [2] <http://www.animations.physics.unsw.edu.au>
- [3] J. Epps, J.R. Smith and J. Wolfe, "A novel instrument to measure acoustic resonances of the vocal tract during speech" in Measurement Science and Technology, 1997, **8**, pp. 1112-1121.
- [4] J.R. Smith, N. Henrich, N. and J. Wolfe, "The acoustic impedance of the Boehm flute: standard and some non-standard fingerings" in Proc. Inst. Acoustics, 1997, **19**, pp. 315-330.
- [5] R. Fearn and J. Wolfe "The relative importance of rate and place: experiments using pitch scaling techniques with cochlear implantees" in Ann. Otolology, Rhinology Laryngology, 2000, **185**, pp. 51-53.
- [6] <http://www.phys.unsw.edu.au/speech>
- [7] A. Botros, J. Smith and J. Wolfe, "The Virtual Flute: An advanced fingering guide generated via machine intelligence", in J. New Music Research, 2006, **35**, 183-196.
- [8] G. Hatsidimitris, and S. Kalyuga, "Guided self-management of transient information in animations through pacing and sequencing strategies," in Ed. Tech. Research Devt, 2013, **61**, pp 91-105.
- [9] R. Low, and J. Sweller, "The modality principle in multimedia learning", in Cambridge handbook of multimedia learning, R. Mayer, ed., Cambridge University Press: 2005, pp 147-158.
- [10] Mayer, R. "Research-Based Principles for Learning with Animation. Research implications for design.," In Learning with Animation, R.K.L.W. Schnotz, ed., Cambridge University Press, 2008, pp. 30-48.
- [11] <http://www.animations.physics.unsw.edu.au/educational-animations>

Author index

Adachi, Seiji 311
Adhikari, Ronjoy 176
Almeida, André 323, 344
Altenmüller, Eckart 115
Ambrazevičius, Rytis 211
Amir, Noam 206
Amir, Ofer 206
Angster, Judit 338
Antunes, José 549, 619, 633
Askenfelt, Anders 133
Auvray, Roman 331

Baddour, Natalie 613
Barlow, Claire 543
Benacchio, Simon 601
Bergeot, Baptiste 344
Besnainou, Charles 167
Bilbao, Stefan 569, 593, 607
Bissinger, George 32
Blaauw, Merlijn 315
Bonada, Jordi 315
Boutin, Henri 483
Brereton, Jude 231
Brock, Casey 365
Budrys, Robertas 211
Buen, Anders 38
Burban, Nicolas 585

Campbell, Murray 534
Carlsson, Peter 129
Carral, Sandra 358
Carvalho, Miguel 549, 619
Castellengo, Michèle 155, 280
Caussé, René 371, 425, 521
Chabassier, Juliette 585
Chadefaux, Delphine 155, 161
Chafe, Chris 244
Chaigne, Antoine 585
Chaldaikis, Achilleas 250
Chen, Jer-Ming 323, 415, 490
Chick, John 534
Chomette, Baptiste 601
Chrysochoidis, Georgios 217
Cottingham, James 365
Coyle, Whitney 350
Curtit, Marthe 371

Da Silva, Andrey 458
Daffern, Helena 231
Dalmont, Jean-Pierre 371, 432
David, Bertrand 54
Davis, Evan 9
de La Cuadra, Patricio 465, 473
Debut, Vincent 549, 619, 633
Delvaux, Bertrand 238
Delviniotis, Dimitrios 217, 242
Demoucron, Matthias 46, 115

Doubrovski, Zjenja 419
Doval, Boris 280
Ducceschi, Michele 607
Dumond, Patrick 613

Elie, Benjamin 54
Eveno, Pauline 371, 495
Fabre, Benoît 161, 182, 331, 465, 473
Fletcher, Neville 483
Fouilhé, Eric 60
Frelat, Joël 167
Fritz, Claudia 24, 109, 123, 140
Fréour, Vincent 502
Fu, Lei 88

Garnier, Maëva 306
Gautier, François 54
Gazengel, Bruno 344, 432
Georgaki, Anastasia 250
George, David 323
Geyer, Clemens Bernhard 514
Giordano, Bruno L. 109
Giordano, Nicholas 386
Goebel, Werner 398
Gough, Colin 66, 75
Graham, Paul 593
Gray, Alan 593
Grothe, Timo 378, 391
Großhauser, Tobias 82, 509
Guastavino, Catherine 123
Guillemain, Philippe 350
Guilloteau, Alexis 350

Hamilton, Brian 593
Haneishi, Eri 275, 311
Hanna, Noel 323
Hansen, John H.L. 293
Hatano, Hiroaki 275
Hatsidimitris, George 668
Henrich, Nathalie 306
Hirschberg, Avraham 403
Hirschberg, Michel 403
Hofmann, Alex 398
Hoge, Katrin 338
Holten, Ad 403
Honda, Kiyoshi 311
Houssay, Anne 60
Howard, David 238
Hélie, Thomas 521
Hufnagl, Bernhard 509

Jia, Ruiqing 88

Kartofelev, Dmitri 626
Kato, Kosuke 264, 269
Kausel, Wilfried 514
Kavoussanakis, Kostas 593
Kawahara, Hideki 256
Kemp, Jonathan 534

Kergomard, Jean 350
Kersaudy, Pierrick 451
Kishimoto, Hiroko 275, 311
Kitahara, Tetsuro 529
Kitamura, Tatsuya 275
Kjellström, Hedvig 133
Kob, Malte 391, 657
Koopmans, Erik 279, 299
Kouroupetroglou, Georgios 217, 242

Lagrée, Pierre-Yves 331
Lamesch, Sylvain 280
Le Carrou, Jean-Loïc 155, 161, 182
Le Conte, Sandie 155, 172, 495
Le Moyne, Sylvie 167, 172
Le Vey, Georges 409
Lehtonen, Heidi-Maria 626
Leman, Marc 46
Li, Weicong 415
Lopes, Nicolas 521
Lorenzoni, Valerio 419

Maestre, Esteban 101
Maloney, Soren 543
Mamou-Mani, Adrien 425, 601
Mansour, Hossein 94
Marques, Miguel 633
Matsubara, Masaki 529
Matsukata, Shogo 529
Matsunaga, Masaru 264, 269
McAdams, Stephen 140, 451
McBride, Jerry 661
McPherson, Andrew 440
Mehrabani, Mahnoosh 293
Menzies, Duncan 440
Meurisse, Thibaut 425
Miklos, Andras 338
Mishra, Manaswi 176
Monteil, Mélodie 557
Moore, Thomas 563
Morise, Masanori 256, 287
Mornell, Adina 509
Morrison, Andrew 563
Munoz, Alberto 432

Nakayama, Masashi 264, 269
Nakiboglu, Gunes 403
Noreland, Daniel 446
Norman, Lisa 534

Ollivier, Francois 167, 172, 601

Palmer, Caroline 279, 299
Paté, Arthur 182
Penttinen, Henri 193
Perry, James 593
Petiot, Jean-Francois 451
Plath, Niko 188
Poitevineau, Jacques 140

Pölkki, Jyrki 193

Qu, Ge 300

Reuter, Christoph 358

Ronen, Irit 206

Roy, Alexandre 161

Rudenko, Oleksii 403

Saitis, Charalampos 109, 123

Saitou, Takeshi 275, 311

Sakakibara, Ken-Ichi 205, 256

Scavone, Gary P. 94, 101, 109, 123, 451, 458, 502, 661

Schoonderwaldt, Erwin 46, 115

Schubert, Peter 299

Sharp, David 425

Shi, Yong 458

Shimokura, Yui 275

Smith, John 306, 323, 415, 483, 490, 668

Smith, Julius O. 101

Spidle, Frances 279, 299

Stoppani, George 16

Stulov, Anatoli 626

Suresh, Gowtham 176

Tahvanainen, Henna 193

Takemoto, Hironori 311

Tann, John 668

Terasawa, Hiroko 529

Terrien, Soizic 465

Theodoridis, Sergios 217, 242

Thippur, Akshaya 133

Thomas, Olivier 557

Tinnsten, Mats 129

Torin, Alberto 569, 593

Torres, Jesus A. 663

Touzé, Cyril 557, 607

Tröster, Gerhard 509

Tzevelekos, Takis 250

Umbert, Marti 315

Vauthrin, Camille 473

Vergez, Christophe 344, 350, 465

Verlinden, Jouke 419

Vitrani, Marie-Aude 161

Välimäki, Vesa 193, 626

Webb, Craig 593, 648

Weilguni, Michael 398

Willems, Jan 403

Wolf, Peter 378

Wolfe, Joe 306, 323, 415, 483, 490, 668

Wollman, Indiana 140

Woodhouse, Jim 3, 94, 147, 199, 543

Woods, Owen 199

Worland, Randy 577

Zambon, Stefano 641

Zhang, Ailin 88, 147

Zietlow, Daniel 563

The Stockholm Music Acoustics Conference (SMAC 2013) continues the series of music acoustics conferences in Stockholm, started 30 years ago. Following the tradition of SMAC 83, SMAC 93, and SMAC 03, SMAC 2013 covers the traditional fields of music acoustics, including musical instruments, singing voice, perception, and physical modeling.

The proceedings of SMAC 2013, July 30 – August 3, 2013, are peer-reviewed and include the contributions from 215 authors from all over the world.

The theme for SMAC this year is “Sound Science, Sound Experience.” During the past five decades, the domain of music acoustics has widened from studies of the acoustics of musical instruments and voice, including basic elements of musical perception and performance, to investigations of how humans experience and interact with sounds and music. Increasingly, the knowledge is put into industrial, societal and psychological perspectives. The age-old dream of bridging science and art has found new and bountiful ground in the field of Sound and Music Computing.

SMAC 2013 is hosted by the Sound and Music Computing Research Group at the Royal Institute of Technology (KTH) in Stockholm, Sweden.