# MUSIC ROOM ACOUSTICS

Papers by M.V. Mathews, I. Bengtsson &
A. Gabrielsson, J. Sundberg, E. Jansson,
W. Lassen Jordan, U. Rosenberg, H. Sjögren

Publications issued by the Royal Swedish Academy
of Music No. 17

1977

Reprinted 1981

Publications issued by
the Royal Swedish Academy of Music 17

# MUSIC ROOM and ACOUSTICS

Papers given at a seminar, organized at
the Royal Institute of Technology in Stockholm by
the Royal Swedish Academy of Music,
the Center for Human Technology, and
the Center for Speech Communication Research
and Musical Acoustics in April 1975

## Contents




ISBN 91-85428-03-5

# PREFACE

Although we almost invariably listen to music in rooms, there is generally little communication between the disciplines of acoustics of music and room acoustics. It seems that this lack of communication reduces the possibilities of our arriving at a good understanding of the basic phenomena underlying the communication of music. On April 26, 1975, a seminar was held at the Royal Institute of Technology (KTH) in Stockholm, organized by the Royal Academy of Music, the Center for Human Technology and the Center for Speech Communication Research and Musical Acoustics, KTH. The purpose was to promote a closer relationship between the fields of acoustics of music and room acoustics by reviewing the state of knowledge in each of these two areas and by exposing important but as yet unsolved problems. The papers given on that occasion are now published in the present book. The editorial team decided that the papers be published in an international language as they do not seem to possess only a strictly Swedish interest. The lack of interaction between acoustics of music and room acoustics is undoubtly international.

Apart from this book (and the impact which it will hopefully have on future research), the seminar also resulted in the formation of a Committee for the Acoustics of Music (Musikakustiska nämnden) under the auspices of the Royal Academy of Music. Its members, nominated by the Academy represent acoustics of music, psycho-acoustics, room acoustics and sound reproduction. The aims of the Committee are (1) to promote research, education and dissemination of information in acoustics of music, (2) to promote contacts between the disciplines of acoustics of music and room acoustics and (3) to assist with advice whenever possible. The editing and publishing of the present book has been one of the first tasks of this committee.

This book is published by the Royal Academy of Music. Si Felicetti of the Department of Speech Communication, KTH, has assisted in the editorial work and Robert McAllister has helped us in checking the English translations. Håkan Sjögren has made the grammophone records of the sound examples. The assistance of these persons is greatly acknowledged by the Academy and by the Committee.

Stockholm December 1976
Johan Sundberg
President of the Committe for Acoustics of Music

# ANALYSIS AND SYNTHESIS OF TIMBRES

by

M. V. Mathews, Bell Laboratories, Murray Hill, New Jersey, USA

In the last decade much progress has been made in understanding some aspects of music. New methods from speech science, from computer science, from psychology as well as the traditional methods of physics have been focused on music. I would like to show you some examples of this work starting with the study of timbre. The general approach is called analysis by synthesis which was adapted from the speech sciences and developed by G. Fant, K. Stevens, and others (see e.g. FANT, 1959 and STEVENS & HOUSE 1972). It starts out with the analysis of the timbre either by physical analysis of the instrument itself, for example the vibrations of a violin, or by signal analysis of the sound wave, for example by Fourier analysis. These analysis techniques, in general, give too much information: important information and nonessential information. The problem is to sort these out. This is done by synthesizing approximations to the sound using only the features that are believed to be important and then by listening to the sounds, the ear being the last judge of whether an adequate approximation to the timbre has been achieved. In this way, we arrive at a simple description of the sound.

Let me illustrate this by a study of the trumpet, done by Jean Claude RISSET (1969). Fig. 1 shows an acoustical analysis of the trumpet. Here we have the various overtones of the spectrum measured and plotted as a function of time. Risset noticed that in the middle of the sound the high frequency overtones are large and the low frequency overtones are small, and at the beginning and the end of the tone the low frequency overtones are relatively large. This observation is not obvious; it took a lot of study on Risset's part, but it proved to be the essen-

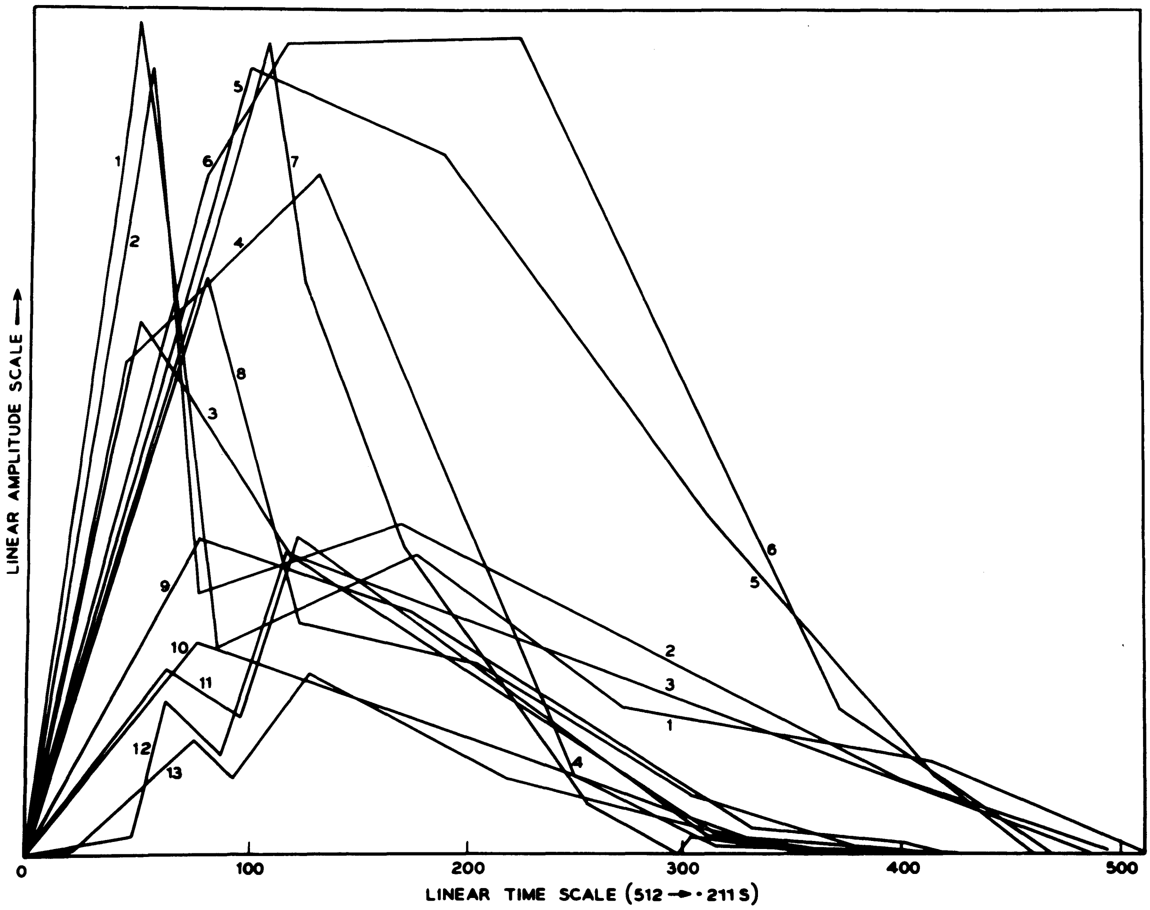



Fig. 1 Acoustical analysis of trumpet timbre. From RISSET & MATHEWS (1969), Physics today, copyright© American Institute of Physics.

tial simplification which describes the timbre of the trumpet. This simplification is shown in Fig. 2, where we simply plot the amplitude of the harmonics as a function of the amplitude of the fundamental. The highest harmonic shown, G7, becomes very large when the fundamental is large but is zero as long as the fundamental is below an amplitude of about 40. The second harmonic is zero up to an amplitude of 25 and then it rises at a smaller rate. These simpler curves characterize an essential aspect of the trumpet sound.

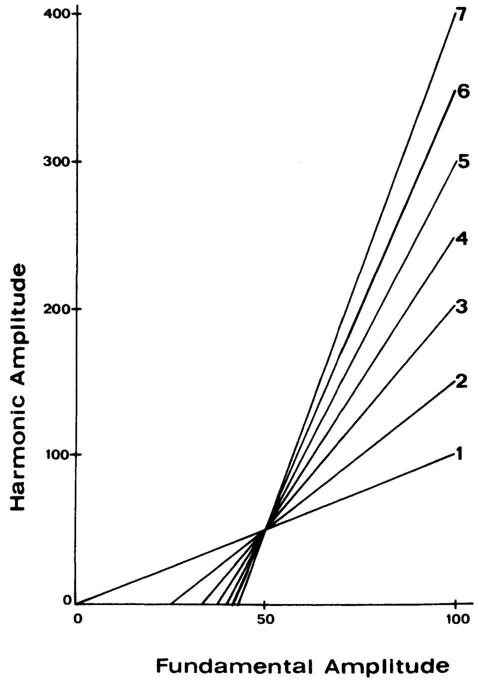



Fig. 2. Functions relating amplitudes of higher harmonics in the trumpet to the amplitude of the fundamental.

The features of these curves were put into a synthesis program. A schematic of the program is shown in Fig. 3. The fundamental and all the harmonics are generated separately with oscillators. The oscillators are shown as semicircular boxes with two controls, the amplitude control is on the left side and the frequency control on the right. At the top of the diagram is an attack and decay generator which controls the build-up and decay of the fundamental. A little overshoot is provided at the beginning of the note. The fundamental amplitude is converted into the amplitudes of the various harmonics by implementations of the functions which you saw in Fig. 2. These amplitudes modulate the various harmonics.

The essential judgment, of course, is: How does this sound? (Sound example.) The first sound example consists of ten notes. Five of these notes have been played with a real trumpet, five of them have been synthesized with this program. The order is TCTCCTCTTC. The average person can get only slightly better (60 %) than a chance score in separating the real and synthetic tones. So, we feel we have an understanding of the tone

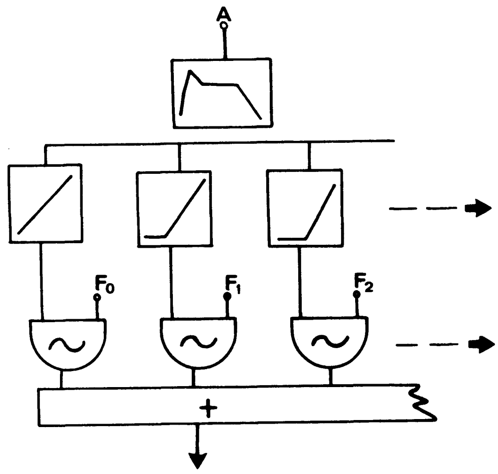


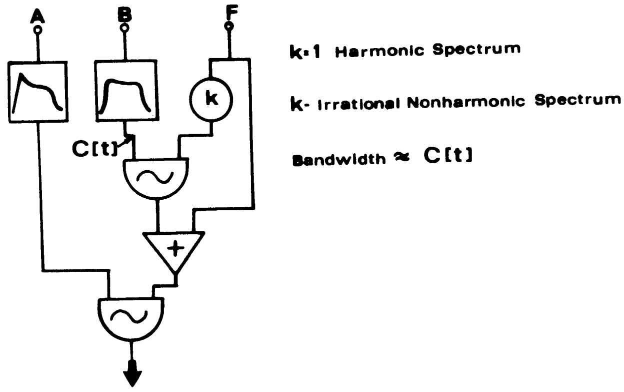

Fig. 3. Schematic of computer program to synthesize brass tones.

of the trumpet. This has been confirmed by further studies of W. WORMAN, A. H. BENADE (1973). They have shown that the characteristic of having more high frequency energy at high amplitudes is a function of the lips of the trumpeter and the way they vibrate and, indeed, is a characteristic of all instruments that are blown with the lips and also characteristic of the wood-wind instruments that use vibrating reeds.

Our understanding of the characteristics of the trumpet tone allows us to make brass sounds in ways that are very different from the normal trumpet. A striking example is the use of frequency modulation to produce brass sound. Frequency modulation techniques have been developed by John CHOWNING (1973).

Figure 4 shows the frequency modulation instrument, again as a computer simulation. This is a more efficient program than the preceding program since it requires fewer oscillators.

Frequency modulation involves the change of the frequency of a carrier oscillator by a modulation signal. The carrier oscillator is shown at the bottom of Fig. 4. Above it, is the modulation oscillator that changes the frequency of the carrier. The constant k determines the ratio of the carrier frequency to the modulation frequency. If k equals




**FM Sound Svnthesis**

Fig. 4. FM instrument

1, the carrier frequency equals the modulation frequency. This is an unusual situation in frequency modulation, but one that is of great musical interest because it produces a harmonic spectrum such as is produced by many normal music instruments. The bandwidth of the spectrum depends upon the amount of modulation which is determined by the function c(t). We can easily change the amount of modulation and thus make a larger bandwidth with more high frequency energy in the middle of the tone which we know is essential for the trumpet quality. This is indeed how we make a brass sound.

What spectra and sounds are obtained for noninteger values of k? If we make k an irrational number, we get a nonharmonic spectrum. Nonharmonic spectra are characteristic of percussion instruments: drums, wood blocks, and so forth. Again, the same kind of control on the bandwidth is possible. The <u>second sound sample</u> (<u>Sound example</u>) contains both percussive sounds and brass sounds which have been made by this technique by Dexter Morrill working at Colgate University.

I want next to turn to a different kind of a timbre which instead of being related to an instrument is related to a property of our hearing. It concerns frequency and pitch. Frequency is a physical parameter--cycles per second of a sound wave. Pitch is a perceptual parameter--one can usually compare two sounds and say one has a higher pitch than the other. In most cases, if we increase the frequency, the pitch becomes higher as shown in Fig. 5a. But there is a pecularity about this relationship: pitches that are produced by frequencies exactly one octave apart are somehow closer together in our perception than they are in the frequency scale. If we not only change the frequency, but also change the harmonic content of the spectrum, the relationship between frequency and pitch need not be this simple curve that is shown in the preceding figure but, rather, some unknown curve which in Fig. 5b is depicted as an increasing spiral.

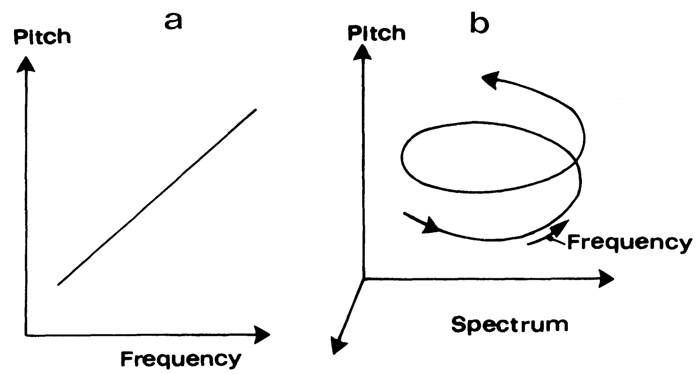



Fig. 5. Relationships between frequency and pitch.

Two points that are an octave apart are quite close together on this spiral. If we synthesize a particular spectrum, perhaps we can get this spiral to collapse so we can create the pitch paradox of a sequence of tones in which each tone has a higher pitch than its predecessor, but actually every twelfth tone is identical. (Sound example.) The third sound example is this sequence. It was done by the psychologist, Roger Shepard, Stanford University. Figure 6 shows how the spectrum was formed. The spectrum of the first tone is shown at the top of the figure and has only octave components. Thus, not all the overtones are present. The second tone in the series is shown in the middle diagram. All the components have moved up a 12th of an octave, but the amplitude of the components are adjusted according to the dotted envelope, so that the high frequency component is smaller and the low frequency component is bigger. By the time we get to the 12th tone of the series, the high frequency component is so small that you cannot hear it and the low frequency component is almost as big as the second component in the first tone. When we move up exactly one octave, we can take out the highest component and put in a very small low frequency component and you do not hear these changes. So, we have formed a circle of tones of ascending pitch and gotten back to our starting point.

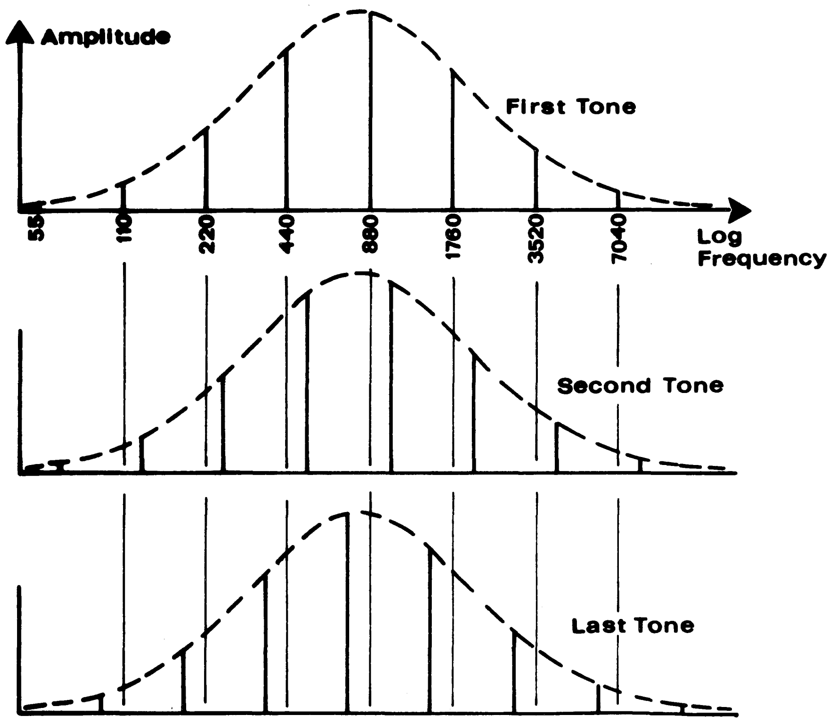



Fig. 6. Spectrum of tones for ever-ascending pitch paradox.

Risset has done this on a continuous basis, as demonstrated in the fourth sound example (Sound example). The continuous illusion is harder to make and you can perhaps hear the components come in and go out in that illusion, but the pitch always descends. K. Knowlton at Bell Telephone Laboratories has applied a similar process to rhythm. The fifth sound example (Sound example) is a rhythm that apparently gets faster and faster, but actually repeats.

Next I want to turn to some studies of the violin. The original study of the trumpet was based on an acoustic analysis of the trumpet signal. These studies have been based on a physical analysis of the violin, or at least started with a physical analysis of the violin body. Further work on the violin has been done by Erik Jansson, see JANSSON & al.(1970).

The modes of vibrations of the violin body have been known to be important for a long time, but they are very difficult to study as Jansson's work clearly shows. They are difficult to study because it is hard to do experiments with pieces of wood.

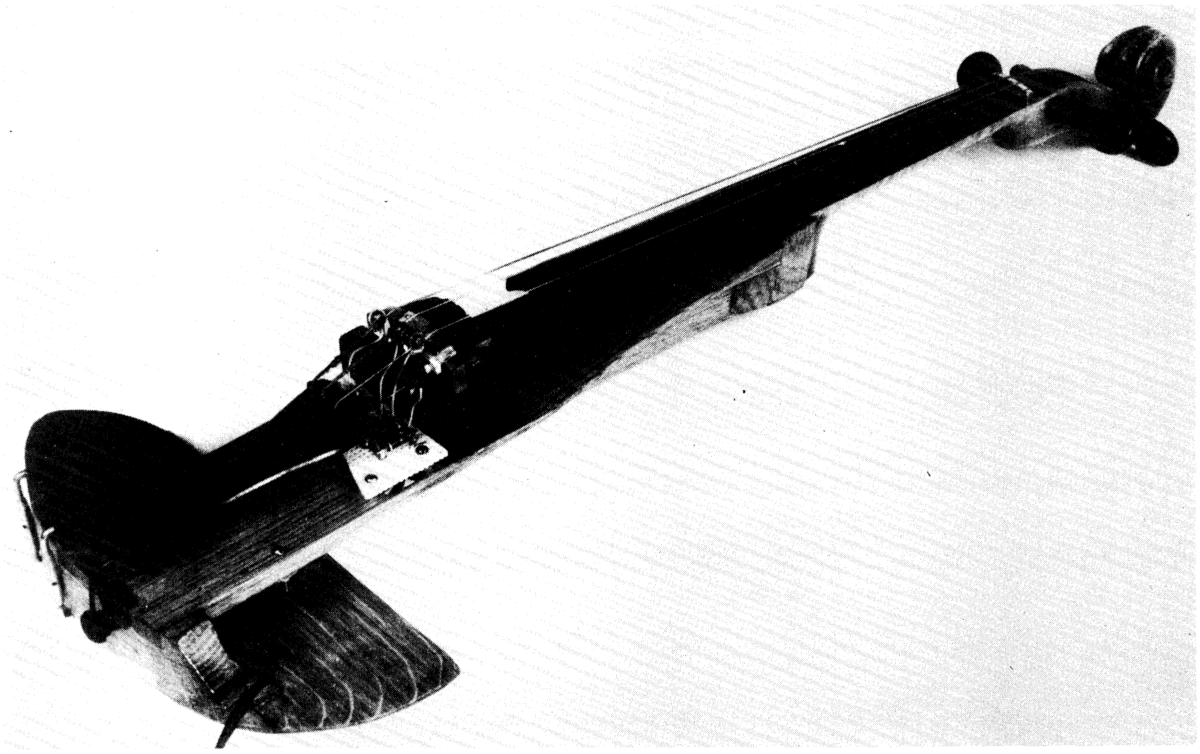


Fig. 7. Violin without body. From MATHEWS & KOHUT (1973), Journal of the Acoustical Society of America.

It takes a long time to change the shape of the wood, and making a violin is very much an art rather than science. I had the idea that we might simulate the resonances of a violin with electric circuits which are easy to make and easy to adjust and see how the sound of the instrument changes as we change these electric circuits in well-known ways. In order to do this, I constructed a violin without any body, as shown in Fig. 7.

Except for the absence of a body, this instrument is a normal violin and is played like a normal violin. The output of the violin strings is taken electrically as shown in Fig. 8 and fed into a set of resonances.

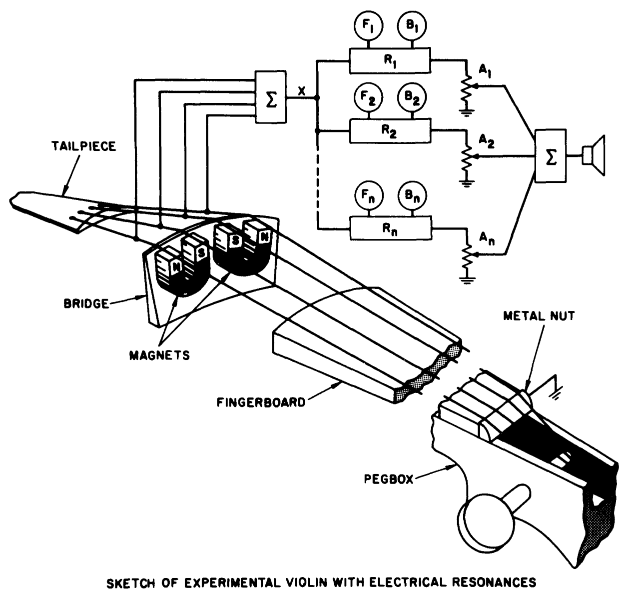



Fig. 8. Circuit schematic for violin tone studies. From MATHEWS & KOHUT (1973), Journal of the Acoustical Society of America.

Physical analysis tells us that the violin has a lot of resonances, 30 or 40, which I constructed and then I tried various experiments, with different frequencies and different Q's (damping factor) for the resonances (MATHEWS & KOHUT, 1973). The frequency response of the resonances is shown in Fig. 9 for four different values of Q. In Case No. I, we have a very low Q; a Q of zero, which eliminates the resonances completely. In Case No. II, we have a medium Q which turns out to be the best, and in Cases No. III and IV we find that too high Q's, to our surprise, sound inferior. Case No. I sounds harsh and unresponsive, Case No. II sounds the best, and Case No. IV sounds nasal or pinched, a characteristic that real violins sometimes suffer from.

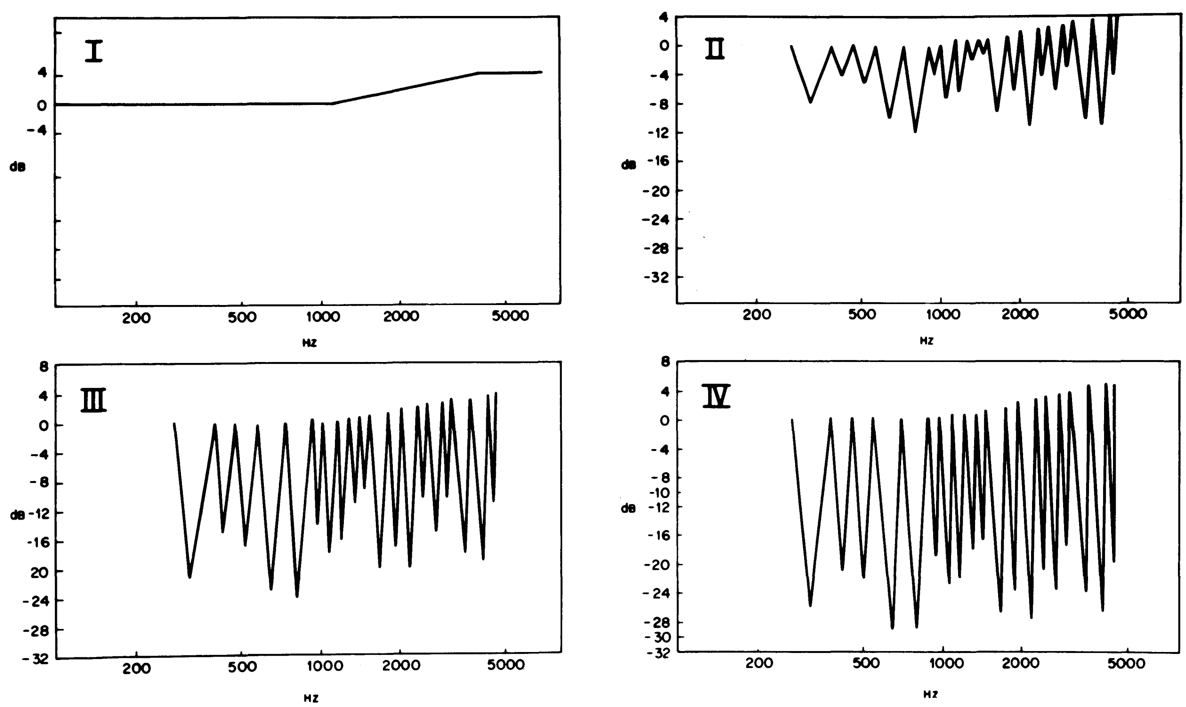



Fig. 9. Frequency response curves for four different values of Q's of violin resonances. From MATHEWS & KOHUT (1973), Journal of the Acoustical Society of America.

Sound example six is a scale repeated three times using, respectively, Q's No. I, No. II, and No. IV (Sound example).

The harshness produced by an absence of resonances is hardly surprising, but the prominence of the pinched sound of No. IV is an interesting discovery. It says that the best violin is not the one that has the strongest resonances, but rather the one which has the right sized resonances, i.e., the best violin has an intermediate value of resonant Q. It means that a secret of making violins may be tuning or adjusting the Q's of the resonances to the correct value.

My own theory, in no sense proved, is that this adjustment is done with the varnish of the violin which is a sticky substance that tends to reduce the Q. Perhaps the violin maker starts at a Q that is too high and keeps adding varnish until he reduces the Q to what he wants. However, my field is electronics. The realm of physical violin making belongs to Jansson and other people, so they have to answer this sort of question.

How good is the electronic violin? Sound example seven is a duet in which one of the violins is the electronic violin with the Case No. II resonances and the other violin is a normal acoustic violin (Sound example).The electronic violin is the first violin to be heard. I believe these two violins blend together very well and there is no real problem to use them together.

We can approximate the sounds of normal violins, but we can also approximate the sounds of very abnormal violins using some of the other techniques that we know. One interesting possibility is to put a variable frequency resonance into the violin. Violins have many resonances, but they are all at fixed frequencies. Why is a variable resonance of interest? There are only a few instruments that have variable resonances, but they are very important instruments. The human voice is the main example, so that putting a variable resonance in a violin gives it some of the qualities, some of this expressivity, which is associated with the human voice. This has been done and is shown in Fig. 10. We use the violin strings as before, we throw away all of the violin resonances, we put in an amplitude detector which simply measures the amplitude of the string vibration and we use this amplitude to control the frequency of a variable frequency resonant filter. This has been used to make

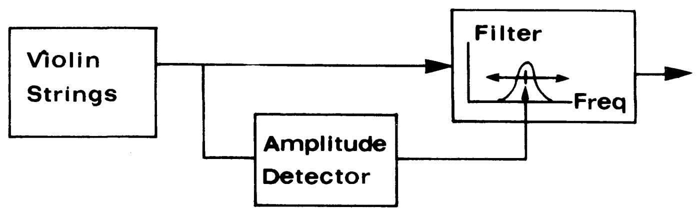



Fig. 10.  Electronic violin with "voice" timbre (variable resonance).

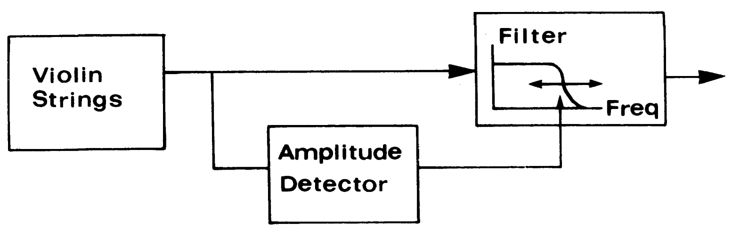



Fig. 11.  Electronic violin with brass timbre (variable low-pass filter).

some rock music by Michal Urbaniak, a Polish violinist (Sound example). Sound example eight is taken from one of his recordings. The violin is the solo voice, but it sounds so different from a normal violin that it is hard to recognize.

Another possibility is to combine the techniques which we used for the trumpet with the violin. Figure 11 shows how we can produce a brass timbre from violin strings using a variable low pass filter. Sound example nine is a sample (Sound example).

The last extension of the violin is an instrument which I call a super-bass. It is of interest because it depends on a characteristic of the human ear and its perception mechanism. The ear has what have been called critical bands of hearing (STEVENS, 1951). The property of these bands which we must take account of is: if more than one strong overtone falls in each critical band of hearing, the timbre that is so produced tends to be perceived as harsh. Low frequency tones have the overtones closely spaced together, thus low frequency sounds tend to be harsh. The violin resonance curves which appear in Fig.9 are irregular; this irregularity reduces the likelihood of having two strong adjacent harmonics and thus reduces the likelihood of having two strong harmonics in the same critical band of hearing. I believe that this property of violin resonances is no accident. I think violins and cellos and bass viols have been developed over the centuries to get smooth timbres and this led to creating irregular resonance patterns.

In sound example ten (Sound example) I have lowered the frequencies of the violin resonances by two octaves. Since violin strings do not vibrate at a low enough frequency, I have used an oscillator which is controlled by a computer to generate the melody for the example which is excerpted from a composition of mine called "Elephants Can Safely Graze".

The last sound example (Sound example) is of computer synthesized singing taken from a mini-opera "Mar-ri-ia-a" by Joseph Olive. Techniques of speech synthesis have been highly developed. Much of the primary work has been by Fant and others here in Stockholm (FANT, 1973). The speech techniques can also be applied to produce singing and one can achieve great control over all the singing parameters. The methods used were linear predictive synthesis and word concatenation (putting words together). These techniques make speech. What has one to do to convert speech into singing? In general, one has to throw away the so-called prosodic features, pitch, loudness, and duration of the speech, and impose the pitches, loudnesses, and durations which are appropriate to the music.

This particular mini-opera concerns a girl scientist who is lonely and builds a computer to talk with her. The computer falls in love with the girl, she responds for a while and then she becomes frightened of their incestuous relationship and takes the computer apart. I'll play a few of the computer's first words followed by a duet. The computer is performed by a DDP-224 machine at Bell Telephone Laboratories, the scientist is sung by Alexandra Ivanoff, a soprano. The libretto is in the appendix.

I have discussed a few examples of musical understanding that we achieve from scientific studies. There are many more that I have not mentioned: composing algorithms, sound in space, real-time computer music, and the sound catalogue, to name a few. Progress has been made in all these areas.

To summarize, I would like to repeat the advice of a pianist friend. To play well, ha said, you must first imagine the music inside your head, then you must make your fingers follow your imagination. Our present scientific and technical expertise can improve existing instruments and can create new instruments which will be more obedient to the imagination of the musician. Understanding the imagination of the musician is a far more difficult task, but new tools in psychology and linguistics can make progress here too. What is needed is a close cooperation of first-rate musicians and first-rate scientists so that the really important musical problems can be studied by the most powerful scientific methods.

References:

BENADE, A.H. (1973): "Physics of Brass Instruments", Scientific American, 229, July, pp. 24-35.

CHOWNING, J. (1973): "Synthesis of Complex Audio Spectra by Means of Frequency Modulation", J. Audio Eng.Soc., 21, pp. 526-534.

FANT, G. (1959): "Acoustic Analysis and Synthesis of Speech with Applications to Swedish". Ericsson Technics No. 1.

FANT, G. (1973): Speech Sounds and Features, MIT Press, Cambridge, Mass.; see also various articles in the Speech Transmission Laboratory, Quarterly Progress and Status Report (STL-QPSR), since 1960.

MATHEWS, M.V. and KOHUT, J. (1973): "Electronic Simulation of Violin Resonances", J.Acoust.Soc.Am. 53, pp. 1620-1626.

JANSSON, E., MOLIN, N-E. and SUNDIN, H. (1970): "Resonances of a Violin Body studied by Hologram Interferometry and Acoustical Methods", Phys. Scripta, 2, pp. 243-256; see also JANSSON's article in this book, p. 82.

RISSET, J.C. and MATHEWS, M.V. (1969): "Analysis of Musical Instrument Tones", Physics Today, 22, Feb., pp. 23-30.

STEVENS, K.N. and HOUSE, A.S. (1972): "Speech Perception", pp. 1-62 in Foundations of Modern Auditory Theory, ed. J. Tobias, Academic Press, New York.

STEVENS, S.S. (Ed.) (1951): Handbook of Experimental Psychology, John Wiley & Sons, New York.

URBANIAK, Michal: "Fusion", Columbia Record KC32852.

# RHYTHM RESEARCH IN UPPSALA

by

Ingmar Bengtsson, Department of Musicology, Uppsala University and Alf Gabrielsson, Department of Psychology, Uppsala University, Sweden.

Abstract

This paper is divided into three parts. Part I, Rhythm research - problems, methods, goals, contains a discussion of the rhythm concept, a brief review of earlier research, and an exposition of general goals for empirical research on musical rhythm. Parts II and III present two research projects going on in Uppsala, one of them dealing with investigations of rhythmic performance (part II), and the other one dealing with investigations of rhythm experience (part III).

## PART I:

### Rhythm research - problems, methods, goals

The concept of rhythm has been the subject of many treatises and discussions since the time of ancient Greece and onwards. It has been used and continues to be used for designating a variety of widely differing phenomena. Besides in music and dancing it is thus often spoken of rhythm in poetry and prose, in art and architecture, in theatre and film, in physiology and various fields of science. No wonder then that there is still no generally accepted definition of "rhythm", although the attempts at definitions may be counted in hundreds. Many examples of definitions and of the wide-spread use of the term are given in an unpublished paper by BENGTSSON (1967) and in BENGTSSON & al. (1969). In the present paper the discussion will be limited to <u>rhythm in connection with music</u> (and sometimes dancing).

Rhythm is generally considered as one of the basic elements in music. There are innumerable discussions on definitions and explanations of musical rhythm, on the relations between rhythm and other elements in music, on rhythm characteristics in the music from various time epoches or in various cultures, on training and education in rhythm etc. In most of this literature there are very differing opinions about the meaning of rhythm, and the term is often used in a vague or suggestive way. This also holds for many related concepts such as "meter", "tempo", "accent", "stress", and others. In actual music practice such terms are often used and understood in an "intuitive" way, which often seems to work in a satisfactory way. However, attempts at more rigorous definitions of these concepts - for the purpose of including them in a real "music theory" (still lacking) and in practical-pedagogical applications of such a theory - are obviously met by many obstacles.

To find a way out of this confused situation is no easy undertaking. There is need of an analysis concerning the use and the meaning of the rhythm concept as well as of empirical investigations on rhythm phenomena in connection with music. Some attempts in this direction are described in the following.

Proposed meaning of "rhythm"

From the survey of rhythm definitions and rhythm research in BENGTSSON & al. (1969; abbreviated to BGT in the following) it is obvious that the rhythm concept is used in very ambiguous ways. Thus rhythm sometimes refers to perceptions/experiences or responses in persons listening to music or sound sequences of some kind. This seems to be a legitimate use of the term (see below). In other cases, however, rhythm seems to refer to properties of the sound stimuli, that is, to physical-acoustical phenomena (for instance, talking about "the rhythm in a sound sequence" per se, independently of a listening subject). In still other cases rhythm somehow refers to properties of the musical notation (for instance, speaking about "the rhythm in a sequence of note-signs").

It seems reasonable to discard the two last-mentioned ways of using the rhythm concept. What is probably meant is that a sequence of sound stimuli may give rise to a rhythm experience/response in a listener - or that a sequence of note-signs somehow (for instance, by means of an actual or imagined performance) may give rise to a rhythm experience/response. Instead of talking about rhythm in the sound stimuli or in the note-signs it is more correct to speak of factors or properties in the sound stimuli or in the notation which may give rise to or otherwise be related to rhythm experience/response.

In our opinion the rhythm concept should in principle be reserved for denoting psychological phenomena: "One may speak of rhythm in terms of a response having experiential behavioral, and physiological aspects and complex relations between these. The aspects may be analytically separated but belong to one and the same 'general response'" (BGT, 1969, pp. 57-58, translated from Swedish). The experiential aspects of the rhythm reaction refer to various perceptual, cognitive, and emotional variables (for instance, the rhythm may be experienced as "rapid", "moving forwards", "dancing", "complex", "marked", "uniform", "vital", "calm" etc.). The behavioral aspects refer to more or less overt movements, such as swaying of the body, "beating the time" with hands or feet, even dance movements. The physiological aspects may be such as changes in the breathing, in the heart rate, in muscular tension etc. There are no sharp limits between the different aspects (this is partly a matter of definition varying with different psychological "schools"), and they are interrelated in very complex ways. The separation into different aspects is mainly intended for purposes of analysis. In reality the rhythm response may be thought of as a "whole", where the person in question is not usually aware of the parts.

Rhythm responses may be elicited by sound sequences (musical or not) with certain characteristics which are adequately described in physical terms like the durations of the sounds, their amplitudes, frequencies, spectra etc. Sound sequences may thus serve as stimuli for rhythm responses, and it is an interesting and important question which relations there are between various properties of the sound sequence stimulus and various properties of the rhythm response. In an analogous way sequences of note-signs may serve as stimuli and could be subjected to studies regarding their relations to rhythm responses. This latter point will not be pursued further here (for some comments see BGT, 1969, p. 58 and 62-63). Of course, rhythm responses may also occur in the absence of such overt stimuli, for instance, by means of imagining the sound of a piece of music. Such "internally generated" rhythm responses also present interesting questions but will not be treated here.

The distinction between stimulus level, notation level, and psychological level also applies to a number of related concepts. For instance, "accent" should refer to the psychological level, that is, perceived accent. There is strictly no "accent" in the sound stimuli per se - however, certain properties of the sound sequence (such as the intensities, durations, frequencies etc. of the different sounds) may give rise to a perceived accent on certain elements in the sequence. The question might then be posed which relations there are between such stimuli properties and the perceived accent(s).

The necessity of distinguishing between what is notated and what is experienced/perceived is often evident in connection with the concept of "measure" or "bar". Consider, for example, a well-known theme from the finale of Schubert's last symphony in C major, Fig. 1. The notated measure ($M_n$) is in 2/4 time. As indicated in the figure, however, the "perceived measure" ($M_\varphi$) may correspond to two or four (perhaps even eight) notated measures. Which case occurs may depend on the actual performance, that is, on various properties of the sounding stimuli, as well as on the "attitude" taken by the listener. More examples are found in BENGTSSON (1966, 1975).

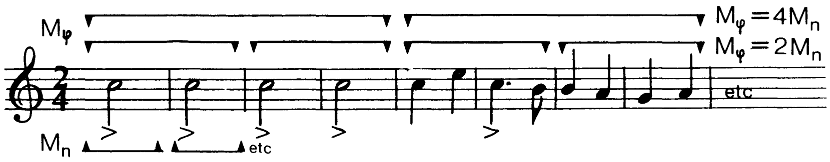



Fig. 1. Notation of a theme from the finale of Schubert's last C major symphony. The length of the notated measure ($M_n$) does probably not coincide with the length of the perceived measure ($M_\varphi$). Rather the perceived measure corresponds to two or four notated measures ($M_\varphi = 2\ M_n$ or $M_\varphi = 4\ M_n$).

It is not to be expected nor requested that the terms and distinctions discussed above should be used in the everyday conversation between musicians. It seems preferable, however, to use them in professional and scientific texts to avoid much unnecessary confusion. To simplify this it has been proposed that the intended meaning of concepts like rhythm, accent, tempo etc. be made by means of various indices or prefixes (for examples, see BGT, 1969; BENGTSSON & al., 1972; GABRIELSSON, 1973a).

"Rhythm$_f$" or "S-rhythm" may denote the sound sequence eliciting a rhythm response. "Rhythm$_n$" or "N-rhythm" may refer to notations (that is, "notated rhythms")."Rhythm$_\varphi$" or simply "rhythm" without any index may refer to the rhythm response/experience. In a similar way "measure$_n$" means "notated measure" and "measure$_\varphi$" means "perceived measure" (see Fig. 1).

## Empirical investigations of rhythm

Experimental investigations of rhythm began shortly after the emergence of scientific psychology in the latter half of the 19th century, and a large number of research reports appeared during the decades around 1900. After about 1920, however, the interest for rhythm research seemed to decrease, and in today's psychology rhythm questions are rarely discussed. A detailed historical account of this research from about 1875 to the late 1960's is given in BGT (1969).

The following enumeration gives some hints of investigated phenomena: "subjective rhythmization" (the fact that one tends to hear a rhythm even in a wholly uniform series of sounds); perception of accent and its relation to various factors in the sound sequence; perception of grouping of sounds to "unities" of some kind; the relative importance of temporal relations and intensity relations in the sound stimuli for eliciting rhythm responses; the role of motor/kinaesthetic components in the rhythm response; the performance of various "rhythmic structures" and so on. In most of these investigations the sound sequences used as stimuli were rather "un-musical" (far too simplified to represent actual music or approximations of that). The most "realistic" studies were experiments on performance characteristics, that is, musical performances were registered by means of various technical equipment to give information about the durations, intensities, frequencies etc. of the tones in the performed music. These studies convincingly illustrated how highly flexible and variable actual musical performance is in relation to the norms implied by the musical notation system: the (seemingly) exact durational relations between different "note-values" (half-notes, quarter-notes etc.) have no counterpart in actual performance, the metronomic tempo indications are followed in very approximate and continuously varying ways, tones following each other in notation may actually "overlap" each other, deviations from "correct pitch" are very frequent etc.

In order to advance the research on musical rhythm it seems necessary to combine earlier approaches in a more effective way. Questions about performances and questions about rhythm responses/experiences must be considered in connection with each other. As regards methods, there must be a continuous interplay between "naturalistic observations" of performances and well-controlled experimental investigations on various questions. The progress during the last decades concerning electronic equipment for synthesis and analysis of sound sequences, the manifold of methods proposed for description and "scaling" of experience dimensions, and the enormous facilities for handling large amounts of data by means of computers seem to present good possibilities for future rhythm research. Some goals for such research are discussed below.

General goals for empirical research on musical rhythm

Some purposes of rhythm research may be discussed in connection with Fig. 2.

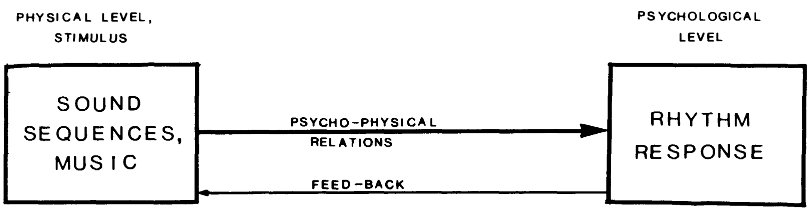



Fig. 2. General framework for empirical rhythm research.

A distinction is made between the physical level or stimulus side on one hand and the psychological level (to which the rhythm response belongs) on the other hand - compare the discussion above about "meaning of rhythm".

The stimuli are, generally expressed, sound sequences of some kind. In much earlier research they were generated mechanically by means of metronomes, pins on rotating cylinders, and the like. Nowadays there are good possibilities for electronic synthesis of sound sequences with well-controlled characteristics. And most natural of all, the stimuli are represented by music of some kind as performed by a musician or a group of musicians. Music stimuli may be taken from existing recordings (gramophone records, tape recordings), or new recordings can be made. New recordings are necessary if it is desirable to control or have knowledge about the conditions under which the performance takes place.

The sound sequences may give rise to rhythm responses in listeners, and a general goal for rhythm research is then, as noted earlier, to find out which relations there are between various properties of the sound sequence stimulus and various properties of the rhythm response. Such relations between physical stimuli and psychological responses are often called psychophysical relations. However, in order to be manageable this very general question has to be split up into a number of more specific questions, some of them pertaining to the stimulus conditions, some to the psychological responses, and still others to the more or less specific relations between stimuli and responses. The rest of this paper may be said to present various such questions within the general framework summarized in Fig. 2.

Rhythm versus non-rhythm

A first question might be: Which are the characteristics of a rhythm response as opposed to a non-rhythm response? Some suggested characteristics for a rhythm response are perceived grouping (the sounds in the sound sequence are not perceived as separate elements but are "organized" into groups of sounds), perceived accent on certain elements in the sound sequence, and perceived regularity of some kind (for instance, regular occurrence of accentuated elements or the regularity implied by the pulse rate, that is, the tempo). Further the rhythm response takes place as a more or less immediate response to the sound stimuli in question and lies within "the psychological present", that is, the psychological "now", which has a duration of some seconds depending on many various conditions. (It seems doubtful to speak of rhythm outside the psychological present, for instance, when talking about "the rhythm of the movements in a symphony" and the like.) The above-mentioned characteristics obviously refer mostly to rhythm experience. As regards characteristic behavioral and physiological aspects of the rhythm response the knowledge is rather diffuse. For some further comments see BGT (1969, pp. 58-61) or BENGTSSON & al. (1972).

In an analogous way, one may ask which are the necessary characteristics of a sound sequence to elicit a rhythm response. This is about the same thing as asking which stimuli conditions are necessary for perceived grouping, perceived accent etc. to appear. Several investigations indicate that if the duration separating the elements in a sound sequence exceeds 2-3 seconds there will in most cases be no perceived grouping or rhythm experience. Consequently the duration separating the elements should be lower than this, and there is ample evidence that the shorter this duration is (within limits), the more elements the perceived group will include. The total duration of the group must lie within the psychological present, the upper limit of which is very difficult to determine, however. As regards perceived accent many investigations indicate that it is related to many different stimulus variables as durations, intensities, frequencies etc. of the elements in the sound sequence - or perhaps rather to the relations in durations, intensities, frequencies etc. between the elements as well as to more complex variables loosely called "melodic and harmonic factors". As noted above, however, perceived grouping and accent may occur even in the case of a wholly uniform sequence of sound stimuli!

The above results were for the most part obtained with short "click sounds" as stimuli, while tones in music as a rule are more continuous (as in legato performance). However, similar phenomena appear for music, too. For instance, in many organ works based on a chorale melody each tone in the cantus firmus often extends for several seconds, and it may be very difficult or even impossible to perceive these tones as "grouped" together in a melody.

## Rhythm responses and their relation to the (musical) stimuli

The problems briefly discussed above concerned characteristics of the rhythm response as opposed to a non-rhythm response. Next we will proceed to the problem how to investigate and describe various kinds of rhythm responses and how to find out what it is in the sound sequence stimulus that brings about such different rhythm responses. We will mainly be concerned with the experiential aspects of the rhythm response and so the term rhythm experience will be used as often as rhythm response. Some examples may make the problem more concrete:

Wherein does the specific rhythmic character of a Viennese waltz from the Strauss dynasty consist? How could this type of rhythm experience be described? And how do the musicians actually play to bring about this well-known and charming rhythm?

How to describe the specific rhythm character of a "polska" from Dalecarlia in Sweden with all its "ornaments" and other intricate features? And what is it in these fiddlers' performance that gives rise to this unmistakable rhythm character?

Which are the rhythm characteristics of various dance types such as waltz, foxtrot, tango, samba, hambo, menuett, rock'nroll and many others? Is it possible to construct some kind of system for a convenient description of their different rhythm characters? And how are these dances actually played to supply these characters?

Som other examples in more condensed form:

It is often said that there are typical and prominent rhythm characteristics in the music of different composers (for instance, Stravinsky or Bartók in our century). It is sometimes stated that music from different periods have different rhythm characters, for instance, that music from the Baroque era has certain rhythm characteristics different from those in, say, the Vienna classicism. Jazz music is said to be more or less "dominated by rhythm" - however, even during the short history of jazz there have appeared rather different rhythm characters (dixieland, swing, bebop etc.). Music listeners often learn to recognize specific rhythm characters associated with a certain performer, a certain orchestra etc. From ethnomusicology we know that musical rhythm is often quite different from what is common in Western music.

In all these cases the same principal questions can be raised: In which way can such various rhythm characteristics - associated with composers, performers, historical epoches, musical genres etc. - be adequately described? And how are these different rhythm characters related to the sounding music - what is it in "the music stimulus" that brings about the different rhythm characters?

In terms of Fig. 2 we are asking, on the psychological level, for some kind of descriptive system to enable us to adequately and comprehensibly describe various forms of rhythm responses/experiences ("rhythm characters"), and, on the physical level, for which features in the sound sequence stimuli which give rise to the various rhythm responses. Or, once again, which psychophysical relations are there between different properties of the sound sequence stimuli and different kinds of rhythm responses? There is also an important feed-back link entering here, see Fig. 2, referring to the fact that the performing musician not only produces sounds - he also listens to them and gets, among other things, a certain rhythm response/experience from his own playing which may modify his performance in various ways so as to "optimize" the rhythm response. It is a sort of control mechanism like that occurring in speech - a speaker listens to his own voice to check that his speech is congruent with his intentions. The feed-back link may also be understood in a more general way, viz. including the responses of present listeners. Their responses, in which ever form they appear, may also affect the musician's way of performing, often profoundly so.

It must be added that the rhythm response is also related to various factors associated with the listener himself. The fact that different individuals may differ widely in their rhythm response to the same music points to the importance of earlier experiences. A listener accustomed to Western art music often has considerable difficulty "grasping" and appreciating rhythm characters present in music from other cultures or in folk music from various countries, including his own. For instance, many Swedes apparently find the rhythm in a "polska" from Dalecarlia in Sweden very strange, while people acquainted with this type of music may enjoy it in a very intensive manner. It seems obvious that the rhythm response, as well as the response to music in general, can be highly modified by learning. The same principle is also found as regards a musician's ability to perform various kinds of music in a way that creates the "proper rhythm character" of the music in question (see below).

The goals for rhythm research discussed above apparently require a long-term project, including necessary economical resources, to be attained. In the following we will report some different approaches followed in the Uppsala rhythm research. Two major avenues of investigation may be discerned, although they in fact overlap each other: one project mainly dealing with studies of performances, and one project mainly dealing with descriptions of rhythm experience.

## PART II:

### Investigations of rhythmic performance

This project was started already in the late 1950's by Bengtsson who gives a detailed account of its background and development in BGT (1969). Observing the confused terminology regarding rhythm, meter, tempo, accent etc. in musicology and the prevailing tendency of many musicologists to study rhythm phenomena mainly in notations of music, Bengtsson concluded that it was necessary to study rhythm phenomena in connection with "living music", that is, "'living music' both in the meaning of sound sequences produced by behaviors of performers and in the meaning of responses in listeners" (BGT, 1969, p. 51). As noted earlier, the sound sequences used in experimental investigations by psychologists were for the most part musically un-interesting. Furthermore, the interchange of ideas between musicologists and psychologists has been very meagre.

The initial problem may be described in the following way. Listeners often make more or less evaluative judgments concerning the experienced rhythm in various types of music or in different performances of the same music. For instance, a certain performance is described as "very rhythmical", "lively", "engaging" etc., while other performances may be characterized as "un-rhythmical", "mechanical", "dead" etc. Such statements often refer to certain music genres, for instance, dance music or folk music for which the way of performance often varies with geographical regions. Some musicians are capable of playing the music in question in a self-evident manner to ensure its proper rhythm character, others are not - but those others may be superior when it comes to performance of another kind of music. The musician who succeeds in rendering "all types of music" with its proper rhythm character is a rare bird indeed. Most musicians have a limited vocabulary in this respect.

The Viennese waltz is a well-known example. The way of playing the "three beats per measure" to bring about the typical Viennese waltz rhythm is an art which is certainly not shared by all musicians or orchestras trying to perform the Strauss waltzes. There is an abundance of similar examples. Skilled performers of Western art music seldom succeed in playing jazz music with the proper "swing" or dance music of some kind to give that irresistible need of dancing or moving to the music. Europeans playing Latin American dance music or singing negro spirituals often sound very strange compared to native performances of this music. And for an outsider to learn to play a "polska" from Dalecarlia in that highly characteristic and self-evident way of the local fiddler requires lots of listening and training to succeed, if ever.

The main problem for the research project may then be formulated as follows: <u>Which are the relevant and important features of the sounding music which give rise to such typical and often easily identified rhythm characteristics</u>? In other words: How do the musicians actually play to bring about these characteristics?

In terms of Fig. 2 the main part of the project is thus centered on an analysis of the actual stimulus, the performed music. It is presupposed that the selected performance elicits the "proper" rhythm response, that it is a representative example of a good performance of that kind of music, as judged by experts or other people acquainted with that type of music.

## The SYVAR-D hypothesis

In accordance with some earlier suggestions and some data it may be hypothesized that an important stimulus factor for the eliciting of various rhythm characters is to be found in some kind of <u>systematic variations as regards tone durations</u> (SYVAR-D). In all musical performance there are certain random variations as well as certain systematic variations. Random variations are unintended and occasional (for instance, due to technical imperfections). Certain kinds of systematic variations are made in a deliberate way to emphasize or point out various structural or "expressive" properties of the music in question, for instance, by means of tempo changes (rubato, ritardando etc.), dynamic changes, intonation etc. However, there are probably also often other kinds of systematic variations of which the performer is not so aware. These are in general a result of continued experience with the music in question. The performer has grown up in a certain musical tradition and/or has been educated in a certain performance conventions, e.g. how to play a Viennese waltz, a "polska" from Dalecarlia, how to make it "swing" in a piece of jazz music etc.) and in general knows that he adopts a special way of playing a piece of music - but he is seldom capable of explaining how he actually makes it.

These "special ways of playing" such "rhythmic dialects", might show up in the form of systematic variations as regards the duration variable. Systematic variations - in relation to what? Some kind of norm or reference is necessary to make the concept meaningful. For much music an easily available and well-known reference would be the duration relations commonly implied by the conventional musical notation. That is, a tone notated by a half-note would be twice as long as one notated as a quarter-note, the latter one twice as long as a tone indicated by an eighth-note etc., all measures in the same meter are equally long and so on. The durations of the various sound events in a performance could thus be studied with this "rational-mechanical" norm derived from the notation as a reference basis: do the duration relations in the performance agree with or do they show any kind of systematic variations (deviations) in relation to this "mechanical" norm?

To take an example, consider the beginning of Mozart's piano sonata in A major (Köchel 331), Fig. 3.

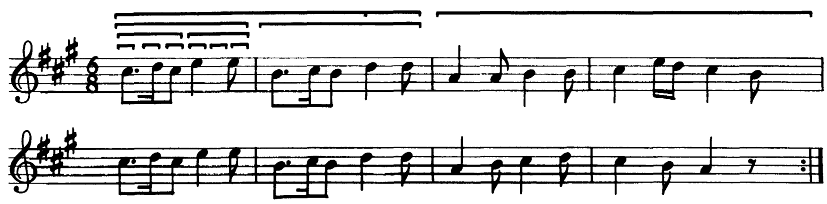



Fig. 3. The melody part of the beginning of Mozart's piano sonata in A major, Köchel no. 331. The "brackets" above the notation refer to different levels (eighth-note level, half-measure level, measure level, level of two measures) as discussed in the text.

There are lots of performance questions which could be discussed for this example. Only some few of them will be mentioned here to illustrate the concept of SYVAR-D.

SYVAR-D may show up on different levels as indicated in the figure. For instance, the ♪.♬♪ notation in bars 1, 2, 4 and 5 would imply a duration relation of 3:1:2 if strictly followed. Will this actually happen in a performance or will there perhaps be any systematic deviation(s) from that norm (as one single example, making a "soft punctuation" so that the tone notated as "punctuated eighth-note" is somewhat shortened and the second tone notated as "sixteenth-note" is somewhat lengthened in relation to what the notation seems to say - or perhaps the second tone is lenghtened at the expense of the third tone)? The ♩♪ notation in the second half of the same bars would imply a 2:1 relation in duration. Will this be the case in the performance or not (it might be expected that this relation is affected by the performer's intention to have the last tone appear as the ending of a musical phrase or as "up-beat" to the following?

On the next level it may be asked whether the ♪.♬♪ notated first half of the measure and the ♩♪ notated second half will be equally long in the performance. A systematic deviation in a direction towards "somewhat longer first half" is possible, as well as its opposite, resulting probably in differently experienced rhythm characters. On still higher levels one may ask whether all notated measures are made equally long, if groups of two notated measures or four notated measures are performed equally long in duration etc. The reader may supply his own suggestions in more examples.

It seems reasonable also to expect that the SYVAR-D's may be dependent on the tempo. They may appear most obviously in a "proper" tempo region, while they may be less pronounced or "destroyed" in too slow or too rapid tempi.

Of course, there may be quite different types of SYVAR-D's for different performers or even for different performances by the same person. This implies that it is necessary to get data from several performances to see whether consistent SYVAR-D's appear in one form or another.

The SYVAR-D hypothesis puts the duration variable(s) in focus in the search for stimulus variables critical for the eliciting of certain rhythm characters. It is probable, however, that other variables are of importance, too - for instance, intensity, frequency, spectrum or more or less complex constellations of all mentioned variables. There may thus be SYVAR in many different respects depending on many various factors (for instance, which instruments are used). For the present, however, the main interest is devoted to the duration variable.

Obviously the "rational-mechanical" duration relations derived from the notation are not applicable for all music: there is much un-notated music and even if notations are tried for such cases they often raise so many problems that their relevance for investigations of SYVAR-D is dubious. In such cases other norms have to be found by means of successive comparisons and/or various statistical procedures.

The concept of duration is in itself ambiguous. With regard to the successive events in a sound sequence one may distinguish between

a) the duration from the beginning of a sound to its end, called "duration in-out", Dio
b) the duration from the beginning of a sound to the beginning of the next sound, "duration in-in", Dii and
c) the duration from the end of a sound to the beginning of the next sound, "duration out-in", Doi.

Dio and Dii may coincide or not depending on whether there is a "no sound duration" (Doi) between the successive sounds or not. It is also possible that Dio is longer than Dii in the case that successive sounds are overlapping. In most cases Dii is probably the most reasonable unit to use for SYVAR-D investigations. However, in many cases the Dio/Doi relation may be of considerable importance so all three measures must continuously be observed. These questions are discussed further in BGT (1969 pp. 101-102) and briefly in BENGTSSON & al. (1972).

Sound analysis equipment

To enable accurate registrations of performed music or other sound sequences various electronic devices for analysis of sound sequences have been developed in cooperation with physicists in Uppsala. The first and most frequently used analyzer was christened with the Swedish girl name

"MONA", hinting at its use for analysis of monophonic sound sequences. MONA may be called a "melody writer" which presents an accurate analysis of the monophonic music in terms of frequency and amplitude variations over time. For analysis of polyphonic music another analyzer christened "POLLY" has been tried in several variants. Its construction has met with many technical difficulties and so POLLY has as yet been used only for pilot studies of various kinds. Detailed descriptions of MONA, POLLY, and some other apparatus, including many illustrative figures are given in BGT (1969 pp. 87-93), BENGTSSON & al. (1972), and BENGTSSON (1967 b).

The reading of the MONA registrations concerning durations presents certain problems as to the exact marking of the beginning and the end of each sound event. Numerous attempts have been made in various ways to establish reasonable and reliable routines, some comments are found in BGT (1969 p. 102) and in BENGTSSON (1967 b).

Some examples of SYVAR-D

It is fairly well-known that the specific rhythm character of a Viennese waltz has something to do with the way the accompanying instruments perform the "after-beats" on the second and third beat of the measure. An early pilot study was made on the Vienna Philharmonics' performance of the first four bars in a Viennese waltz including only the accompanying instruments before the entrance of the melody. An obvious SYVAR-D pattern appeared: "short first beat, long second beat". That is, compared to the "mechanical" norm derived from the notation, the first beat was considerably shortened, while the second was considerably lengthened (this is also clearly heard if the recording is played at half the correct speed). While the duration relation between the two beats would be 1:1 according to the notation, it was in fact partly displaced to the ratio 2:3 in the performance. It was also evident that a kind of SYVAR-D appeared on the level corresponding to two notated measures so that the "intended measure", and probably also the "perceived measure", may be said to comprise two measures in the notation. This example is shown in detail in BGT (1969, p. 97) and BENGTSSON & al. (1972).

A more recent analysis of the first eight measures of another Viennese waltz ("Kettenbrücke Waltz" by J. Strauss senior) performed by the

Boskowsky Ensemble (Philips SGL5757) confirms and simultaneously adds some nuances to the picture. This analysis was made by means of POLLY and only a fraction of the abundant data from this analysis can be presented here. The notation appears in Fig. 4 and the music i given as sound example one (Sound example). Under the notation there is a schematic representation of the performance. The upper "line" represents the melody as played by the first violin. The two lines in the middle (with "triangles") represent the accompanying chords appearing on the second and third beat of each bar, the upper line corresponding to the second violin and the lower to the viola. At the bottom, finally, there is the bass part. The heavy solid vertical lines represent the borderlines between the measures if they had been equally long. The markings between the two lines with "triangles" represent the borderlines between the three beats within the measure, if they had been equally long. The dashed vertical lines extending from the bass part upwards represent the entrance of the first beat in each measure for the bass part and may thus be considered as the border-lines between "measures" in the performance.

It is seen that this latter border-line lags behind the "rational" border-lines for bars 2-5, while it anticipates them in bars 6-8. This reflects tempo changes frequently occurring in Viennese waltzes: a "slow" beginning, then acceleration, and then slower again towards the end of the phrase. The same phenomenon can be seen if you instead follow the entrance of the first beat of the melody part. It starts in fact "too early", lags behind (bars 3-4) and then anticipates the "rational" border-lines (bars 5-8). As one might expect it also "leads" before the bass part except for the last bars where it is slightly overtaken by the bass.

Studying the two middle parts reveals some differences between the performance of the second violin and the viola, see also Fig. 5. The pattern of "short first beat" (actually a rest) and "long second beat" appears most obviously in the viola: its second beat comes "too early" (both in relation to the "rational" border-lines and to border-lines between first and second beat as derived from the dashed border-lines). The duration of the third beat conforms rather well to what follows from the "mechanical" norm (see Fig. 5), except in bars 4 and 6 (the

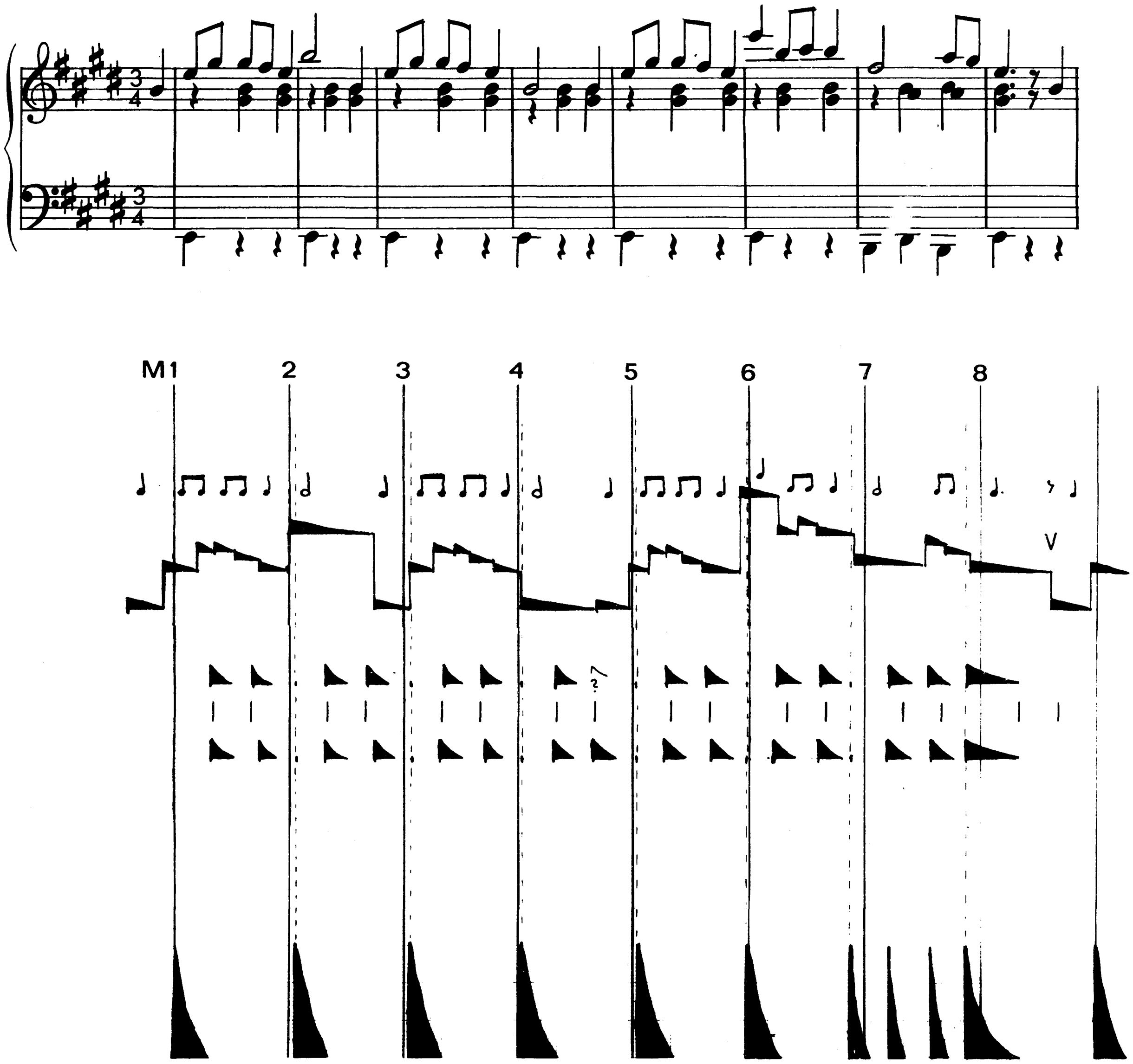



Fig. 4. Upper part: Notation of the beginning of the "Kettenbrücke Waltz" by J. Strauss senior. Lower part: Schematic representation of the performance, see text for explanation.

prolonged third beat in bar 4 is apparently made for the case of a long "up-beat" to the following phrase). In the second violin, however, the pattern seems more variable. The first beat is shortened, as for the viola, while the second and third beats are both more or less prolonged, the third beat especially in bars 2 and 4 where it may be considered as an "up-beat" and where the second violin actually plays the same tone as in the melody. However, towards the end of the example (bars 6-7) the pattern for the second violin conforms to that followed by the viola.

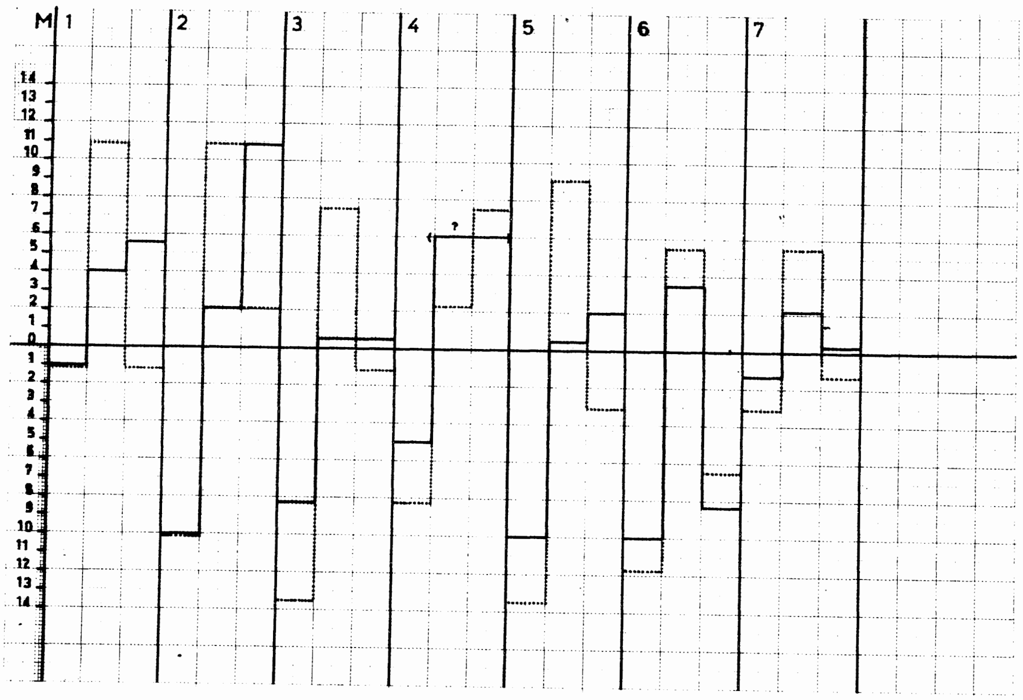



Fig. 5. Comparison between the performance of the second violin (solid lines) and of the viola (dashed lines).
The horisontal line in the middle (marked zero) represents the case that all beats are made equally long (= "mechanical" norm). Deviations upwards from this line mean that the corresponding beats are lengthened and deviations downwards that they are shortened in relation to what would follow from the "mechanical" norm. The deviations are expressed in pro mille of the total duration of the first eight bars.

Looking at the duration relations for beats in the melody, see Fig. 6, reveals that in general the first and the second beats are longer than the third beat which is shortened (except for bar 7 where the third beat on the contrary is longer due to a ritardando towards the end). In comparison with that the duration relations for the accompanying instruments (averaged over the second violin and the viola in this figure) in most cases follow the pattern "short first beat" - "long second beat" - "intermediate third beat", except for bars 2 and 4 as noted above. Both for the melody and for the accompaniment there is a tendency to a SYVAR-D on the level corresponding to two notated measures, especially for bars 1-4. Look at the "profile" for the melody in bars 1-2 (Fig. 6) and then find a similar profile in bars 3-4. A similar two-bar grouping also occurs for the accompaniment, see the profiles for the accompaniment in Fig. 6. The reasons for a two-bar grouping in the performance are rather evident from the notation (Fig. 4).

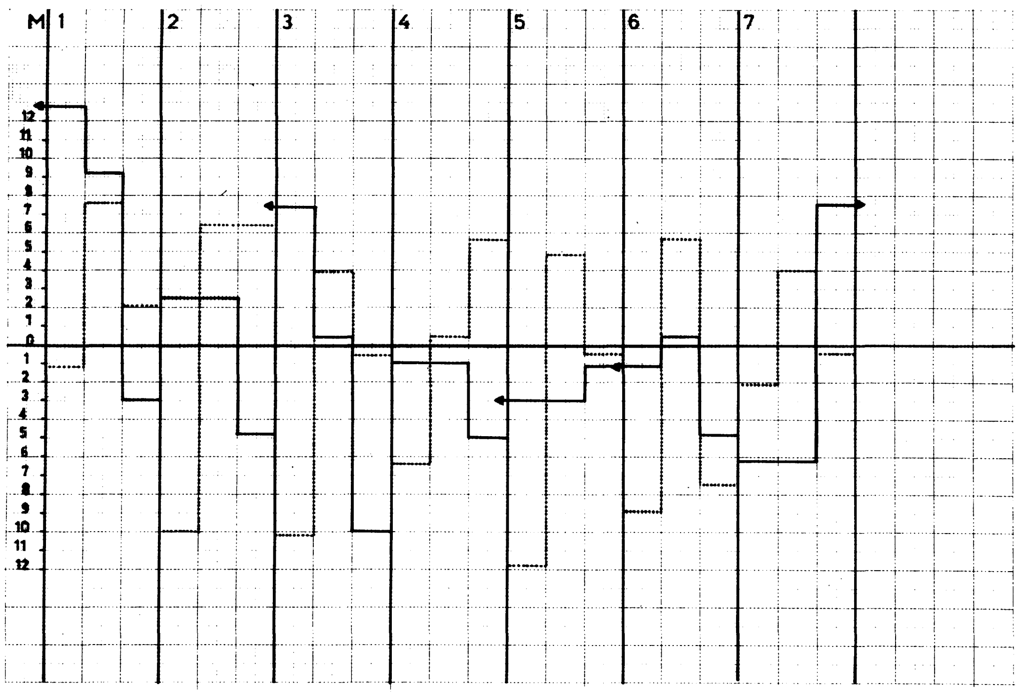



Fig. 6. Comparison between the performance of the melody (solid lines) and of the accompanying instruments (dashed lines representing the average over the second violin and the viola). Explanations as in Fig. 5.

Much more could be discussed but the comments here may suffice to illustrate the general reasoning. A careful listening to the music is recommended and listening at half the speed may highlight certain points.

Fig. 7 and sound example two (Sound example) is taken from Swedish folk musik in Uppland, Eric Sahlström performing a "bridal march" by Gelotte

on the "key-fiddle" (Swedish: "nyckelharpa").
Several measures of this piece has the notation ♩ ♫ ♩ ♫ (as in the first two bars), and it is the mean values for the durations in this pattern which are shown in a computation scheme above the notation. The lower part of this scheme shows the values which should occur according to the "mechanical" norm. Thus the first quarter-note would occupy 25 % of the total duration of the measure, each of the following eighth-notes would occupy 12.5 % etc. Expressed in milliseconds (above the per cent figures) this means that the quarter-note would correspond to 288 msec and the eighth-note to 144 msec (in the mean tempo used here). The actual values in the performance appear in the upper part of the scheme. As seen there both quarter-notes are considerably lengthened in the performance (compared to the "mechanical" norm) so

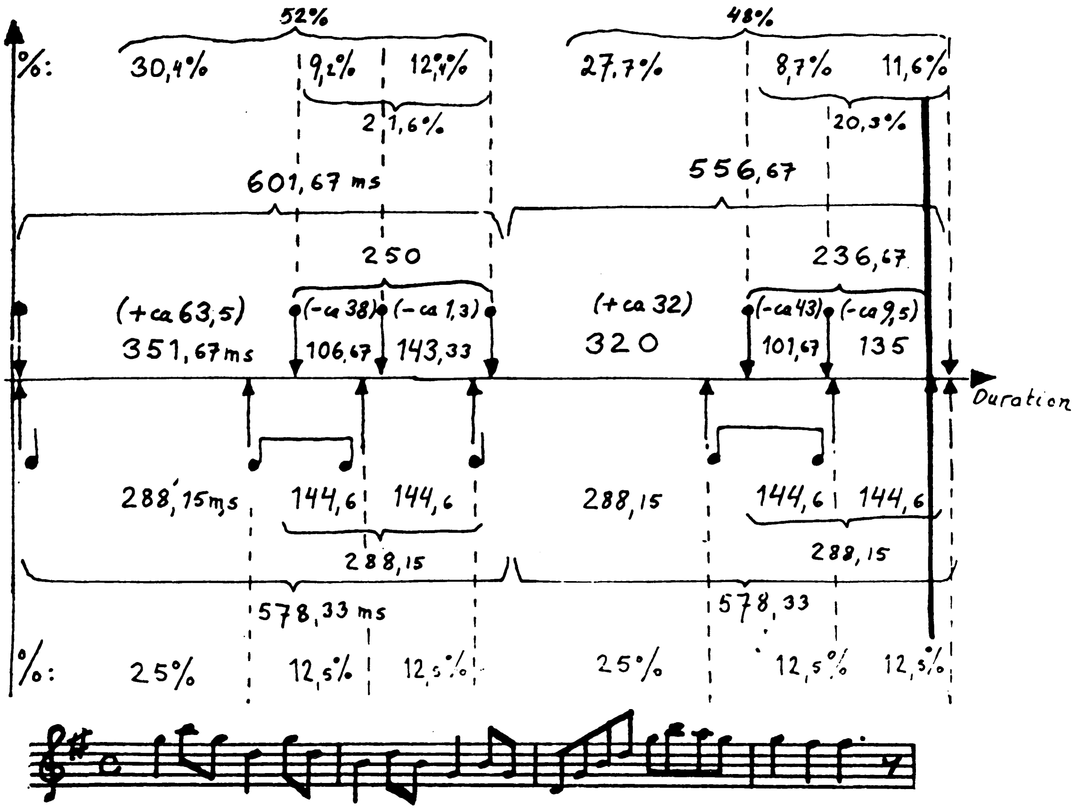



Fig. 7. Notation and scheme representing performance of a "brudmarsch" played on a key-fiddle. See text for explanation. From BENGTSSON & al. (1969), Svensk Tidskrift för Musikforskning.

that the first and third beat in the measure occupy about 30 % and 28 %, respectively. The second and fourth beats are correspondingly shortened to slightly more than 20 %, and within these beats there is a "short - long" relation between the two eighth-notes: the first one is considerably shorter than the second one (both of them yet shorter than 12.5 %). On the "half-measure" level there is thus a SYVAR-D of the type "long - short". On the "beat-level" there is on the contrary a "short - long" pattern. It seems no doubt that these patterns are of high relevance for the rhythm response elicited when you listen to this music. Using half the speed makes the duration relations described above fairly obvious for the listener. Some more comments are given in BGT (1969 p. 98).

Fig. 8 and sound example three (Sound example) comes from Swedish folk music, Evert Åhs playing a "polska" from Dalecarlia on a "spelpipa". Listen to the music and try to follow it in the preliminary note transcription in Fig. 8. Of course, un-notated music of this kind presents many problems for SYVAR-D investigations and in fact several different approaches could be taken for the computations. BENGTSSON (1968) has discussed this example in detail and made computations with the transcription in Fig. 8 as a reference basis, that is, with a triple time meter and the many "ornamental tones" considered as up-beats. Among the results the most interesting was a specific SYVAR-D pattern for the beats within measures: a "short first beat" and a "long third beat", the second beat in most cases falling between these extremes. The details may be studied by means of the duration values (in msec) which appear in the transcription.

Single examples as those above can only serve as suggestions. Of course, a far larger material is necessary in order to get statistically reliable data and to increase the representativity. During 1974-75 an investigation on a larger scale was performed in Uppsala. Six instrumentalists (pianists, flutists, clarinetists) took part in an experiment in which they played 28 different melodies taken from various music genres, all the way from Western art music to popular tunes and pop. About half of these should first be played by ear, that is, without any notation present. Then all 28 melodies were played according to two or three different notations: one the correct or original notation,

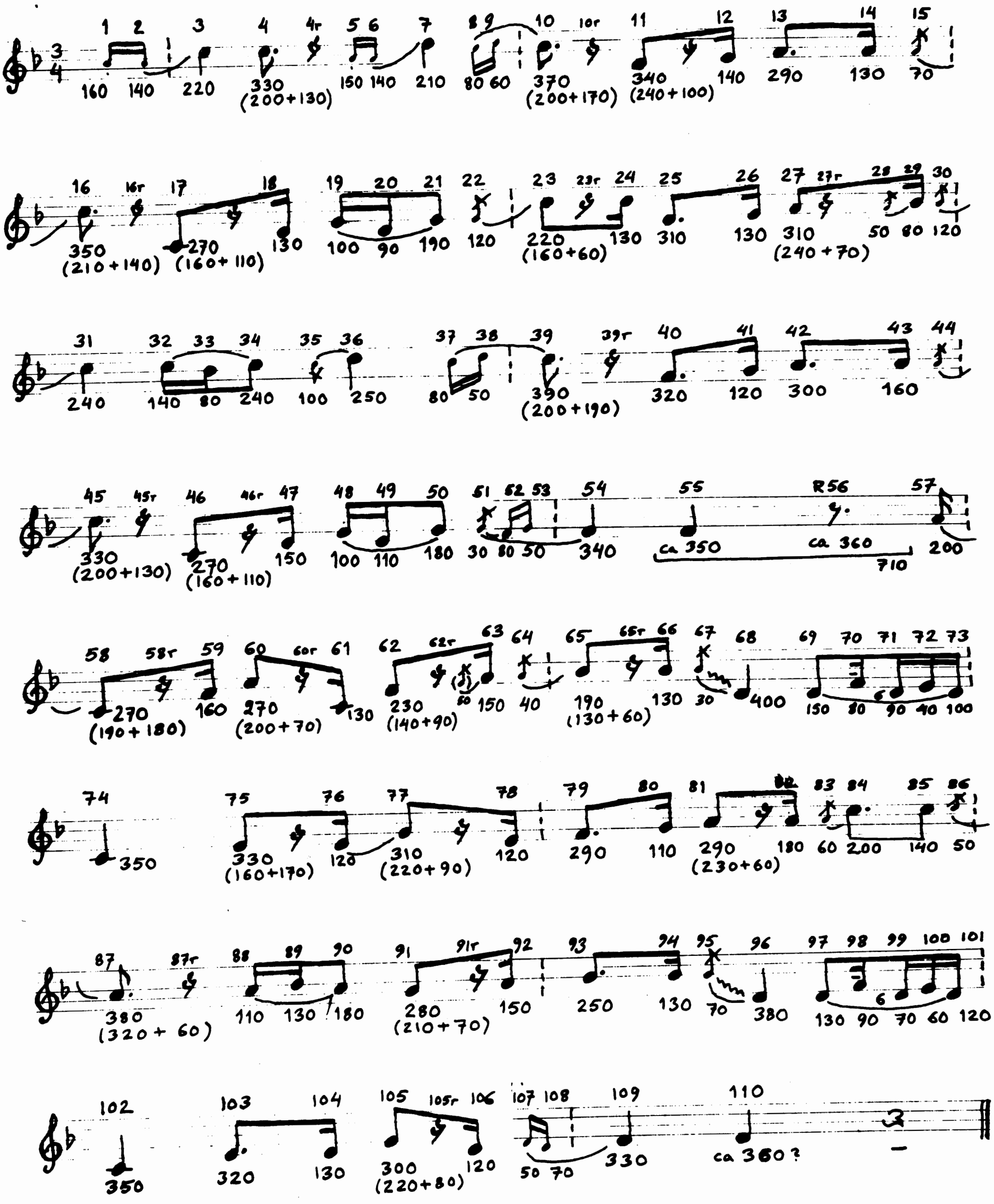


Fig. 8. Preliminary note transcription of a "polska" played on a "spelpipa". Duration values in msec are added. From BENGTSSON & al. (1972), Studia Instrumentorum Musicae Popularis II.

one or two in some other notation (no information was given as to which notation was correct or original). All performances were recorded and registered by means of the MONA analyzer resulting in a wealth of data, which presently are subjected to various analyses.

One of the greatest problems for such research is the enormous amount of data, which may be treated in many different ways and which must be strongly reduced to include only "the important data". Unfortunately, however, the criteria for what is "important" data are not self-evident but have to be successively constructed on basis of various statistical considerations, judgments from listening, and other points. Of course, use of computers is an absolute necessity. A program for computer treatment of duration values in music performance has been written and is presently under revision. (BENGTSSON 1967 b, 1970; CASTMAN & al. 1970, 1974.)

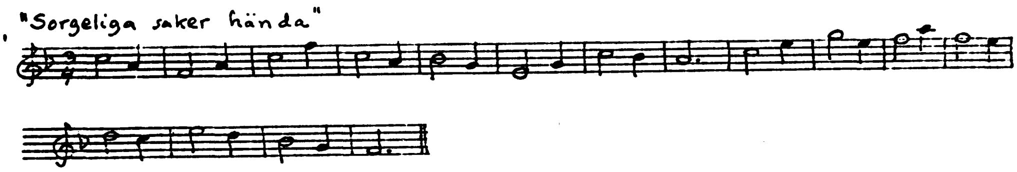



Fig. 9. Notation of the tune "Sorgeliga saker hända".

A detailed report of the conditions and results from this investigation will appear later. Only a single example is briefly commented upon here, the performance of a well-known Swedish tune from the nineteenth century "Sorgeliga saker hända". Its notation appears in Fig. 9. Even such a simple example with the same two note-values in almost all measures presents a lot of problems when interpreting the duration data from the different performances. One thing was immediately apparent. The 2:1 relation in duration between half-note and quarter-note is something which appears almost exclusively as exceptions in the performances. In average over all performances the longer tone (half-note) is about 1.75 as long as the shorter tone (quarter-note). However, the relation varies from as low values as about 1.35:1 to some cases around 2:1, different for different performers and different in different measures of the tune. Presently graphs of the type shown in Figs. 5-6 are constructed to study possible SYVAR-D's on various other levels in the performances.

Although the SYVAR project is mainly centered on analyses of performances it was clear from the outset that it also had to include judgments concerning the rhythm responses elicited by different performances. This is necessary in order to get an understanding of the psychophysical relations between different performances of a piece of music and the corresponding rhythm responses. Then another problem enters, viz. how to analyze and describe various aspects of the rhythm response. This problem will be treated in the following sections. Presently preliminary "listening tests" are performed for some selected performances from the above-mentioned experiment.

An obvious way to follow up results from SYVAR analyses is by means of synthesized sound sequences. There are by now good facilities for an accurate generation of sound sequences with desired characteristics. Hypotheses concerning SYVAR's and concerning the relations to the response side may thus be investigated by generating synthesized performances which are systematically varied in different aspects believed to be critical for the elicited rhythm response. Here, too, it is necessary to enter into the problem of how the rhythm response could be analyzed and described in an adequate way for such purposes. As yet only some few syntheses have been made, mainly for demonstration purposes.

## PART III:

### Investigations of rhythm experience

The problem how to analyze and describe the rhythm response has been repeatedly mentioned in earlier sections - how to describe the rhythm character of a Viennese waltz, of a "polska", of dances in general, of music from different epoches or cultures, of music by different composers, of different performances of a piece of music etc. In terms of Fig. 2 we now turn to an analysis on the psychological level and we will concentrate on the experiential aspects of the rhythm response, leaving most of the behavioral and physiological aspects aside.

Rhythm experience has been much discussed by music theorists as well as by psychologists. In experimental psychology methods such as analytic introspection, phenomenological description, and reproduction of various stimulus patterns have been used, in one way or another, with the purpose of analyzing and describing the components of rhythm experience. There are some interesting results but on the whole rhythm experience has proved to be very elusive to analysis. A detailed historical account is given in BGT (1969).

It seems no doubt that rhythm experience is a multidimensional phenomenon, that is, there are a large number of different characteristics or dimensions in which rhythms may differ. Such dimensions are often suggested in earlier literature and research. For instance, rhythms are said to differ in meter, in complexity, in patterns of accentuation, in various motion characteristics (rhythms may be characterized as "walking", "dancing", "rocking" etc.), and in emotional aspects (rhythms may sound "vital", "engaging", "excited" etc.). A system for description of rhythm experience would thus probably include a number of different dimensions, each of which points to a relevant aspect of the rhythm experience. Within the framework of such a system it might then be possible to describe the characteristics of various rhythms - and to describe experienced differences between rhythms - by referring to their positions in the different dimensions.

A research project on these problems was started in the late 1960's and was reported in a number of papers. (GABRIELSSON 1973a, 1973b, 1973c, 1973d, 1974a, 1974b) The purpose of the project was to try to find out some relevant dimensions of rhythm experience as investigated in different situations and with different methods. It was also expected that some information would be obtained concerning the psychophysical relations between the sound sequences used as stimuli and the resulting dimensions. This latter point was, however, secondary in importance.

Sound sequence stimuli

The sound sequences used as stimuli will for brevity be designated as S-rhythms, the prefix S referring to "stimulus" (S-rhythms thus mean sound sequences eliciting a rhythm response of some kind). They were taken from three different categories: monophonic S-rhythms, polyphonic S-rhythms, and real music.

To get the monophonic S-rhythms a pianist or a drummer was presented a number of "notated rhythms" or N-rhythms (N referring to notation) and was instructed to perform these for four measures in immediate sequence. The performances were recorded on tape and then used as S-rhythms in several experiments. In all about 70 N-rhythms were used ranging from very "simple" ones in 4/4 or 3/4 time to rather "complex" examples including syncopations, ties, and "irregular" meters as 11/8 time. Some examples are shown in Figs. 10-11. Some few of them appear as example four (Sound example).

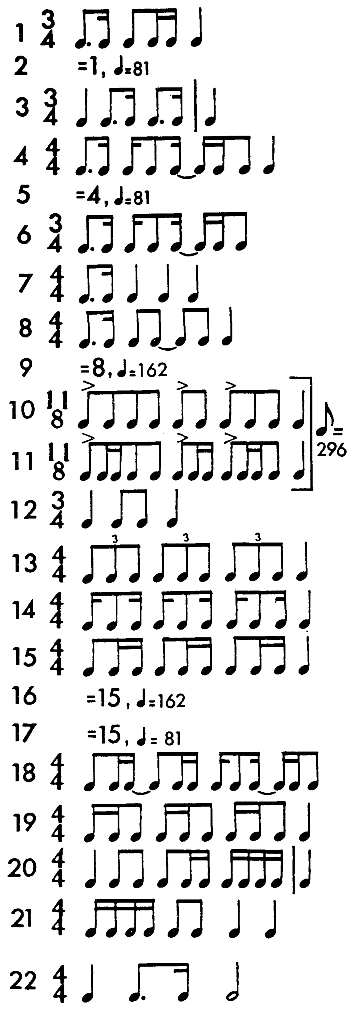



Fig. 10. Examples of monophonic N-rhythms used for experiments on rhythm experience. Unless otherwise notated, the metronomic tempo was 108 quarter-notes per minute. From GABRIELSSON (1973), Scandinavian Journal of Psychology.

The polyphonic S-rhythms were taken from an electronic "rhythm box" of the type which is now commonly built into electronic organs and the like. The "rhythm box" synthesizes spectra representing various percussion instruments (bass drum, snare drum, congas, claves, maracas, cymbals etc.) and produces sound sequences said to represent dances as waltz, foxtrot, tango, swing, rock'nroll, bossanova, rhumba, samba, and many others. The speed may be varied within a wide range. An analysis of the produced sound sequences shows that the duration relations follow the "rational-mechanical" scheme implied by the notation, that is, a sound notated as quarter-note is exactly twice as long as a sound notated as eighth-note etc. Some sounding examples appear as example five (Sound example).

The examples representing real music were taken from recordings of dance music, some 20 dance types of different origins (Viennese waltz, Slow waltz, Swedish waltz ("gammalvals"), dixieland, swing, foxtrot, slowfox, mambo, cha-cha-cha, samba, rhumba, bossanova, tango, rock'n-roll, pop, hambo, schottis, polska in two variants, and a march). Some of these appear in highly abbreviated versions as example six (Sound example). In general each excerpt lasted for about 1 1/2 minute.

Listeners and judgment methods

The listeners were for the most part musicians (with at least four years of musical performance, as a rule much more) and in some experiments also non-musicians. In order to get a description of how the listeners experienced the different S-rhythms, the listeners were instructed to make judgments according to one or more of the following three methods.

a) Similarity ratings. The subjects listened to the S-rhythms presented pairwise in succession and judged how similar they were to each other. The judgments were made on a scale from (for instance) 100, denoting "perfect similarity, no difference between the two rhythms", down to 0, denoting a "minimum similarity". The resulting data are thus each listener's judgments of the similarity between all S-rhythms taken pairwise as expressed on the 0-100 scale.

These data were analyzed according to a "distance model" which is common in multidimensional scaling. (Multidimensional scaling refers to methods designed for analysis of which dimensions or "components" are included in experiences of some kind, for instance in rhythm experience as here.) The rhythms are thought of as lying in a space. The distances between them are estimated from the given similarity judgments: the more similar two rhythms are judged to be, the nearer they lie to each other in the space and vice versa. The judged similarities between the rhythms may thus be thought of as reflecting the distances between the rhythms in a "rhythm experience space", and the axes of this space represent possible dimensions in the rhythm experience. The analysis is totally computerized.

As a simple example see Fig. 11. Six S-rhythms were used, and the corresponding N-rhythms appear in the figure. Three dimensions were expected to appear. One of them should reflect the "duration pattern" (N-rhythms 1, 3 and 5 with eighth-notes in the middle versus N-rhythms 2, 4 and 6 with eighth-notes in the beginning) and this happened in dimension I (the horisontal axis). One dimension should reflect the meter (N-rhythms 1 and 2 in 4/4 time versus the others in 3/4 time), which happened in dimension II (the axis extending in "depth"). The third dimension (the height) reflected the difference in metronomic tempo (N-rhythms 1-4 versus N-rhythms 5-6). It was also apparent from the data that different listeners gave highly different weights to these different dimensions when they judged the rhythms for similarity. In this experiment the dimensions could be reasonably predicted and it served, in fact, as a "model experiment" to test whether the method would work.

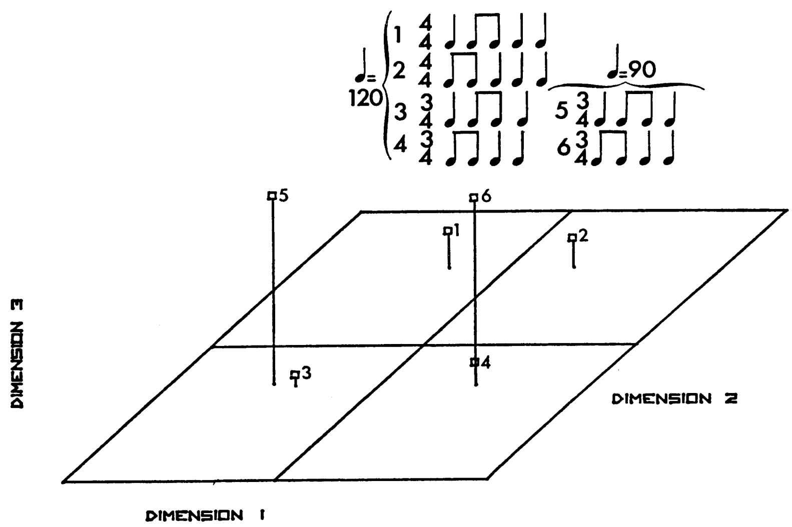



Fig. 11. Six rhythms placed in a three-dimensional "experience space" resulting from an analysis of the judged similarities between the rhythms. The distances between the rhythms (designated by small squares) reflect the similarities between them, and the three axes in the space represent possible dimensions in the experience of these rhythms. From GABRIELSSON (1974), Scandinavian Journal of Psychology.

b) <u>Adjective ratings</u>. The subjects listened to the S-rhythms, one at a time, and judged them on a large number of adjective scales. The listener thus had a list of adjectives (in randomized order) for each rhythm and put a number from 0 to 9 for each of the adjectives, 9 denoting that the rhythm in question had a "maximum" of the quality designated by the adjective and 0 denoting that it had nothing of this quality. Originally about 400 adjectives were selected and sent to a large number of musicians who rated them for their appropriateness for describing experienced rhythm. On the basis of such ratings from 141 musicians 80-90 adjectives were chosen for use in the experiments.

The data from such an experiment thus consist of each subject's ratings in each of the adjective scales for all rhythms in the experiment. These data are analyzed for their degree of co-variation (correlation) and then subjected to so-called factor analysis. In this case factor analysis aims at reducing the large number of adjective scales to a considerably smaller number of "fundamental" dimensions, which thus would represent possible dimensions in the rhythm experience. The analysis starts by computing the correlation between the different adjectives (in average over the subjects). Correlations between certain adjectives may indicate that these adjectives all reflect a more fundamental "factor" (dimension), the meaning of which can be interpreted by observing what is common in meaning for these adjectives. For instance, in Fig. 12 let adjectives No. 1, 2, 3, 6 and 7 be "intricate", "irregular", "complex", "syncopated", and "un-uniform", respectively. A high co-variation between them would lead to the interpretation that they all reflect, in one way or another, a more fundamental dimension which might be termed "complexity" or "variation". The analysis is wholly computerized, except for the final interpretation of the psychological meaning of the dimensions. The 80-90 adjectives used here could in general be reduced to 3-6 such fundamental dimensions.

c) Free verbal descriptions. In most of the experiments the subjects were also asked to make free verbal descriptions in their own words of how they experienced the rhythms (this is essentially the same as phenomenological descriptions mentioned above). The adjectives or other characterizing terms appearing in such descriptions were systematized and related to the results from the similarity and adjective ratings. In most cases these relations were very obvious.

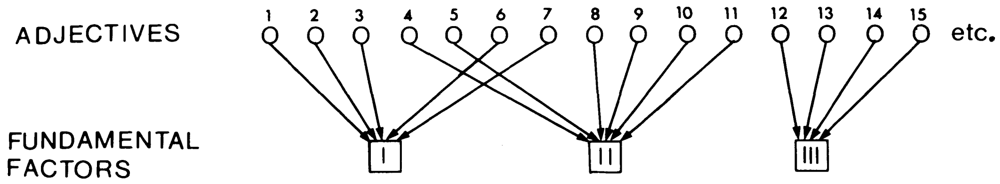



Fig. 12. A simplified representation of the meaning of factor analysis on correlations between adjectives, see text for explanation.

Since it could be expected that the judgments and the results would be dependent on the context of S-rhythms in each single experiment as well as on many other conditions, it was decided to make a series of moderately-sized experiments rather than a few larger experiments. Thus, in all 22 experiments were performed, each of them including 6-20 S-rhythms and 12-25 listeners. Detailed accounts of the methods and results in each experiment are given in the above-mentioned reports.

Resulting dimensions

In each single experiment generally three to six dimensions could be interpreted. As expected they varied with the context of S-rhythms and to a certain degree also with judgment methods (there were, however, only slight differences between the results from musicians and non-musicians). Summarizing the dimensions found in all experiments may be grouped in three classes as shown in Table 1. They are briefly described in the following (for more detailed accounts the reader is referred to the original reports).

Table 1.

Summary of suggested dimensions in rhythm experience

| "STRUCTURE" | "MOVEMENT/MOTION" | "EMOTION" |
|---|---|---|
| Meter | Rapidity | Vital - dull |
| Accent on first beat | Tempo | Excited - calm |
| Type of basic pattern | Forward movement | Rigid - flexible |
| Prominence of basic pattern, accentuation/ clearness | Movement characters (dancing - walking, floating - stuttering, rocking - knocking, solemn - swinging etc.) | Solemn - playful |
| Uniformity - variation, Simplicity - complexity | | |
| Duration pattern at different beats | | |
| Cognitive - perceptual | Perceptual - emotional | Emotional |

1. One group of dimensions is concerned with what may be called "structural properties" of the rhythms. The following dimensions belonged to this group.

a) "Meter". This refers, of course, to the perceived character of two, or three, or four etc. "beats per measure". In the present experiments there were for the most part only two or three levels in this dimension, viz. triple meter and quadruple or duple meter. It was noted that the perceived meter is no discrete variable but rather a continuous one: there are not only differences between triple meter and quadruple meter etc. but also differences in the degree of "tripleness", "quadrupleness" etc. For instance, the character of "three beats per measure" is very obvious when listening to N-rhythm No. 12 in Fig. 10 but much less obvious for N-rhythm 6 in the same figure. The same difference as regards "four beats per measure" is apparent when comparing N-rhythm 7 with N-rhythm 18 in this figure. Factors which may contribute to a weakened meter character may be syncopations, "diffuse" accents, and obscure pulse or tempo. Quite obviously many dance types differ in the degree of perceived "tripleness", "quadrupleness" etc. It was apparent, for instance, that the triple meter character of a certain "polska" from Dalecarlia was very unclear for many listeners.

b) "Degree of accent on the first beat". It was noted that the perceived accent was related to many various factors in the S-rhythms. For the monophonic S-rhythms beaten on the drum the accent was in general related to the intensity and/or the duration of the corresponding sound event. For the polyphonic S-rhythms and for the music examples other factors contribute, too, including melodic and harmonic factors specific for each example.

c) "Type of basic pattern". This refers to a perceived (or perhaps sometimes "imagined") pattern which may be thought of as "underlying" the rythm actually heard. Thus, the rhythm actually heard may be perceived as a "variation" or "filling out" of the underlying basic pattern. The basic pattern may sometimes coincide with the pattern of accented and unaccented beats but this is only one example from a variety of possibilities, which will vary depending on the specific context. For instance, in one experiment with the polyphonic S-rhythms the rock'nroll rhythm could be heard as a variant of the foxtrot rhythm, while the beguine rhythm was more reconcilable with a habanera rhythm as a basic pattern.

d) "Degree of marked basic pattern", "accentuation/clearness".

This refers to the perceptual prominence of a basic pattern, irrespective of which type the pattern is. A high level in this dimension probably corresponds to expressions as "strongly marked rhythm", "accentuated rhythm", "clear/firm/steady rhythm" etc. The opposite would thus be a "loose" or somehow "diffuse" rhythm. In the polyphonic S-rhythms and in the music examples the prominence of the basic pattern seems related to perceptually strong accents and to the prominence of the accompanying instruments which are often the main carriers of the basic pattern (consider, for example, the accompaniment given by the lower strings in a Viennese waltz or by a guitar in a Swedish waltz). For example, in one experiment here the "marked/accentuated" character was most evident for a march, hambo and Swedish waltz, while an unaccompanied polska and a slow "soft" waltz was opposite in character.

e) "Uniformity - variation", "simplicity - complexity". The meaning of this dimension is rather self-evident. In general rhythms judged as "uniform" were also judged as "simple", and "varied" rhythms as relatively more "complex". This does not mean that the two labels for this dimension could be generally substituted - for instance, two relatively "varied" rhythms may sometimes lie highly different on a "simplicity-complexity" continuum. For examples of rhythms which were judged as rather "uniform/simple" see No. 7, 12 and 15 in Fig. 10, while No. 4, 6, 10, 11 and 18 were judged as more "varied/complex". The perceived "complexity" may be dependent on the metronomic tempo, too - for instance N-rhythm No. 8 in Fig. 10 was perceived more complex in a tempo around 160 quarter-notes per minute than in a more moderate tempo.

f) In certain experiments there were one or two dimensions which could be related to the duration pattern of the sound events at different beats. An example of such a dimension is given in Fig. 11.

The above-mentioned dimensions seem to reflect cognitive-perceptual aspects of the rhythm experience. There is, of course, no sharp limit between cognitive and perceptual processes, and their relative importance in the present context certainly varies from person to person. For many subjects here there were undoubtedly a great deal of cognitive processes involved, for instance when judging the "meter" or the "complexity" of a rhythm.

2. A second group of dimensions refers to predominantly perceptual or perceptual-emotional aspects of the rhythm experience and is more or less concerned with "movement/motion properties" of the rhythms.

a) "Rapidity" and "tempo". "Rapidity" is the more general term, denoting the perceived rapidity of the rhythm "as a whole", while "tempo" denotes a special case of rapidity, viz. the perceived rapidity of the pulse. For a given metronomic tempo the perceived rapidity of the rhythm "as a whole" is related, among other things, to the number of sound events per unity of time. For instance, N-rhythm 15 in Fig. 10 is perceived as more rapid than N-rhythm 7 due to its larger sound event density. However, both of them may be perceived as having the same tempo (= rapidity of the underlying pulse). When the S-rhythms differ in metronomic tempo as well as in sound event density the perceived rapidity seems to depend on both these factors in complex ways, which probably vary in different contexts. Other factors may contribute, too. For instance, a complex pattern, masking the underlying pulse or basic pattern, might be perceived as more rapid than another pattern, equivalent in metronomic tempo and sound event density (compare No. 6 and No. 1 in Fig. 10). And in music melodic and harmonic factors very probably influence the perceived rapidity. It should also be pointed out that the tempo may often be perceived at different "levels" - for instance, at the "quarter-note level" <u>or</u> at the "eighth-note level" in the same S-rhythm.

b) "<u>Forward movement/motion</u>". This refers to a distinction between rhythms which are perceived to go on, or even accelerate, in their motion up to the first beat in the following measure, and rhythms which "stop" or "decelerate" somewhere in the measure - compare, for instance, N-rhythms 20 and 22 in Fig. 10. There are, of course, intermediate positions between these extremes: the motion may be momentarily retarded (by some longer "note-value") before going on again etc. This dimension is apparently related to the distribution of sound events over the measure.

c) "<u>Movement/motion characters</u>". This stands as a general term for a number of (bipolar) dimensions, tentatively labelled by adjective opposites as "dancing - walking", "floating - stuttering" (or "rocking - knocking", "graceful - thumping", "flexible - rugged"), "solemn - swinging", and others. Experiences of motion in relation to rhythm experience have been discussed for a long time but have proved to be very elusive to analysis. Nevertheless they are of great importance and are often referred to in writings about music and rhythm as well as in many free descriptions of the subjects in these experiments. In one experiment here a march rhythm was contrasted to various dance rhythms on a "walking - dancing" dimension. In another experiment the staccato patterns for tango and habanera in the rhythm box was contrasted to more legato-like patterns as bossanova on a "stuttering - floating" dimension.

3. There is finally a third group of dimensions mainly referring to emotional aspects of the rhythm experience and also tentatively labelled with adjective opposites.

a) "Vital - dull". One end of this dimension is characterized by adjectives as vital, lively, inspiriting, gay, energetic, full of verve and the like, the other end by adjectives as dull, restrained, heavy, hesitating, and static. In general it seems that the "vitality" of a rhythm is positively correlated with its rapitidy/tempo (the correlation is, however, not perfect).

b) "Excited - calm". One end of this dimension is described by adjectives as excited, restless, intense, tense, exciting, aggressive, hard, wild, and violent, while the other end is suggested by adjectives as calm, soft, smoothed out, relaxed, soothing, and restrained. A relatively higher rapidity/tempo, a higher loudness level, pronounced syncopations, "hard" percussion instruments etc. might be factors which are related to an "excited/intense/hard" character.

c) "Rigid - flexible". One end of this dimension is characterized by adjectives as mechanical, monotonous, static, steady, and firm, the other end by adjectives as flexible and free. It may be thought of as a kind of emotional counterpart to the "uniformity-variation" dimension and appeared especially in some experiments with S-rhythms from the rhythm box, some of which undoubtedly sound very "rigid" in character.

d) "Solemn - playful" was a dimension suggested in some analysis of experiments with real music.

Concluding comments

The dimensions discussed above should only be understood as suggestions for possible dimensions in rhythm experience, which have to be tested for their validity and generality with other stimuli, other subjects, other methods etc. The distinction between cognitive, perceptual, and emotional aspects of the rhythm experience conforms to much psychological thinking but should be taken with caution since the borderlines are diffuse and since several of the dimensions may be classified in different ways. The psycho-physical relations behind the dimensions are only loosely suggested here and there in the text above.

These reservations being made, it seems evident, however, that many of the dimensions agree with dimensions mentioned in much literature by musicologists, music theorists, musicians, and others. They may serve as a basis for continued research on description of rhythm experience and related matters. Numerous revisions will undoubtedly occur. Presently they may be investigated for their use in connection with listeners' judgments of different performances in the present large experiment within the SYVAR-D project as described above.

Referring finally once again to the principal scheme for rhythm research in Fig. 2 it may be said that the ongoing rhythm research in Uppsala has resulted in some new information both as regards the understanding of the rhythm response and as regards the properties of musical stimuli or performances which are relevant for eliciting the "proper" or intended rhythm character of the music in question. The work described in this paper is only a beginning but it is hoped that some crucial issues have been made clear and that the work will contribute to an increased interest and increased research in the fascinating field of musical rhythm.

References:

BENGTSSON, I. (1966): "Taktstrecken och punkteringspunkterna i vårt öra", pp. 11-33 in Gottfrid Boon-Sällskapets Minnesskrift, Stockholm.

BENGTSSON, I. (1967a): "Anteckningar om rytm och rytmiskt beteende med särskild hänsyn till musik", Del I, Uppsala (Mimeo).

BENGTSSON, I. (1967b): "On Melody Registration and 'Mona'", in Elektronische Datenverarbeitung in der Musikwissenschaft, ed. H. Heckmann, Gustav Bosse Verlag, Regensburg.

BENGTSSON, I. (1968): "Anteckningar om 20 sekunder svensk folkmusik", in Festskrift til Olav Gurvin, Harald Lyche, Oslo.

BENGTSSON, I. (1970): "Studies with the Aid of Computers of 'Music as Performed'", pp. 443-444 and 448-449 in Report on the IMS Congress Ljubljana 1967, Kassel & Ljubljana.

BENGTSSON, I. (1975): "Empirische Rhythmusforschung in Uppsala", pp. 195-220 in Hamburger Jahrbuch für Musikwissenschaft, Bd. 1, 1974, Verlag K.D. Wagner, Hamburg.

BENGTSSON, I., GABRIELSSON, A., and THORSÉN, S.M. (1969): "Empirisk rytmforskning", Svensk tidskrift för musikforskning, 51, pp. 49-118.

BENGTSSON, I., TOVE, P.A., and THORSÉN, S.M. (1972): "Sound Analysis Equipment and Rhythm Research Ideas at the Institute of Musicology in Uppsala", pp. 53-76 in Studia Instrumentorum Musicae Popularis, Vol. 2, ed. E. Stockmann, Musikhistoriska Muséet, Stockholm.

CASTMAN, B., LARSSON, L.E., and BENGTSSON, I. (1970, 1974): "Ett dataprogram (RHYTHMSYVARD) för bearbetning av rytmiska durationsvärden", Institutionen för Musikvetenskap, Uppsala (Mimeo).

GABRIELSSON, A. (1973a): "Similarity Ratings and Dimension Analyses of Auditory Rhythm Patterns, I ", Scand. J. of Psychology, 14, pp. 138-160.

GABRIELSSON, A. (1973b): "Similarity Ratings and Dimension Analyses of Auditory Rhythm Patterns, II ", Scand. J. of Psychology, 14, pp. 161-176.

GABRIELSSON, A. (1973c): "Adjective Ratings and Dimension Analyses of Auditory Rhythm Patterns", Scand. J. of Psychology, 14, pp. 244-260.

GABRIELSSON, A. (1973d): "Studies in Rhythm", Acta Universitatis Uppsaliensis: Abstracts of Uppsala Dissertations from the Faculty of Social Sciences, 7.

GABRIELSSON, A. (1974a): "Performance of Rhythm Patterns", Scand. J. of Psychology, 15, pp. 63-72.

GABRIELSSON, A. (1974b): "An Empirical Comparison Between Some Models for Multidimensional Scaling", Scand. J. of Psychology, 15, pp. 73-80.

# SINGING AND TIMBRE

by

Johan Sundberg, Center for Speech Communication Research and Musical Acoustics, Royal Institute of Technology (KTH), Stockholm, Sweden.

## Introduction

The aim of acoustics of music is to develop explanatory theories of musical sounds. This is a very old branch of science, perhaps the oldest dealing with music. Actually, Pythagoras was one of the first in the field. He revealed relationships between the sound of a two-tone-chord and the sounding lengths of the strings, which he used in order to generate the two tones. He found that the more complicated the string-length ratio was, the more dissonant or "apart-sounding" was the sound. Inversely, consonant chords were obtained only when the string length ratios could be expressed in terms of small integers, such as 1:2 (octave), 2:3 (fifth), 3:4 (fourth). Evidently, Pythagoras combined physics and music as he established these relationships between the string-length ratio and the sound.

In the centuries after Pythagoras the progress in the acoustics of music was slow. It took until the 19th century before decisive contributions were made. LORD RAYLEIGH provided the basic theoretical framework of modern acoustics in his classical Theory of Sound (1878). Hermann VON HELMHOLTZ did important pioneer work in the acoustics of music when he determined the acoustic properties of various musical sounds by means of ingenious inventions and experiments. His great book (1863) had the title Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik, thus stressing the interdisciplinary character of the acoustics of music, and von Helmholtz was convinced that music could be fully understood only in the combined light from physics and the physiology of hearing. In this way he may be considered the father of modern acoustics of music. If so, Pythagoras should be called the grand-father.

In spite of the contributions from Lord Rayleigh and von Helmholtz, the tools for handling and analyzing acoustic signals were tedious to use in the 19th century. However, towards the middle of the present century the picture changed radically. Electronic amplifiers, taperecorders, spectrum analyzers and other electro-acoustical devices were invented. Storing and analyzing sounds became easy tasks. Thus, the conditions for a fruitful combination of acoustics and music improved substantially.

These new technical possibilities have not been neglected. A number of research centers have been established in the last few decades. The fact that several of them are affiliated with the musical-instrument industry mirrors a primary goal of these centers. In manufacturing musical instruments on an industrial scale, knowledge is required as to how an instrument of high quality can be obtained as rapidly and inexpensively as possible. The condition then, is a <u>theory of the instrument</u>. But for practical purposes such a theory is insufficient. It just tells us how to build the instrument in order to obtain an instrument generating sounds with certain <u>acoustic</u> properties, but it does not tell what acoustic properties that characterize a high quality instrument. It is necessary to include a <u>theory of instrument quality</u> in an explanatory theory of musical sounds, i.e. to consider the musical function along with the auditory perception of the sounds from the instrument. This is true in all research in the acoustics of music. The present paper will demonstrate this by extending the perspective beyond the scope of musical-instrument industries and consider recent contributions to the research in singing and related research in auditory perception.

## <u>Theory of voice</u>

Let us first recapitulate the basic theory of voice production and thereby restrict the presentation to non-nasalized vowels, see e.g. FANT, (1968). The voice organ can be regarded as a system consisting of an oscillator (the vocal folds) and a tube resonator (the pharyngeal and buccal cavities, in short called the <u>vocal tract</u>). The tube resonator has a very important function. If a sine wave sweeping from low to high frequencies is fed with constant amplitude into the resonator, an amplitude which is strongly frequency-dependent can be observed at the opposite end of the resonator. As seen in Fig. 1 the amplitude reaches

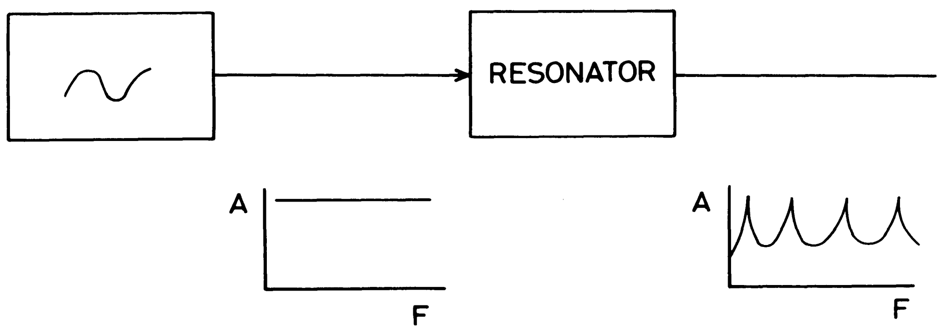



Fig. 1. Schematic illustration of resonator properties. A sine wave making a glissando from low to high frequencies, F, with constant amplitude, A, is fed into a resonator. The resonator output is characterized by a highly frequency-dependent amplitude. Those frequencies which give maximal amplitudes in the resonator output are called the resonance frequencies, or, if the resonator is the vocal tract, the formant frequencies.

maximum at certain frequencies called resonance frequencies in general and <u>formant frequencies</u> in the case of the voice organ. Fig. 2 demonstrates in a schematical fashion how the formants shape the sound radiated from the lips. The vibrating vocal folds generate a complex sound consisting of a great number of harmonic partials. Those partials, which lie closest to a formant in frequency, are most emphasized by the resonator and are radiated with greater amplitudes from the lip opening than other partials lying further away from a formant. The formant frequencies are determined by the positions of the articulators, i.e. the lips, the jaw, the tongue, and the larynx. Also, the vocal tract morphology is important, i.e. the individual peculiarities of the vocal tract. The signal from the oscillator is basically the same regardless of the vowel produced. What differs between vowels is the combination of formant frequencies. As a formant enhances those partials, which lie close to it in frequency, each vowel is characterized by strong partials in certain frequency regions regardless of the pitch. These frequency regions correspond to the formant frequencies. For instance, the vowel (i) as in the word "beat" is characterized by strong partials in the

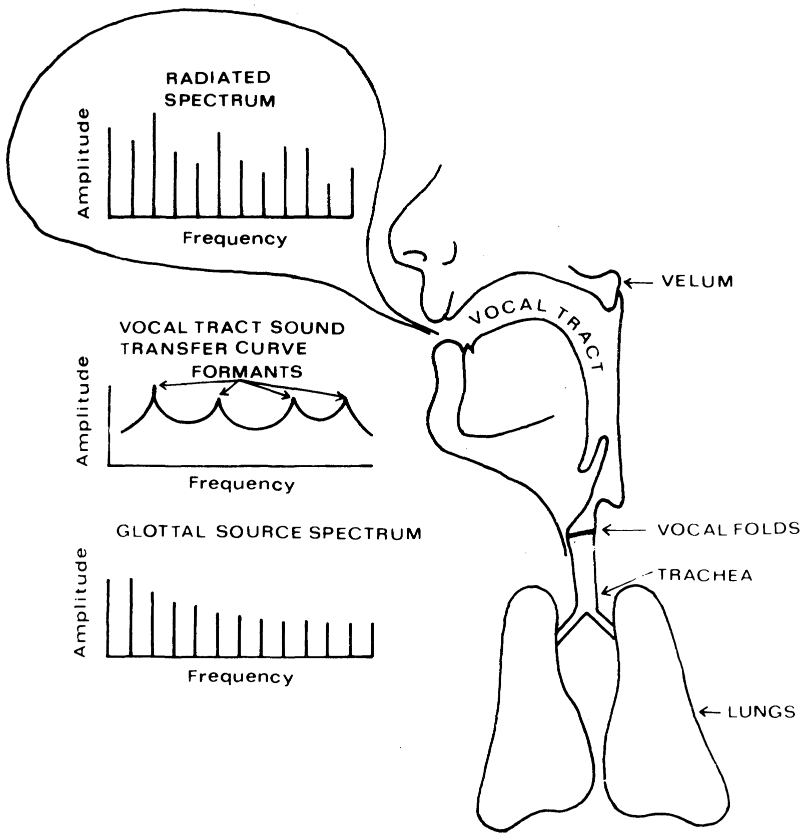



Fig. 2. Schematical illustration of the production of voiced sounds. The lungs provide an overpressure under the glottis which causes the vocal folds to vibrate. This vibration generates a complex tone consisting of a number of harmonics, the amplitudes of which decrease with the number of the harmonic. This complex tone propagates through the vocal tract which is a resonator. Those harmonics, which lie closest to a formant frequency (e.g. the second and the sixth) are radiated from the mouth opening with higher amplitudes than other harmonics. From SUNDBERG (1974b), Musiklivet Vår Sång.

frequency region around 250 and 2000 Hz, respectively, because the two lowest formants are approximately located to these frequencies. The corresponding formant frequencies for the vowel (a) (as in "part") are approximately 500 and 1000 Hz.

It is rather simple to construct an electrical model of the voice organ. Such a model is composed of electronical components and is called a synthesizer. The oscillator of the voice corresponding to the vibrating vocal folds is substituted by a signal generator, and the vocal tract by a set of electrical resonance circuits connected in a sequence, or cascaded. Once the four or five lowest formant frequencies are known, the frequency curve of the vocal tract is in essence predictable, and such frequency curves are provided in an analoguous way by the electrical resonance circuits in the synthesizer. Thus, if the resonance frequencies of such a synthesizer are adjusted to the formant frequencies of a given vowel sound, the circuit system offers the same frequency curve as a vocal tract characterized by these formant frequencies. Therefore, if we send a well-defined complex tone through our synthesizer, which we have previously adjusted to the formant frequencies of a given vowel, we will obtain this very vowel provided that our complex tone is acoustically identical to the tone generated by the vocal folds.

The soprano's jaw opening

Let us now return to the singing voice and first consider the case of sopranos. In comparing their singing and speech, three observations can be made. In singing, (1) the sound is generally much louder (2) the jaw opening seems to depend on the pitch rather than on the vowel, and (3) the vowel intended is often difficult to identify, particularly in high notes. (ONDRÁČKOVÁ, 1973; SIMON & al., 1972, STUMPF, 1926.) It is incumbent on acoustics of music to explain why speech and singing differ in these respects.

Fig. 3 illustrates the typical dependence of the lip and jaw opening on the pitch: the higher the pitch, the wider the opening. An increase of the jaw opening is particularly influential on the frequency of the first formant, (LINDBLOM & SUNDBERG,1971). Actually, it has been shown that the singer adjusts the frequency of her first formant to the frequency of the pitch, i.e. to the fundamental frequency. This seems

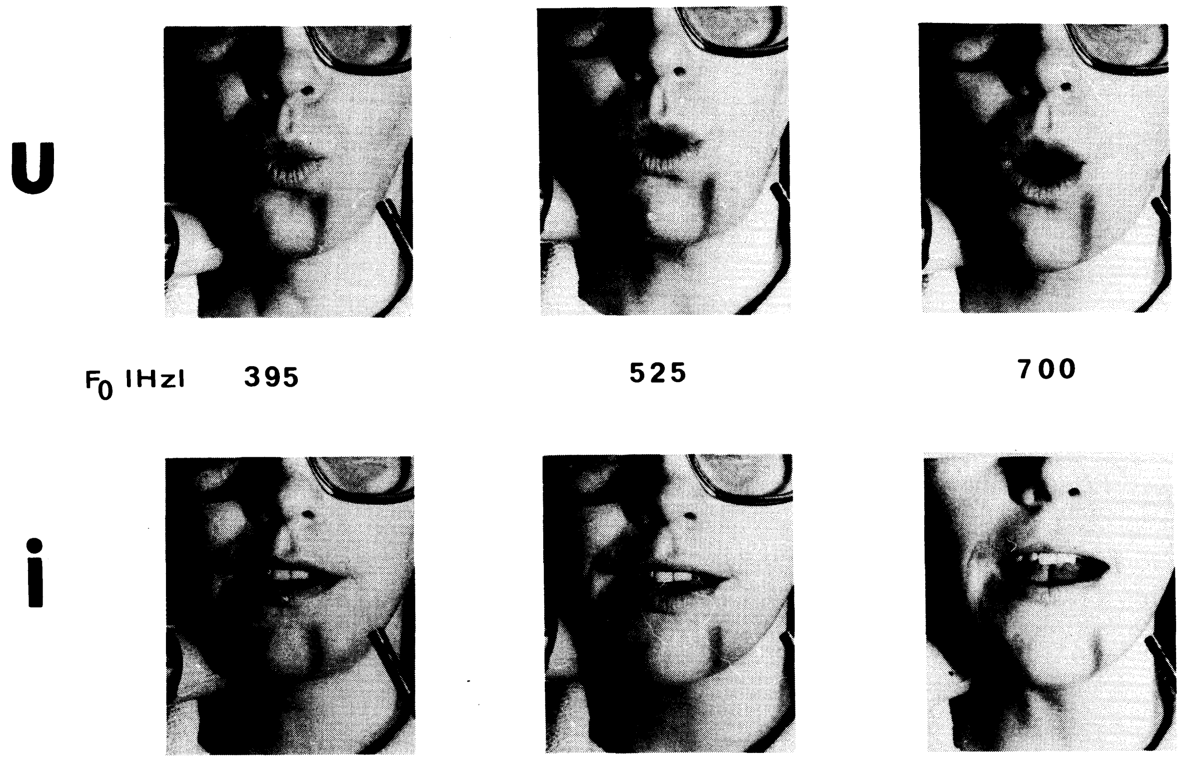



Fig. 3. Photos of the lip opening of a soprano singing the vowels (u) and (i) (upper and lower series) at the fundamental frequencies ($F_0$) indicated. The lip and jaw opening are seen to increase with rising fundamental frequency.

to be the acoustical purpose of the pitch-dependent jaw opening. The next question is why a soprano makes this frequency match. If she sings a high-pitched note, e.g. an $A_5$, the partials of the complex tone generated by the vocal-fold vibrations fall on multiples of 880 Hz: $1 \cdot 880 = 880$, $2 \cdot 880 = 1760$, $3 \cdot 880 = 2640$ Hz and so on. In the case of an (u) (as in boot) the normal frequency of the first formant is 300 Hz, approximately. With a jaw opening normal for speech, the singer's lowest formant would then be located at a frequency far below that of the lowest partial. Then, the possibility of a resonance effect would be wasted, as the resonance would occur in a frequency region, where there is no sound to give resonance to. With a suitably wide jaw opening she can adjust the resonance to the frequency of the fundamental, i.e. the lowest partial. The result is that this partial gains considerably in amplitude, which makes the sound loud, as illustrated schematically in Fig. 4. If she next sings the note $F_5$, the fundamental frequency is close to 700 Hz. If she keeps the same jaw opening and first formant frequency as she had for $A_5$ at 880 Hz, a marked drop in loudness would occur, as the first partial moves away from the resonance maximum, cf. Fig. 2. To compensate for such amplitude variations by simply raising vocal effort is not a desirable solution to this dilemma, as it would place a strain on the vocal folds. The entire problem can be avoided if the soprano reduces the jaw opening so that the first formant again matches the new fundamental frequency.

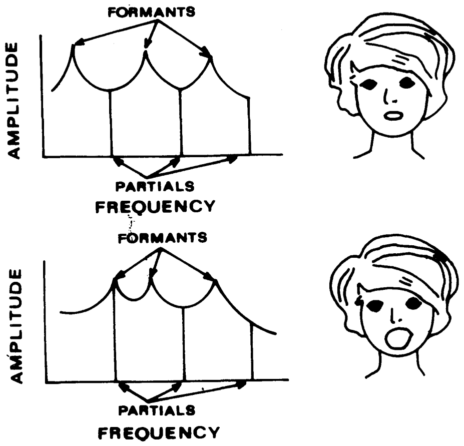


Fig. 4. Schematical illustration of the formant strategy in soprano singing at high pitches. In the upper case the soprano uses a small jaw opening. Then, the first formant appears at a frequency far below the frequency of the lowest harmonic in the vowel sung. The result is a rather low amplitude of that harmonic. In the lower case the jaw opening is adjusted so that the first formant matches the frequency of the fundamental. The result is a considerable gain in amplitude of that partial. From SUNDBERG (1974b). Musiklivet Vår Sång.

This effect is illustrated in the first sound example. First it gives the soprano's own production of a vowel at four pitches. The slightly synthetic impression is due to the fact that the natural onsets and decays of the tones have been made uniform by an electronic gate. Then follows synthesized versions of the same sounds. Here, then, the formant frequencies have been adjusted to the values that the soprano used herself; for instance, the first formant frequency follows the fundamental frequency. Finally comes another set of synthesized vowels, but here the formant frequencies are kept constant, regardless of the pitch. The formant frequencies are those which the soprano used in singing the vowel at the lowest pitch in the first example. The example thus demonstrates what would happen if the soprano kept the same articulation regardless of the pitch she sung. The result would be that the loudness would vary considerably between pitches.

Varying the first formant so that it tracks the fundamental frequency by adjusting the jaw opening raises negligible demands on effort, and at the same time it gives maximum strength to the sounds produced. Thus, the pitch-dependent articulation illustrated in Fig. 3 can be explained with reference to vocal economy. It should be expected to occur in a trained singer above that point in the scale, where the fundamental frequency of the pitch being sung passes above the normal frequency of the first formant. Presumably, this situation occurs also in altos and tenors. (SUNDBERG, 1975)

As mentioned, each individual vowel is associated with a given combination of formant frequencies. Thus, by abandoning the formant-frequency values normal in speech, the soprano also abandons the quality of the intended vowel. This seems to be of minor importance, however, because when the pitch is high, the vowel quality cannot be maintained even with the correct formant frequencies. Thus, the soprano seems to choose the best possible solution.

These results appear to explain why a soprano's speech and singing differ in the three respects mentioned. It is noteworthy that this explanation is based on evidence provided by a study of acoustical features, articulation, and perception. It seems that the possibility of answering the pertinent questions presupposes involvement of all these three aspects.

A partial reaches a maximum relative amplitude when it matches a formant in frequency. Other things being equal such partials dominate the spectrum. Thus, in the higher pitches in a soprano we would expect to find the spectrum dominated by the fundamental. In an acoustic study of female registers, LARGE (1968) found that the female chest and mid-registers differed with respect to the balance between the fundamental and the higher harmonics. Register differences are generally assumed to stem only from the vocal-cord function as opposed to articulation. In view of the results mentioned above it seems worthwhile to examine more thoroughly the role of articulation in what is called registers in the female voice.

The male voice and its "singing formant"

To economize with one's own voice is absolutely imperative for all professional singers. Only those who manage to do this and still be heard will survive as singers. However, male singers cannot use the same solution as the soprano, because a male singer sings pitches with a fundamental frequency lower than the normal first-formant frequency. The solution male singers use involves modest but characteristic changes of the vowel qualities normally encountered in speech. For instance, the vowel (i) is shifted towards the vowel (y) (similar to the German für) and the vowel (ɛ) (as in head) towards the vowel (ɜ) (as in heard). (APPELMAN, 1967) The vowel quality becomes more "dark", slightly resembling the effect we experience when we listen to someone who speaks and yawns at the same time. Often the term "covering" seems to refer to this phenomenon. The timbral effect seems to be associated with a lowered position of the larynx and an expansion of the lower pharynx cavity, see eg. LARGE (1972).

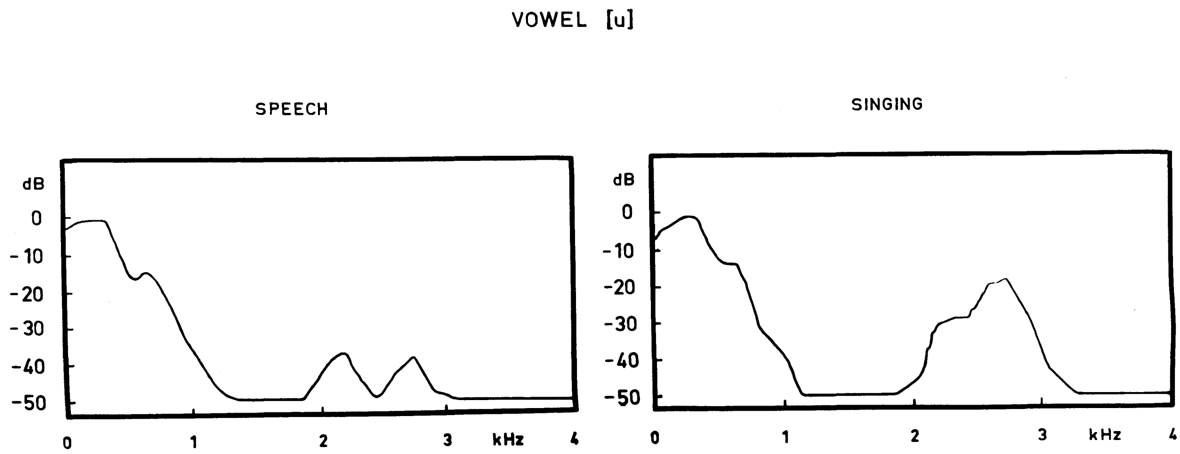



Fig. 5. Spectrum contours (envelopes) of the vowel (u) spoken (left) and sung (right) by a professional opera singer. The amplitudes of the harmonics between 2 and 3 kHz give a marked peak in singing as compared with speech. This peak is called the "singing formant", and it appears to characterize all voiced sounds in male professional opera singers. From SUNDBERG (1974a), Journal of the Acoustical Society of America.

Fig. 5 illustrates the typical spectral differences between male speech and singing as found in the vowel (u). The sound level near 3000 Hz is seen to be 20 dB higher in singing than in speech. Vocal pedagogues tend to speak about "head voice" or "singing in the mask" to describe this effect. (GIBIAN,1972) This spectrum peak is frequently called the "singing formant" and its presence, regardless of vowel and dynamics, is considered a quality criterion for singers. (WINCKEL,1953) The "singing formant" can be explained acoustically with reference to an extra formant associated with the larynx tube. The larynx tube can behave as an independent resonator provided that its outlet in the pharynx is less than 1/6 of the pharyngeal cross-sectional area. It appears that this condition can be met if the larynx is lowered, because such a lowering seems to expand the pharynx. The extra larynx-tube formant can be tuned to a frequency close to 3 kHz, i.e. between the frequencies of the third and fourth formants in normal speech. The acoustic effect on the spectrum envelope is illustrated in Fig. 6. The envelope gets a peak of 20 dB close to 3 kHz when the extra formant is present.

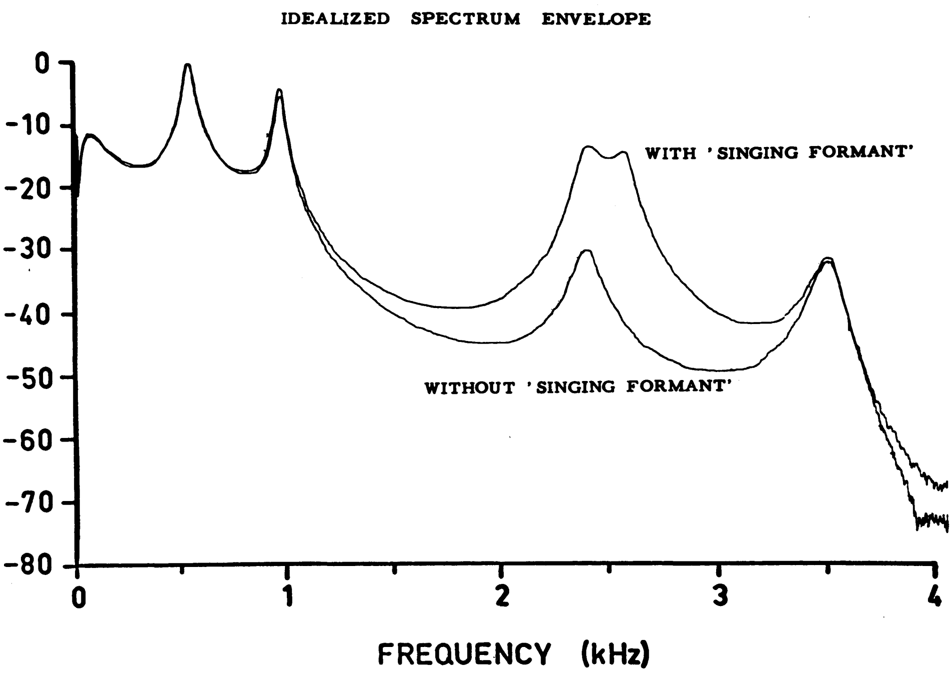



Fig. 6. Acoustic explanation of the "singing formant". The figure shows frequency curves of the vocal tract with four and five formants. The three lowest and the highest formant frequencies are the same in both cases and the extra formant is close to the third. The consequence of adding the extra formant is that the resonance curve gains about 20 dB in the frequency region of the extra formant. This means that the harmonics in that frequency region will be 20 dB stronger due to the extra formant, other things being equal.

It can be shown that the lengthening of the vocal tract and the expansion of the lower pharynx, resulting from a lowering of the larynx, affect the lower formant frequencies as well. These changes account for the main shifts in vowel quality between normal speech and professional male, singing. (SUNDBERG, 1974a)

The lowest part of the pharynx is called the sinus piriformes. It is a pear-shaped cavity which surrounds the lateral and posterior parts of the larynx tube. Acoustically the sinus piriformes act as a side branch or "appendix" of the main vocal tract extending from the vocal folds to the lip opening. Such a side branch absorbes the sound energy transmitted through the vocal tract at the frequency of its resonance. The result is that very little sound energy is radiated at that frequency, which consequently appears as a sudden dip, or minimum in the frequency curve of the vocal tract. The dip due to the sinus piriformes is between 3 and 4 kHz, and in professional singers a dip in the spectrum envelope is often observed in this frequency region. Thus, this dip presumably reflect expanded sinus piriformes. WINCKEL (1952) claims that good voices have fewer high-frequency partials than poor voices. It is likely that this observation is related to the spectrum-envelope dip caused by the lowering of the larynx, because a lowered larynx seems to characterize good male singers.

The next question is why this increase in the spectral amplitude is so desirable in singing. The answer seems to be dependent on the acoustical conditions under which a singer works: he is generally accompanied by an orchestral accompaniment which may be quite loud. The sound of the orchestra is strongest in the frequency region around 450 Hz, as seen in Fig. 7. Towards higher frequencies, the sound is weaker. The singer's high-amplitude partials near 3 kHz form a marked peak in the spectrum, as can be seen in the same figure.

The perceptual effect of this is that the singer's voice is much easier to discern against a background of an orchestral accompaniment. This is demonstrated in Sound example two. First a singer sings some bars of a song six times. The first, third, and sixth times the acoustic effect of his extra formant (i.e. the larynx tube resonance) has been eliminated by electronic means, a so called inverse filter.

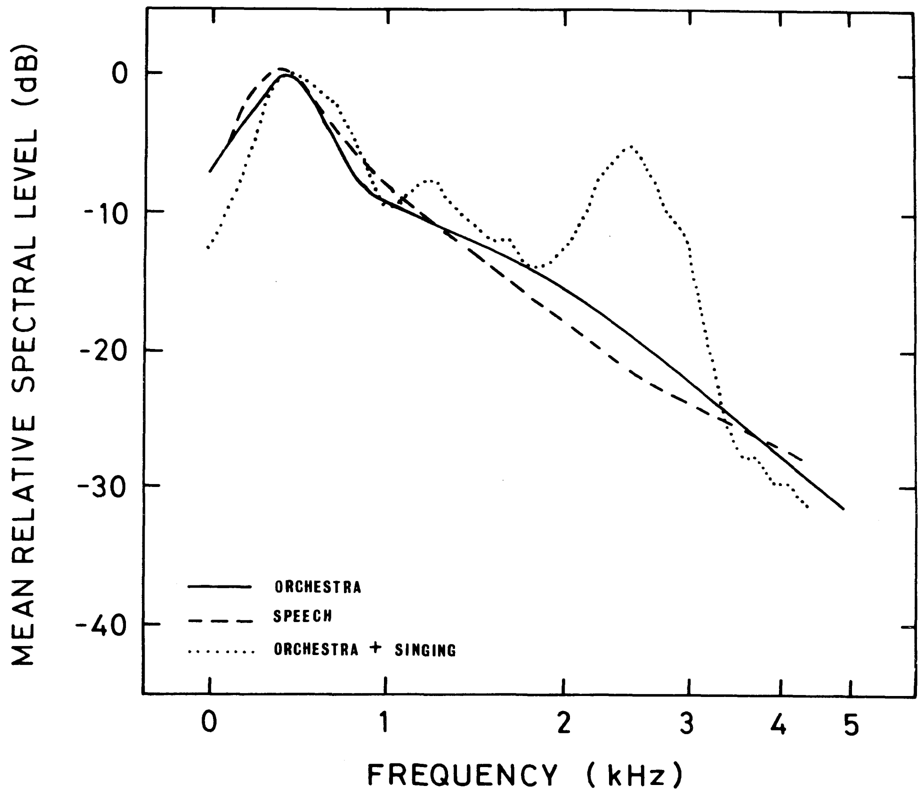



Fig. 7. Averaged spectra of the sound of a symphony orchestra (solid curve), normal speech (dashed curve), and a singer accompanied by an orchestra (dotted curve). The average spectrum of the orchestra is strikingly similar to that of normal speech. The "singing formant" is seen as a broad peak being the only major difference between orchestra with and without a singer soloist. A singer's voice accompanied by an orchestra is easier to discern when it has a "singing formant" than when it lacks it, as is the case in the normal speaking voice. From SUNDBERG (1972), Report of the 11th Congress of the International Musicological Society.

It would correspond to a singing without a "singing formant" (but with a lowered larynx). Then follows noise which has the same distribution of spectral energy as the average sound from a modern symphony orchestra. Next, this same noise is presented together with the singing with and without the "singing formant". In listening it becomes apparent that the singer's voice can be discerned more easily when the voice has the "singing formant".

We may then conclude that, once again, vocal economy seems to be the leading principle being the motivation of the special singing technique which generates vowels with high-amplitude partials near 3 kHz. This conclusion is supported by the fact that such a technique does not seem to be adopted when the accompaniment is relatively soft in dynamic level, such as lute or guitar, or when there is a sound-amplifying equipment at hand to take care of the audibility problems, as in pop music, for instance. It can also be mentioned that the voice of a trained opera singer will be heard as an individual in a choir. This suggests that voice training should have partly different goals depending on whether the student aims at a career as a choir singer, an opera soloist, or a microphone-aided soloist.

The investigations of male singing reviewed above explain why spoken and sung vowels differ in the way illustrated in Fig. 5. As in the case of the pitch-dependent jaw opening in sopranos, the explanation considers the acoustics of vowel production, the articulation, the musical function of a solo singer, and the auditory perception. It appears that our understanding of why male speech and singing differ in the way discussed would be incomplete if any of these aspects were disregarded in the explanation.

## What is voice timbre?

The problem of audibility is not the only secret in singing, which musical acoustics should elucidate. Another interesting aspect is the personality of the voice. Next we will consider the research done in what may be regarded as the first step towards a description of personal voice characteristics, namely the classification of the male voices into bass, baritone, and tenor. The timbral differences are independent of the vowel and pitch involved, i.e. a tenor, baritone, and bass singing the same vowel on the same pitch differ with respect to voice timbre. The acoustic differences between these voice types have been examined. It was revealed that the formant frequencies are decisive factors in voice classification. For most vowels it was found that the lower the voice class is, the lower are the formant frequencies. These differences are easy to demonstrate and synthesize. Sound example three consists of the vowels (a, i, u) synthesized with the formant frequencies which were found to be typical of a tenor, a baritone, and a bass, respectively.

It seems clear that the set of formant frequencies is a very important acoustical correlate of voice type. (CLEVELAND,1975)

What, then, is the origin of these formant-frequency differences between basses, baritones, and tenors? It is a well-established fact that female and male voices differ with respect to formant frequencies, and that these differences stem from dissimilarities in the mouth and pharynx-cavity lengths. Thus, the pharynx and mouth have been estimated to be on the order of 25 % and 15 % shorter in females than in males, respectively. (FANT,1973) Fig. 8 compares the formant-frequency differences between female and male subjects with those found between

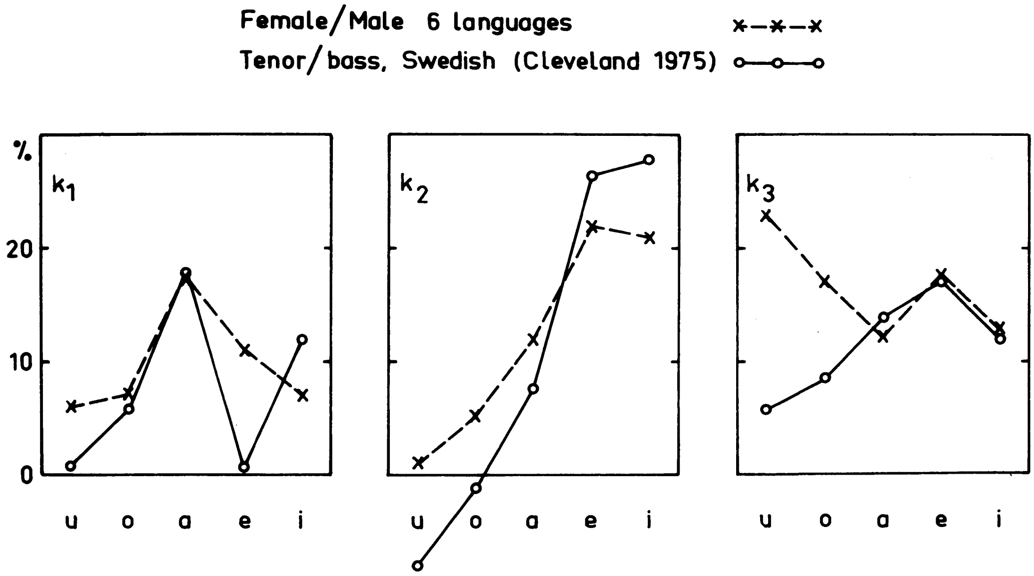



Fig. 8. Percentual differences in the three lowest formant frequencies $F_1$ (left), $F_2$ (middle), and $F_3$ (right) between a bass and a tenor (dashed line) and male and female voices (solid line). The values differ between the vowels indicated, but the tenor/bass and the female/male differences are seen to be very similar. This suggests that the reasons for the differences are similar, i.e. that it is the vocal-tract shape and length which account for the bass/tenor formant-frequency differences.

tenors and basses. The differences vary with vowel but are strikingly similar between sex and voice type. This strongly suggests that the formant-frequency differences between voice types have an origin analogous to those between sexes. This would mean that tenors and basses differ as regards the vocal-tract dimensions. Thus, it seems that voice classification with respect to voice type may be possible to base on vocal-tract dimensions. However, another factor necessary to voice classification is the pitch range. Again inferring from differences observed between female and male subjects, the pitch range may be correlated with the vocal-cord length. (HOLLIEN, 1960) Possibly, in the future, voice classification can indeed efficiently be assisted by purely objective data on the morphology of the voice organ. In any case we have good reasons to postulate that the vocal-tract dimensions should match the pitch range in good voices.

Apparently, this is contradicted by the fact that there are many singers whose voice type changes, e.g. from baritone to tenor. The explanation would be that a person is able to change his vocal-tract length, and, particularly, his pharynx length. The effective length of the mouth can be shortened if the mouth corners are retracted, and the pharynx length depends on the larynx height. Thus, it would be interesting to know if a change of the voice classification is accompanied by a change of the vocal-tract length. This is a question for future research.

Even if two singers possessing the same voice classification sing the same vowel on the same pitch, we will hear a timbre difference which enables us to hear that this is singer X and that is singer Y. As a matter of fact, a good deal of these personal voice characteristics seems to be due to the formant frequencies. This is supported by <u>sound example four</u> giving a real and a synthesized vowel of a baritone singer. It seems that the essence of the singer's personal voice quality is obtained when his formant frequencies are reproduced in the synthesis. This suggests that the signal generated by the vocal folds are rather similar between singers and not very important in determining the differences between singers. (SUNDBERG 1973) The predominating factors seem to be the formant frequencies, which are the acoustic consequences of the morphology and the articulatory habits of the singer.

Obviously, singing is constituted not only by sustained vowels, but also by transitions between vowels and between pitches. Trained singers have been found to perform wide melodic pitch changes more rapidly than untrained voices. (SUNDBERG, forthcoming) Moreover the speed of the fundamental frequency undulations due to the vibrato has been found to be related to the maximum speed of pitch change. (VENNARD, 1971) This is only one of several signs of the mastery which a singer has to develop in the operation of the vocal folds. And, probably, the manner in which a singer performs pitch transitions belong to the very typical personal characteristics of his voice.

Rough and smooth timbre

As mentioned, the timbre differences between bass, baritone, and tenor appear to depend on the formant frequencies to a large extent and, consequently, on the vocal-tract dimensions. Therefore it would be logical to assume that the identification of the sex of a speaker or singer should rely on the formant frequencies as well, because, as mentioned, the vocal-tract-dimension differences between females and males lead to formant frequency differences. However, according to COLEMAN (1973) the average voice pitch is a far more important factor than the three lowest formant frequencies. We may all agree that the timbre differences between female and male voices can be described in terms of "smoothness", female voices sounding "smoother" than male voices. Is the voice fundamental frequency related to "smoothness" then? The timbre dimension "rough/smooth" has recently been successfully examined by hearing scientists, and it seems that the results may contribute to the acoustic description of femaleness and maleness in the voice timbre, as well as to other timbre questions.

As is well known, we perceive sound when the basilar membrane in the cochlea is set into vibration. Each region along the basilar membrane corresponds to a specific frequency region. For instance, the lowest frequencies set the uppermost part of the membrane into a maximum amplitude vibration, and the highest frequencies excite the lowermost end of the membrane (close to the oval window), maximally. If the ear is fed with a complex tone containing several partials, each partial will excite its proper place on the membrane according to its frequency. The spacing of the partials along the mambrane has been shown to be of decisive importance to the rough-smooth timbre dimension. (TERHARDT, 1974) Partials falling closer than about 1.3 mm on the basilar membrane contribute to the roughness. This critical distance of 1.3 mm corresponds to a small-frequency range, called the critical band. The width of these critical bands is given as a function of the center frequency in Fig. 9. For frequencies lower than 450 Hz, the critical band is around 100 Hz, and for higher frequencies its width is close to a minor third. Most musical instruments generate harmonic spectra, i.e. the frequency of the n:th partial equals n times the frequency of the lowest partial. Thus, the frequency distance between each pair of adjacent partials is constant in terms of Hz but decreases in terms of critical bands. Between the fifth and sixth partials, there is an interval of a minor third. Consequently, these partials fall into the same critical band, provided that the partials are higher than 450 Hz. This condition is fulfilled when the fundamental frequency is higher than $\frac{450}{5}$ = 90 Hz. In these cases, then, the partials higher than the fifth may contribute to the roughness of the timbre. The stronger such pairs of adjacent partials are, the rougher is the timbre.

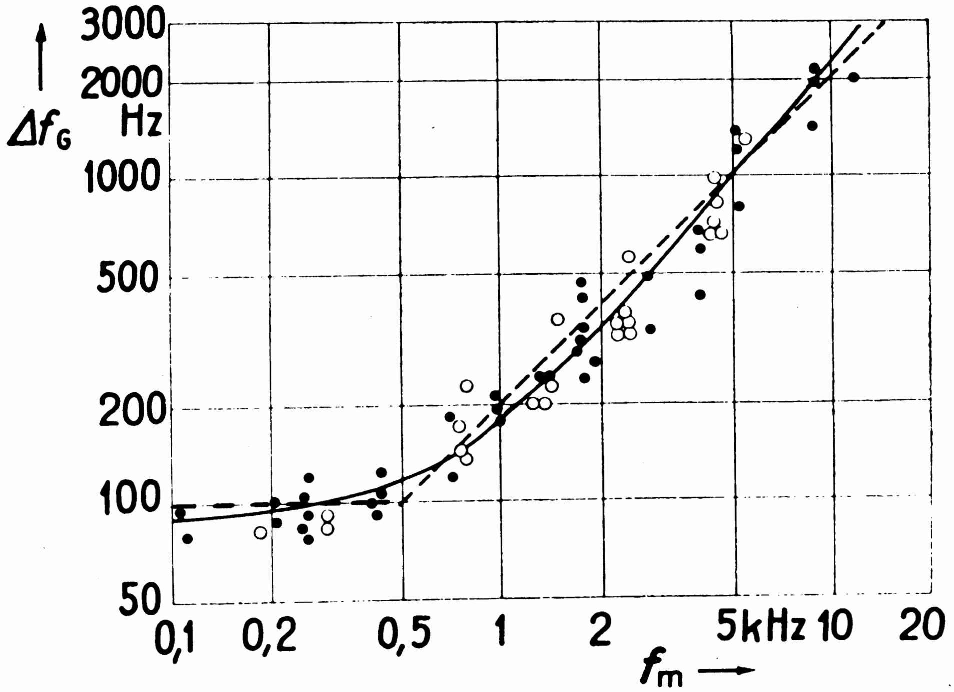



Fig. 9. The width of the critical bands as a function of their center frequency. The dashed line shows an approximation: the critical band is around 100 Hz for frequencies lower than 450 Hz and a minor third for higher frequencies. The different symbols refer to different experimental conditions, see ZWICKER & FELDTKELLER: Das Ohr als Nachrichtenempfänger, (2) S. Hirzel Verlag, 1967, p 72 f, from which the figure is taken.

The effect is demonstrated in sound example five. The first sound contains partials no 1 and 2 in a harmonic spectrum. These partials excite separate critical bands. The same applies to the second sound comprising partials no 1, 2, 3 and 4. In the third sound partials no 5 and 6 are added, and these partials excite the same critical band. The timbre is becomming slightly rough. This effect is considerably enhanced in the fourth sound, which is constituted by partials no 1-8. In each of these sounds all partials are of equal acoustic amplitude and each sound is presented three times in succession.

Rough/smooth and male/female voice

How is this related to the roughness/smoothness in the male and female voice, then? This is not thoroughly examined yet, but the solution seems to be within reach. Let us first consider the question why fundamental frequency seems to be important to sex identification by voice. Disregarding substantial individual variations, we may say that female voices speak with a fundemantal frequency lying about one octave higher than that of male voices, on the average. As mentioned above, the formant-frequency differences between the sexes are much smaller. Therefore, the average number of partials per formant is smaller in the case of female voices. The average number of partials per formant must determine the likelihood of two adjacent partials being of equal and strong amplitude, or in other words, the roughness. In this way it is possible that one can explain a dependence of roughness on the fundamental frequency in speech. But even in a case where we compare an alto and a tenor singing the same vowel on the same pitch we observe a timbre difference, which many of us would certainly describe as a difference in roughness. Such a roughness difference could probably be explained in a similar manner. A tenor would have a longer vocal tract and hence lower formant frequencies than an alto. If the formant frequencies are lower, they will fall closer to each other on the frequency scale, particularly as regards the higher formants, which are hard to move in frequency by articulatory gestures. In the case of the tenor, then, the higher formants would tend to emphasize adjacent partials, i.e. partials falling into the same critical band, whereas in the case of the alto non-adjacent partials would be emphasized, i.e. partials falling into separate critical bands.

The effect is illustrated in sound example six. A vowel is repeated six times. The first, third, and fifth times the third and fourth formants are close enough in frequency so as to emphasize adjacent partials in the spectrum (the 9th and the 10th). In the remaining vowels these two formants are more widely separated in frequency so that non-adjacent partials are enhanced, namely the 9th and the 11th. These partials fall into separate critical bands. In the first mentioned case there are 9 partials distributed under 4 formants, and consequently, the mean number of partials per formant is 9/4=2.25. The corresponding value for the other case is 11/4=2.75. Once again we find reasons for suspecting the average number of partials per formant to play an important role in voice timbre. Even though the explanatory power of that number remains to be explored, Terhardt's theory of roughness seems to open up new ways to a deeper understanding of timbre in singing.

## Rough/smooth and organ timbre

Needless to say, timbre is an important factor in most musical instruments. Let us digress and see how Terhardt's theory fits into the timbre of the organ, an instrument where the timbre is of extremely great musical importance. An organ contains several stops, each of which includes a chromatically tuned series of pipes. All pipes within a stop possess similar timbre characteristics. Some stops are constituted by "closed" pipes, which are closed at the upper end. Such pipes generate spectra containing odd-numbered partials only. Hence, the frequency distance between adjacent partials are twice as wide as in other, "open" stops with pipes generating spectra containing both even-numbered and odd-numbered partials. Terhardt's theory predicts a higher degree of roughness in open stops than in closed stops because of the difference in the density of the partials. The effect can be experienced in sound example seven. An excerpt of a Bach-choral is played first on a closed stop (Gedakt 8'), then on an open stop (Principal 8').

According to Terhardt roughness will occur not only because of the density of the partials but also when the amplitude of partials exciting the same critical band are equal and strong. In the sound from organ flue pipes, the amplitudes of the partials generally fall with the number of the partial. Most stops are tuned to octaves, so that the sounding pitch corresponds to either the key played or to one of its (mostly higher) octaves. Some stops, however, are tuned to an octave+a fifth. The effect on roughness of adding octave stops is small because partials exciting different critical bands are enhanced, as can be seen from Fig. 10. Adding a fifth stop would increase roughness, on the other hand. It enhances the 10:th partial of the lowest stop, while the octave stops enhance its neighbour partials, as can be seen in the same figure.

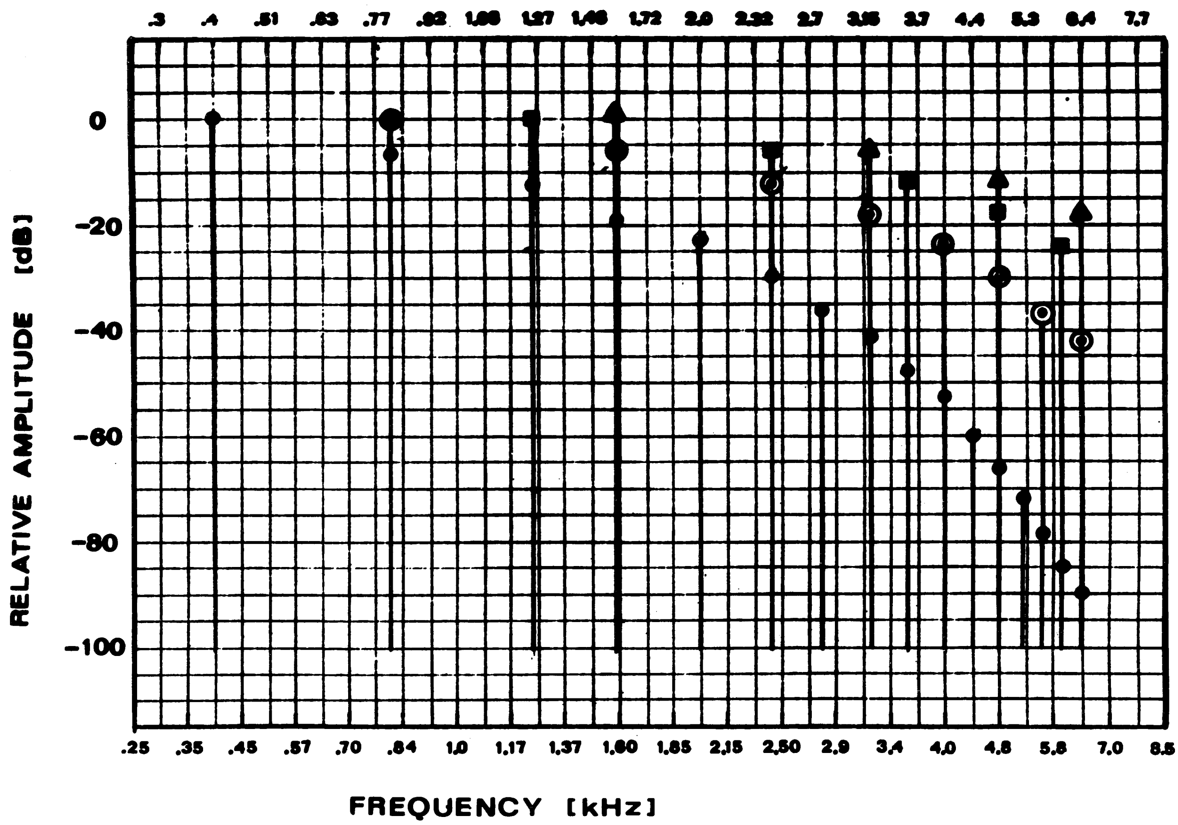



Fig. 10. Schematical representation of the spectra produced by a couple of organ stops. The key played is approximately corresponding to $G_4$. The frequency scale is made so that the distance between one division equals one-half of a critical bandwidth. It is seen that the density of the partials increases the higher up in the spectrum we go. The amplitudes of the partials in each spectrum decrease by 6 dB per partial. The stops are

| | | |
|---|---|---|
| ● | Principal | 8' |
| o | Octava | 4' |
| ▲ | Octava | 2' |
| ■ | Quinta | $2\frac{2}{3}'$ |

Note the dense cluster of strong partials between 3.2 and 4 kHz arising when also the fifth stop is sounding.

The organ builder actually determines the roughness, among other things, when he designs and tunes the pipes, because the number and amplitudes of the partials depend on the physical properties of the pipe. (MEYER, 1960; SUNDBERG, 1966; ISING, 1971) But also, a given stop can give very differing timbres depending on the acoustic properties of the room in which the stop is placed. In such cases a difference in roughness seems to describe the timbre differences rather well. If we consider the room as a huge violin sounding box, we realize that there is a very high number of resonances in any room, and that the density of the resonances along a frequency scale will increase with the size and the sound reflexion in the room. Thus, as in the case of the violin, (cf. MATHEWS, 1976) the likelihood of two partials falling into the same critical band to be of equal and strong amplitude is smaller in a large and reverberant room than in a small and non-reverberant room. Hence, a given stop is likely to sound less rough in a large and reverberant room. Here again we lack experimental support for the conclusion, but there is no doubt that organ timbre seems to offer fruitful soil to the acoustics of music. Presumably the concept of roughness and its acoustical correlate will prove to be extremely useful for organ builders, players, and for composers of organ music.

Rough/smooth and consonance/dissonance

Roughness has also been shown to be related to the problem mentioned initially, consonance and dissonance. PLOMP & LEVELT (1965) showed that dyads of sinusoids sound maximally dissonant if they are separated by a quarter of a critical band, as shown in Fig. 11. These findings are in rather close agreement with a theory which von Helmholtz put forward and which was rejected by most music theorists. Von Helmholtz claimed that maximum dissonance is obtained when the frequencies of the sounding tones differ by 30 to 40 Hz, and 30 to 40 Hz is a rather fair approximation of a quarter of a critical bandwidth as long as we disregard frequencies higher than 1000 Hz. The findings of Plomp & Levelt have rather unexpected consequences. In the case of pure tones, consisting of one partial only the frequency difference in critical bands, (and thus not the frequency ratio) is the factor determining the degree of dissonance. Thus (1) the major third between sinusoids of 100 and 125 Hz sounds quite dissonant, and (2) the major seventh between sinusoids of 440 and 831 Hz sounds almost as consonant as (3) the octave between sinusoids of 440 and 880 Hz, because the frequency distance is then wide enough in terms of critical bands. These three cases are presented in sound example eight.

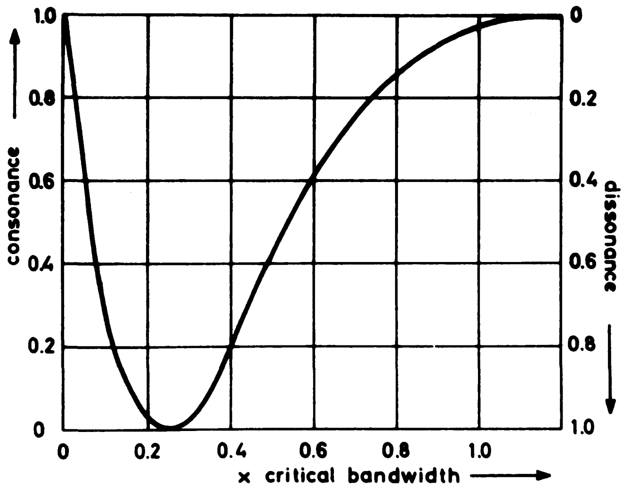



Fig. 11. Idealized results of experiments where musically untrained subjects were asked to estimate the degree of consonance between two simultaneously sounding sine waves. When the frequency distance between the sine waves corresponds to 0,25 of a critical bandwidth, the degree of dissonance reaches a maximum (equivalent to minimum consonance). From PLOMP & LEVELT (1965), Journal of the Acoustical Society of America.

However, these somewhat odd effects disappear as soon as we turn to complex tones. This can be heard in the same example, where the same three dyads are repeated with complex tones instead of sinusoids. In the case of complex tones, the frequency distance between all harmonics contribute, and pairs of partials separated by a quarter of a critical band contribute maximally to the dissonance effect. If we consider complex tones with harmonic spectra only, we will find that, as a rule, tones with simple fundamental frequency ratios will sound more consonant than tones with more complicated ratios. Thus, old Pythagoras' observations have eventually found an explanation, after more than 2000 years!

The theory of consonance of Plomp & Levelt and Terhardt's theory of roughness certainly resemble each other in important respects, as they both relate a timbre quality to a property of the sense of hearing, i.e. the critical bandwidth. And, actually, Terhardt brings about some evidence for the conclusion that there is no sensory difference between roughness and dissonance: under certain experimental conditions they cannot be separated from each other.

It is likely that these findings on timbre will prove to be extremely important to our music culture, not least to the consciously working sound designer or electronic music composer.

## Outlook

Some of the more recent advances in musical acoustics as well as in auditory perception have been mentioned aboved. This research has arisen at the point where physics and auditory perception meet with music and it is there musical acoustics has its domain. The new tools and knowledge which have recently been developed in this field indicate that music culture can profit rather extensively from musical acoustics. Also, the reasons for musicologists to avoid acoustical aspects of music have been eliminated in important respects. If music scientists realize the new and pregnant possibilities inherent in a combination of physics and auditory perception, it seems that significant evolution and progress can be made within musicology to the benefit of our music culture in general.

References:

APPELMAN, D.R. (1967): The Science of Vocal Pedagogy, Indiana University Press, Bloomington, London.

CLEVELAND, T. (1975): The Acoustic Properties of Voice Timbre Types and their Importance in the Determination of Voice Classification in Male Singers. Diss. Univ. of Southern California, USA.

COLEMAN, R.O. (1973): "A Comparison of the Contributions of two Vocal Characteristics to the Perception of Maleness and Femaleness in the Voice", STL-QPSR * No. 2-3, pp. 13-22.

FANT, G. (1968): "Analysis and Synthesis of Speech Processes", pp. 173-277 in Manual of Phonetics (ed. B. Malmberg), North-Holland Publ. Co., Amsterdam.

FANT, G. (1973): Speech Sounds and Features, The MIT Press, Cambridge, Mass.

GIBIAN, G.L. (1972): "Synthesis of Sung Vowels", Mass.Inst.Techn., QPR No. 104, Jan. 15, pp. 243-247.

von HELMHOLTZ, H. (1863): Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik, Verlag F. Vieweg & Sohn, Braunschweig.

HOLLIEN, H. (1960): "Some Laryngeal Correlates of Vocal Pitch", J. Speech and Hearing Res. 30:1, pp. 52-66.

ISING, H. (1971): "Erforschung und Planung des Orgelklanges", Walcher Hausmitteilung Nr 42, pp. 38-57.

LARGE, J.W. (1968): "An Acoustical Study of Isoparametric Tones in the Female Chest and Middle Registers in Singing", Nat.Ass. Teachers of Singing, Dec. pp. 12-15.

LARGE, J.W. (1972): "Towards an Integrated Physiologic-Acoustic Theory of Vocal Registers", Nat.Ass. Teachers of Singing, Feb/Mar pp. 18-36.

LINDBLOM, B. and SUNDBERG, J. (1971): "Acoustical Consequences of Lip, Tongue, Jaw, and Larynx Movement", J. Acoust.Soc.Am. 50:2, pp. 1166-1179.

MATHEWS, M.: "Analysis and Synthesis of Timbres", in this book, pp. 4-19.

MEYER, J. (1960): Über Resonanzeigenschaften und Einschwungvorgänge von labialen Orgelpfeifen, Diss. Braunschweig.

ONDRÁCKOVÁ, J. (1973): The Physiological Activity of the Speech Organs, Mouton,The Hague.

PLOMP, R. and LEVELT, W.J.M. (1965): "Tonal Consonance and Critical Bandwidth", J. Acoust.Soc.Am. 38, pp. 548-560.

RAYLEIGH, Lord (1878): Theory of Sound, MacMillan and Co, London.

---

* Speech Transmission Laboratory Quarterly Progress and Status Report

SIMON, P., LIPS, H. and BROCH, G. (1972): "Étude sonagraphique et analyse acoustique du temps réel de la voix chantée à partir de différentes techniques vocales", Travaux de l'Institut de Phonétique de Strasbourg 4, pp. 219-276.

STUMPF, C. (1926): *Die Sprachlaute, experimentell-phonetische Untersuchungen*, Verlag von Julius Springer, Berlin.

SUNDBERG, J. (1966): *Mensurens betydelse i öppna labialpipor*, Almqvist & Wiksell, Uppsala.

SUNDBERG, J. (1972): "Production and Function of the 'Singing Formant'", pp. 679-688 in *Report of the 11th Congr. of the Internat. Musicological Soc.*, (ed. H. Glahn, S. Sørensen, P. Ryom) *Vol. II*, Edition Wilhelm Hansen, Copenhagen (1974).

SUNDBERG, J. (1973): "The Source Spectrum in Professional Singing", Fol.Phon. *25*, pp. 71-90.

SUNDBERG, J. (1974a): "Articulatory Interpretation of the 'Singing Formant'", J. Acoust.Soc.Am. *55*, pp. 838-844.

SUNDBERG, J. (1974b): "Körsångare - vad vet du om ditt instrument, människorösten", Musiklivet Vår Sång Nr 4, pp. 6-11.

SUNDBERG, J. (1975): "Formant Technique in a Professional Female Singer", Acustica *32*, pp. 89-96.

SUNDBERG, J. (forthcoming): "Maximum Speed of Pitch Changes in Trained and Untrained Voices".

TERHARDT, E. (1974): "On the Perception of Periodic Sound Fluctuations (Roughness)", Acustica *30*, pp. 201-213.

VENNARD, W. (1971): "The Relation between Vibrato and Vocal Ornamentation", J. Acoust.Soc.Am. *49*, p. 137(A).

WINCKEL, F. (1952): "Elektroakustische Untersuchungen an der menschlichen Stimme", Fol.Phon. *4* (Sep.), pp. 93-113.

WINCKEL, F. (1953): "Physikalische Kriterien für objektive Stimmbeurteilung", Fol.Phon. *5* (Sep.), pp. 232-252.

ZWICKER, E. and FELDTKELLER, R. (1967): *Das Ohr als Nachrichtenempfänger*, (2), S. Hirzel Verlag, Stuttgart.

# ON THE ACOUSTICS OF MUSICAL INSTRUMENTS

by

Erik V. Jansson, Center for Speech Communication Research and Musical Acoustics, Royal Institute of Technology (KTH), Stockholm, Sweden.

## Introduction

The musical instruments constitute a link of a human communication chain: composer - player - instrument - listener. The composer groups together demands and directions for the player in the written music. Depending on these demands and directions the player must set specific demands on his instrument to be able to realize his interpretation of the music in an enjoyable way for the listener. In general the player cannot modify his instrument. Therefore the instrument must be made in such a way that it can satisfy the composer and the listener. Furthermore the player wants to be able to express himself with his instrument by playing it differently. This means that the limits set by the instrument must not be too narrow for the player. The instrument maker should meet and preferably foresee these demands of the music - those of the composer, the player and the listener.

The problem of making good instruments divides in a natural way into finding answers to the three following questions: What qualities are wanted? How do different instrument parts influence the qualities? How are instruments with the desired qualities designed? The answer to the first question lies in the psychoacoustic domain, i.e. the understanding of the relation between produced musical tones in physical terms and the perceived tonal qualities. The answer to the second question is sought by the tools of physics: i.e. the understanding of how the instrument works, and how the properties, the shapes and the joining together of different materials influence the produced musical tones. The answer to the third question is of a technical nature, i.e. how to use the understanding of the production of musical tones from physics and the understanding of the tonal quality from the psychoacoustics. This answer should be transformed into rules concerning how instruments of good quality should be made, and how such qualities should be measured.

Traditionally the instrument makers have collected much experience about how good instruments should be made, some of which can not be explained by science today. But their experience is partly hard to transfer to other craftsmen and their design rules often include adjustments in the final stages. Therefore design rules firmly rooted in physics and psychoacoustics as sketched in this paper can be of good help for the single instrument maker. They are necessary in order to guarantee good instruments when rationally made in large numbers. The acoustical investigations of musical instruments aim at finding such rules. This paper will review recent achievements in this area of the acoustics of music.

The intention of the author is to present the information in a way easy to read but still correct, suitable as a nonformal introduction to the subject. Therefore no thorough explanations are offered, but rather implications to give an intuitive feeling for the phenomena involved. The terminology is chosen to be explanatory rather than strict in a scientific sense. It is the hope of the author to provide an understanding of which physical qualities are wanted of good instruments, how some musical instruments work and what can be said of how such instruments should be designed. Two things particularly important for good instruments, intonation and timbre, have been selected as topics for this paper. Intonation will be discussed mainly in connection with wind instruments, and timbre in connection with the string instruments, the violin and the guitar. The information on each topic is presented in three parts, first the psychoacoustic part, secondly the physical part and finally the technical part corresponding to the three questions asked above. The information presented is not sufficient for calculating designs. For such purposes as well as for more detailed information the reader is directed to the references listed which hopefully will give the information wanted.

## Intonation and wind instruments

The musical tones are not single tones in a physical sense, they consist of a chord. This musical-tone-chord is made up of several single tones with specific interval relations. The musical intervals between the lowest and the second lowest, the second lowest and the third lowest, etc. tones are an octave, a fifth, a fourth etc. respectively. The musical pitch is closely related to the physical frequency Hz, of the lowest tone.

The magnitudes of the intervals mentioned are equally measured in the physical frequency Hz, which means that the musical tone-chord forms a so-called harmonic spectrum.

Our culture has adopted a simple and physically well defined scale, the equally tempered scale. In this scale the octave corresponds to a doubling of the physical frequency. Although being a musical compromise the scale is generally accepted. Still it poses some difficulties in connection with the function and the physics of musical instruments. Keyboard instruments as the piano and the organ are tuned according to this scale. Careful measurements show, however, that eventhough the organ is tuned very closely to the equally tempered scale, the piano is tuned with stretched intervals, see Fig. 1 and 2. (SUNDBERG, 1967 and MARTIN & WARD, 1961) This difference in tuning can be explained by means of the function or the physics of the instruments. Every tone of the organ is made up of a strictly harmonic spectrum and therefore the frequencies of the different tones are multiples of the fundamental frequency. On the piano, every note is made up by a slightly inharmonic spectrum and the frequencies of the different tones are slightly higher than multiples of the fundamental frequency. To prevent octave chords from beating, the organ should be tuned with "correct" intervals, but the piano with slightly larger intervals.

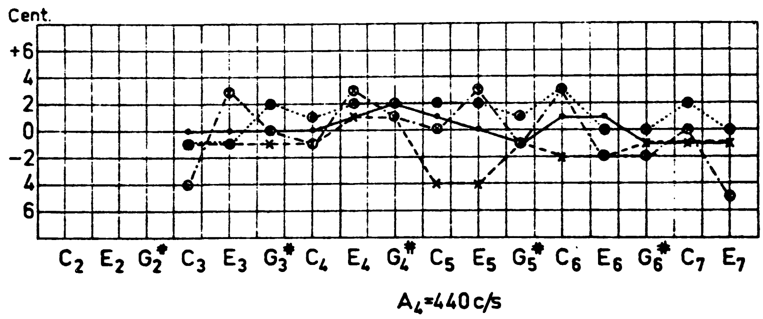



Fig. 1. Deviations in cents from the equal tempered scale in different organ stops. ● and ⊙ refer to two diaposon stops; x and ⊗ refer to two flute stops. From SUNDBERG (1967), Svensk Tidskrift för Musikforskning.

Under certain conditions the ear prefers stretched intervals to "correct" intervals. Thus, even the octave, the basic interval of all scales, is sometimes preferred as sligthly stretched. If two tones, spaced one octave apart, are played simultaneously, the chord sounds the best with the "correct" tuning. If, on the other hand the same tones are played in succession, the interval definitely sounds the best if it is slightly stretched. This can easily be demonstrated with a violin.

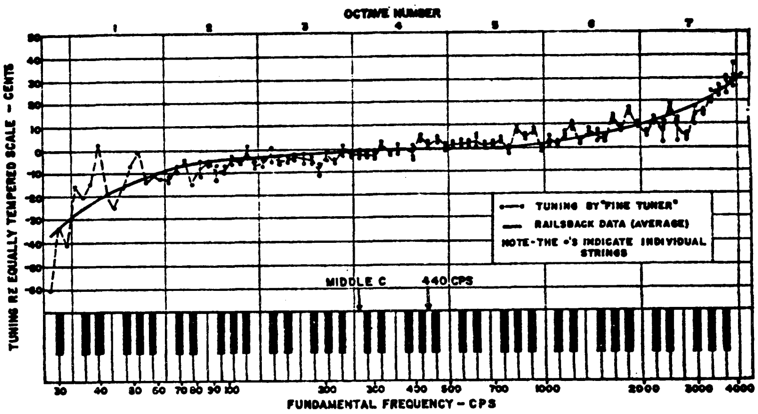



Fig. 2. Deviations in cents from the equal tempered scale in pianos. From MARTIN & WARD (1961), Journal of the Acoustical Society of America.

Thus in playing the player may want to make final adjustments of the intonation because of physical and psychoacoustical reasons. But the physics of his instrument, for instance a flute, may enforce the player to make fairly large adjustments to achieve good intonation. In Fig. 3 the full line represents tones played on flute. (FRANSSON, 1963) Each note was played separately and the player was not instructed what note to play next until he had finished playing the first. Thus he played the frequencies the instrument responded to with no intonation correction. The tones played do not describe an equally tempered scale - the full line in Fig. 3 is neither horizontal nor straight. The physical measurements of the resonance frequencies confirm the non-equal tempered scale properties of the flute. The tones played are furthermore on the average a semitone flat. Thus the player has to make two kinds of adjustments to play in tune with other instruments. First he must adjust the average intonation by tuning, i.e. by changing the physics of his instrument. Secondly he must in addition adjust the intonation while playing as much as a quartertone to keep separate notes of a musical piece in tune.

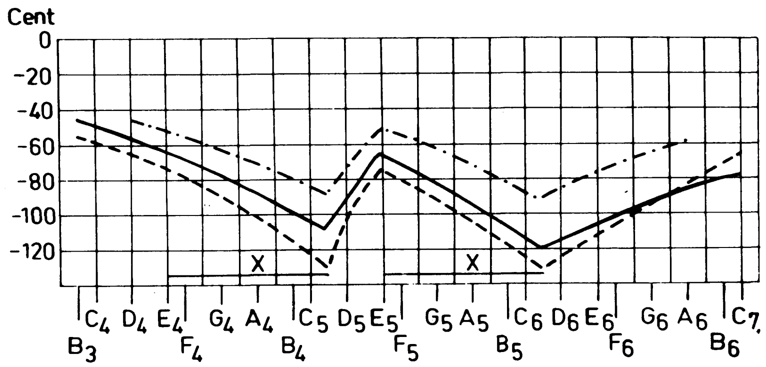



Fig. 3. Relation between resonance frequency and corresponding blown tones. ——— mean values for five flutes, —·—·— STL-Ionophone measurements on partly covered narrow embouchure and – – – STL-Ionophone measurements on partly covered normal embouchure. From FRANSSON (1963).

Let us continue to study musical instruments, instruments in which their constructions, their physics, determines the frequencies of tones played. (BENADE, 1976) Typical such instruments are the wind instruments. They consist of tubing with a more or less complex shape. The frequency of the played tone is determined by the acoustic length of this tubing. This length is approximately equal to the distance from the "blowing" end to the first open side hole, or, if there is no side hole to the bell opening. Thus the total length of the tubing determines the lowest note that can be played. This way of looking at the wind instruments is qualitatively a correct description of the function of the tubing. On a long instrument, as the basoon, low notes can be played, and on a short one, as the piccola flute, only high notes. Shortening the acoustical length of a tubing by opening side holes results in a higher note. A more accurate description is, however, needed to predict the frequency of played notes with an accuracy sufficient for musical demands.

Before giving a rough sketch of such a description, let us very shortly look into the fundamentals of the function of brass wind instruments. The principles also apply to reed instruments. In Fig. 4 a typical setup is described for experimental investigations. A capillary feeds "constant" sound flow (velocity) into the mouthpiece of a trumpet and a small microphone registers the resulting pressure variations, the sound pressure, in the mouth piece. (BENADE, 1973) The frequency of the "constant" flow sound is slowly increased and the resulting sound pressure (air

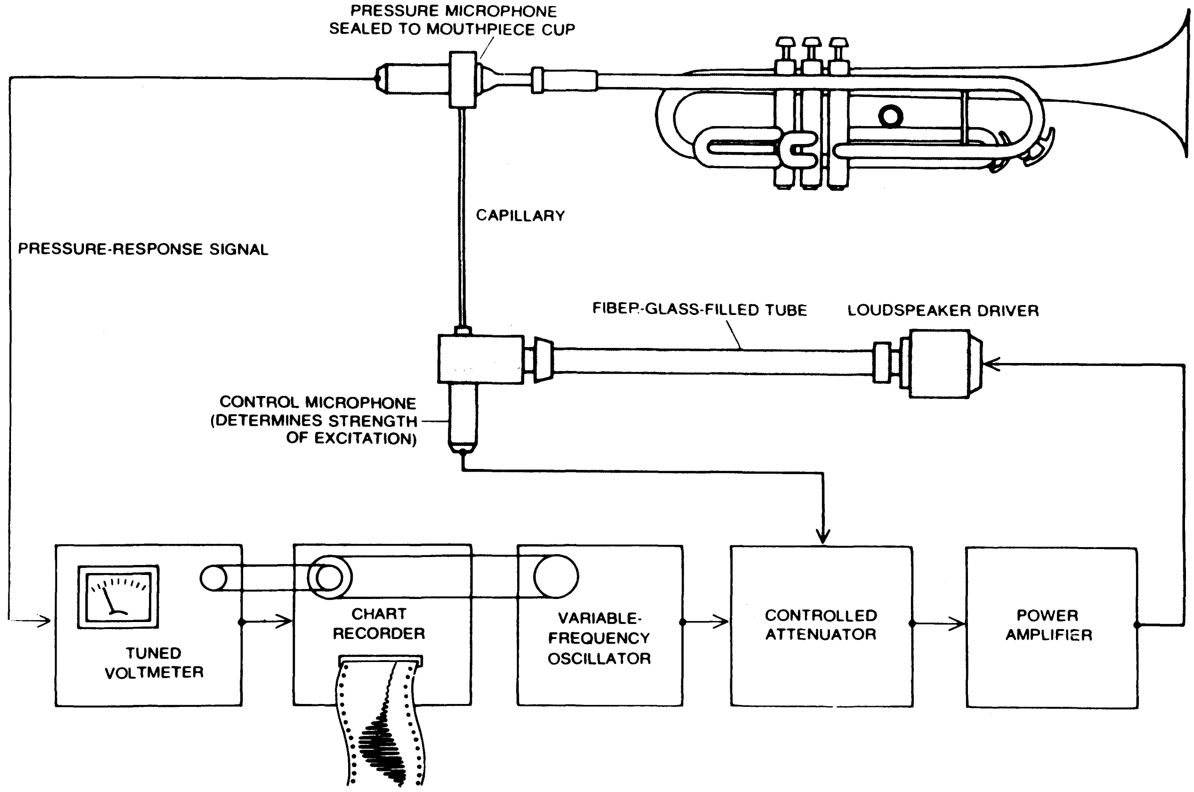



Fig. 4. Impedance-measuring apparatus uses the driver from a horn loudspeaker as a pump to feed a flow stimulus through a capillary into the mouthpiece cup of the instrument under study. A control microphone sends signals to an attenuator to ensure that the acoustic stimulus entering the capillary remains constant. The pressure response of the instrument, and thus its input impedance, is detected by a second microphone that forms the closure of the mouthpiece cup. The signal from the microphone goes to a frequency-selective voltmeter coupled by a chain drive to oscillator. A chart recorder coupled to the voltmeter plots the resonance curves. From "The Physics of Brasses" by A.H.BENADE. 

pressure variations) is plotted by a chart recorder. In this way a record is obtained of how much sound pressure is built up by a given sound flow supplied by the lip vibrations of the player. The built up sound pressure will in its turn influence the lip vibrations, especially if it is high, i.e. the instrument will give a reaction to the action of the player's lips. Similar and highly accurate measurements can also be made with less electronics by means of the STL-Ionophone, which we have been using for the last ten years. (FRANSSON & JANSSON, 1975)

A chart obtained in the way described shows very marked peaks, so called resonance peaks, cf. Fig. 5. At the frequencies of these peaks the instrument will give large reaction on the lip vibrations. The instrument can monitor the lip vibrations in such a way that oscillations are maintained and the musical tone is given. The resonance peaks are simply related to the acoustical lengths. It is well known that the frequencies of played tones are at least close to the peak frequencies. The simple classical theory (the 'linear' excitation theory) says that the instrument plays at the frequencies of the peaks. To test the validity of this theory a horn was designed with a rather special spacing of the resonance peaks, see Fig. 5. (BENADE & GANS, 1968) The resonance peaks were spaced so as to avoid all interger ratios between peak frequencies. When a player tried to play this very peculiar horn, it turned out that the lip sound was amplified at the peak frequencies, but it was almost impossible to produce a sustained note on the instrument. This result contradicts the classi-

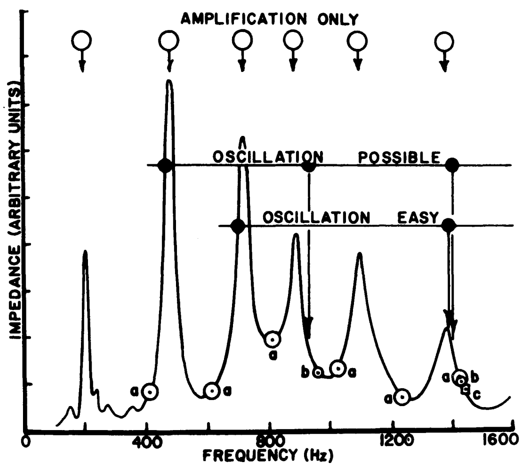



Fig. 5. Measured impedance curve of the tacet-horn. The circles marked "a" indicate the low impedance values observed at frequencies that are harmonics of the lowest resonance frequency. The circles "b" and "c" indicate similarly the impedances at the harmonics of the second and third resonances. From BENADE & GANS (1968), Annals of the New York Academy of Sciences.

cal theory that the fundamental frequency of a played note always equals a peak frequency. However, at the frequency marked by the filled dot just below the second resonance peak in Fig. 5, it was possible to play, and at the frequency marked by the filled dot below the third peak, it was easy to play. In these later cases the higher partials of the played tones fall close to higher resonance peaks. These results suggest that several partials of the "tone" played interact with several resonance peaks to sustain the "tone". Furthermore a large summed interaction gives easily played tones. In physical terms, the function is in agreement with self-sustained nonlinear oscillations. For further studies the reader is referred to BENADE & GANS (1968) and to WORMAN (1971). Thus we would expect that the frequency of a played note is not entirely determined by the frequency of the closest resonance peak.

Let us look at gross features of a real instrument, the clarinet, which has resonance peaks typically positioned as in Fig. 6. The clarinet has a register key to facilitate playing in its upper register. When this register key is closed the resonances are positioned so that odd-numbered partials of a tone played fall at resonance peaks. This is demonstrated in the upper part of Fig. 6, which shows five resonances interacting with five partials giving a large summed interaction. For a tone played in the upper register, only two resonances and two partials interact giving a considerably smaller interaction. Thus the lower register is favoured with the register key closed. When the register key is opened the lowest resonance is most effected and its frequency is shifted as shown in the lower part of Fig. 6. This means that a musical tone with its fundamental frequency equal to the

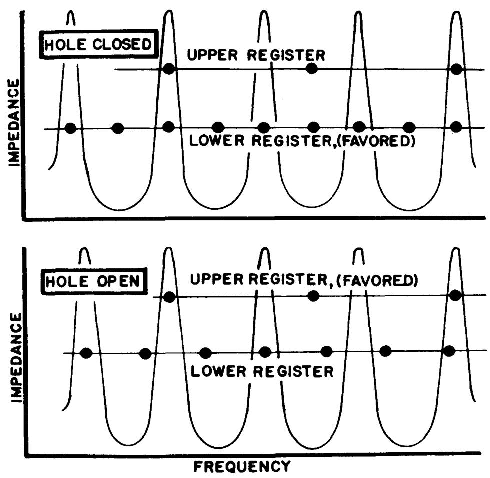



Fig. 6. Schematic diagram of the relation of the impedance curve to the played-note harmonics for a clarinet with its register key open and closed. From BENADE & GANS (1968), Annals of the New York Academy of Sciences.

first peak frequency gives interaction for only one partial and one resonance and a small summed interaction. The response to the tone of the upper register is essentially unchanged giving a larger summed interaction. Thus the upper register is favoured with the register key open.

Let us now look at another example. Fig. 7 shows a resonance curve for a trombone. From the markings on the frequency axis it is easily seen that the higher resonances are not positioned at multiple frequencies of the first resonance frequency. This means that a tone played with its fundamental frequency equal to the first resonance frequency, gives one partial interacting with one resonance peak and small summed interaction is obtained. Thus it is hard to play this tone. On the other hand, if the second peak (labelled 2) is chosen to interact with the second partial, then several higher partials and higher resonances will interact and a large summed interaction, is obtained. The lowest tone used in playing has a fundamental frequency equal to half that of the second resonance, thus using the large summed interaction. This tone is called the pedal tone.

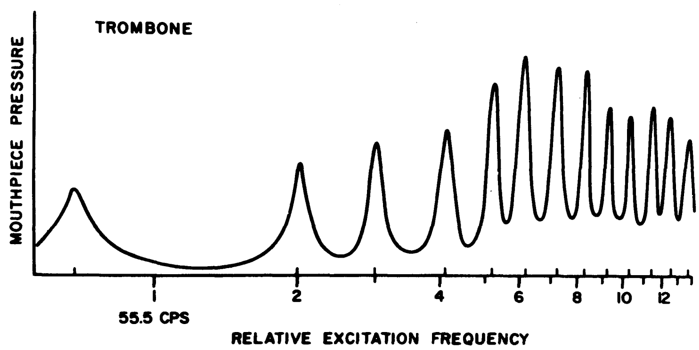



Fig. 7. Resonance curve for a trombone. Reprinted from "THE ACOUSTICAL FOUNDATIONS OF MUSIC" by John Backus. By permission of W.W. Norton & Company, Inc. Copyright © 1969 by W.W. Norton & Company, Inc.

The two examples chosen indicate that we for a complete description must take into account the heights of the resonance peaks and the strengths of the partials. For instance, the phenomenon with the register key of the clarinet applies to loud and medium loud playing but not to soft playing. (BENADE, 1976)

A closely related approach has been taken by WOGRAM (1972). He shows that the frequencies of the tones which players prefer to play on a trombone agree well with the frequencies at which the total sound energy (of all partials) stored in the instrument is maximum. He proves experimentally that these frequencies can be measured in a straightforward way by "blowing" the trombone with a specially designed air siren. The numerical results show that the discrepancies between played frequencies and resonance frequencies are considerable: half a semitone step, on the average. The discrepancies between played frequencies and frequencies measured by means of his siren are small: less than 1/10 of a semitone step. The discrepancies between played and theoretically calculated frequencies are smaller than 1/5 of a semitone step on the average. An alternative way to calculate the properties of bells combined with an extension of classical horn theory to spherical waves is given in BENADE & JANSSON (1974).

The information presented above can be transformed into general design rules. A good instrument should, with normal playing, give an intonation corresponding to the equally tempered scale. This must be achieved in designing the tubing of the instrument, because the material of the walls is of second order importance as long as it is smooth, tight, and rigid. One possible solution, indicated by Benade, is to position, or "line up" the resonance peaks so that their frequencies form a harmonic series. (BENADE, 1976; BENADE & GANS, 1968) Furthermore, the instrument should allow the player to deviate moderately from the equally tempered scale by different ways of blowing, as for the flute, cf. Fig. 3.

Let us again look at a brass instrument, the trumpet. The trumpet consists of four different parts, the mouthpiece, the conical mouthpipe, the cylindrical tubing and the flaring bell. The nonharmonic resonance frequencies of a cylindrical tube are perturbed by means of the three other parts into series of approximately harmonic positioned resonance frequencies starting with the second. This clearly explains why it is necessary to design all parts of an instrument together. A specific mouthpiece which is excellent in one trumpet will not be equally good in all trumpets. Theoretically, a brass instrument can be designed to give harmonically spaced resonance frequencies.*) For such an instrument cylindrical tubing may be added or subtracted without destroying the properties of the instrument. In Fig. 8 the acoustical lengths of a trumpet bell and a trumpet mouthpiece are sketched. A little calculation shows that it is indeed hard to line up all the resonance frequencies of an ordinary trumpet over its total frequency range (a 50 % addition of acoustical length is needed between the first and second resonance). However, in order to line up the second and higher resonances in the continued harmonic series it is sufficient with a 33, 20, 16 % etc. addition. This can be achieved with reasonable accuracy.

The given rules are straightforward but special design charts are needed for the non-mathematical designer. The simple rule of thumb is that adding or subtracting a specific amount of tubing gives a proportional change in the frequency of the tone played. However, this rule is not very exact, because the end corrections of the mouthpiece and the bell vary with frequency, cf. Fig. 8. Moreover, several resonances affect the played tone frequency. The present lack of practical design rules may become serious, as the pitch standard has a tendency to be raised.

---

*) The acoustical length (equal to a quarter wavelength cylindrical pipe) should obey the relation

$$L_{acoust}\ (F,F_1) = (C/4F_1) + (C/4F_1)x(1-F_1/F)$$

where the first term represents a constant for given first resonance frequency $F_1$ (not equal to the cylindrical tubing though), the second term the "end correction" depending on the frequency F, and the velocity of sound C. The relation can be applied partly, for instance beginning with the second resonance.

This will cause players to look for instruments which the maker cannot design from his empirical experience.

In testing the quality of an instrument it is necessary to measure its intonation. One way of doing this is by means of a special blowing machine, as suggested by WOGRAM (1972). But the same thing can be obtained directly from played music as well. A method for automatic notation of played music has been developed. (SUNDBERG & TJERNLUND, 1971) This method can be used to give test records of instruments showing the note played on a tonal system together with its excursion from the standard value with a discrepancy line (cf. Fig. 9). In such a way individual test records can easily be supplied for each instrument tested (with considerably simpler equipment than we are using presently).

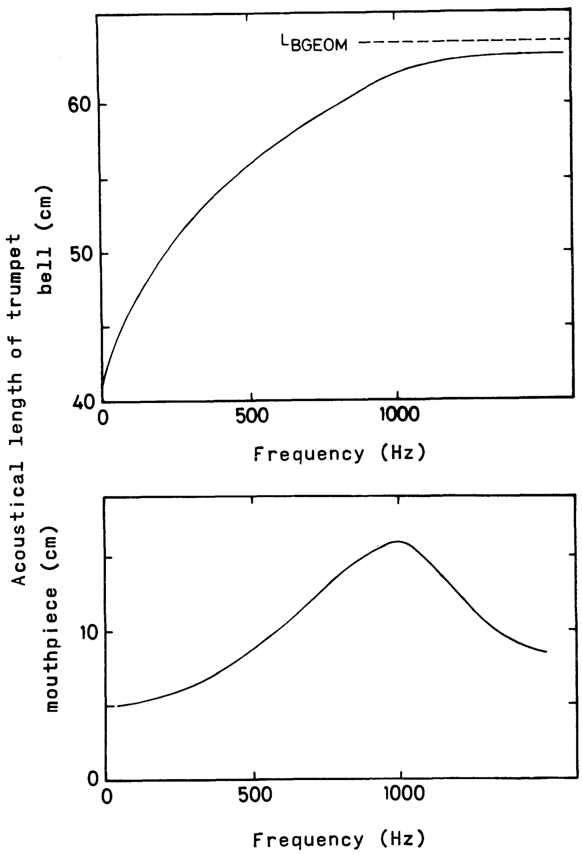



Fig. 8. Acoustical length as function of frequency for upper diagram: a trumpet bell (adapted from BENADE & JANSSON, 1974), lower diagram: a trumpet mouthpiece (adapted from BENADE, 1976).

Timbre and string instruments

The common musical tones are not single tones in the physical sense, they are made up by chords, as stated above. The strength of the different tones of the single musical tone-chord (the amplitude of the partials) gives it a specific quality, which is usually referred to as timbre. In the following we shall restrict timbre to denote this perceptual quality and we shall discuss it in connection with string instruments.

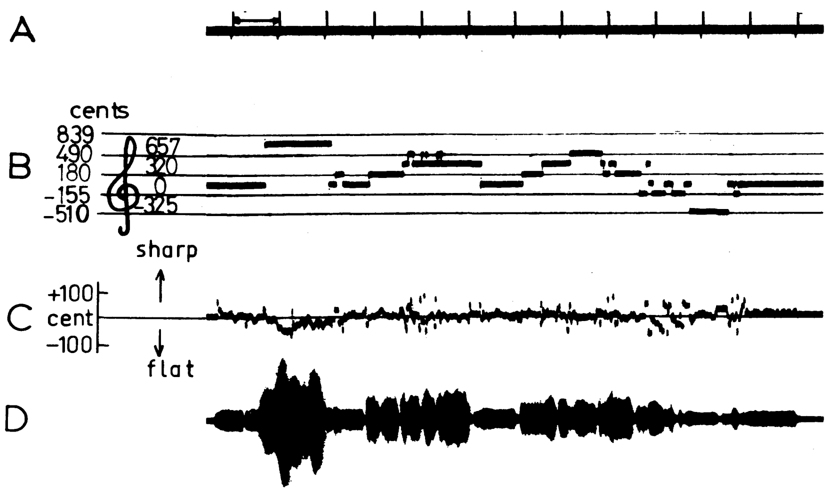



Fig. 9. Automatic notation of played music with deviations in cents given. From SUNDBERG & TJERNLUND (1971), Proceedings of the VIIth International Congress on Acoustics.

The resonance boxes of the string instruments, such as the violin and the guitar, act as amplifiers of the weak sound which is provided by the strings. A small portion of this weak sound leaks to the resonance box, which amplifies it and gives to it the musical quality which we hear. In the violin the amplification may amount to 30 dB. (JANSSON, 1966) This amplification is very frequency dependent and not at all similar to the straight and even response we usually find in good electronic amplifiers. Still, players usually want their instruments to have an "even response". This would suggest that the even response of the electronic amplifier would be ideal even in a violin.

By comparison of physical analysis (the impulse response in 1/3-octave filter bands) with tonal-quality judgements by an expert group YANKOVSKII (1966) found that the best violins have a domeshaped response with a principal maximum at 1250 Hz. Furthermore the best violins had harmonically spaced peaks at 250, 500, 800 and 1250 Hz. Yankovskii also managed to separate the physical responses corresponding to "classical mean", bright, noble, nasal timbre etc. cf. Fig. 10. LOTTERMOSER (1968) used a spectrum analyzer approximating the loudness evaluation of the ear. His findings confirmed that the responses of excellent violins are far from even.

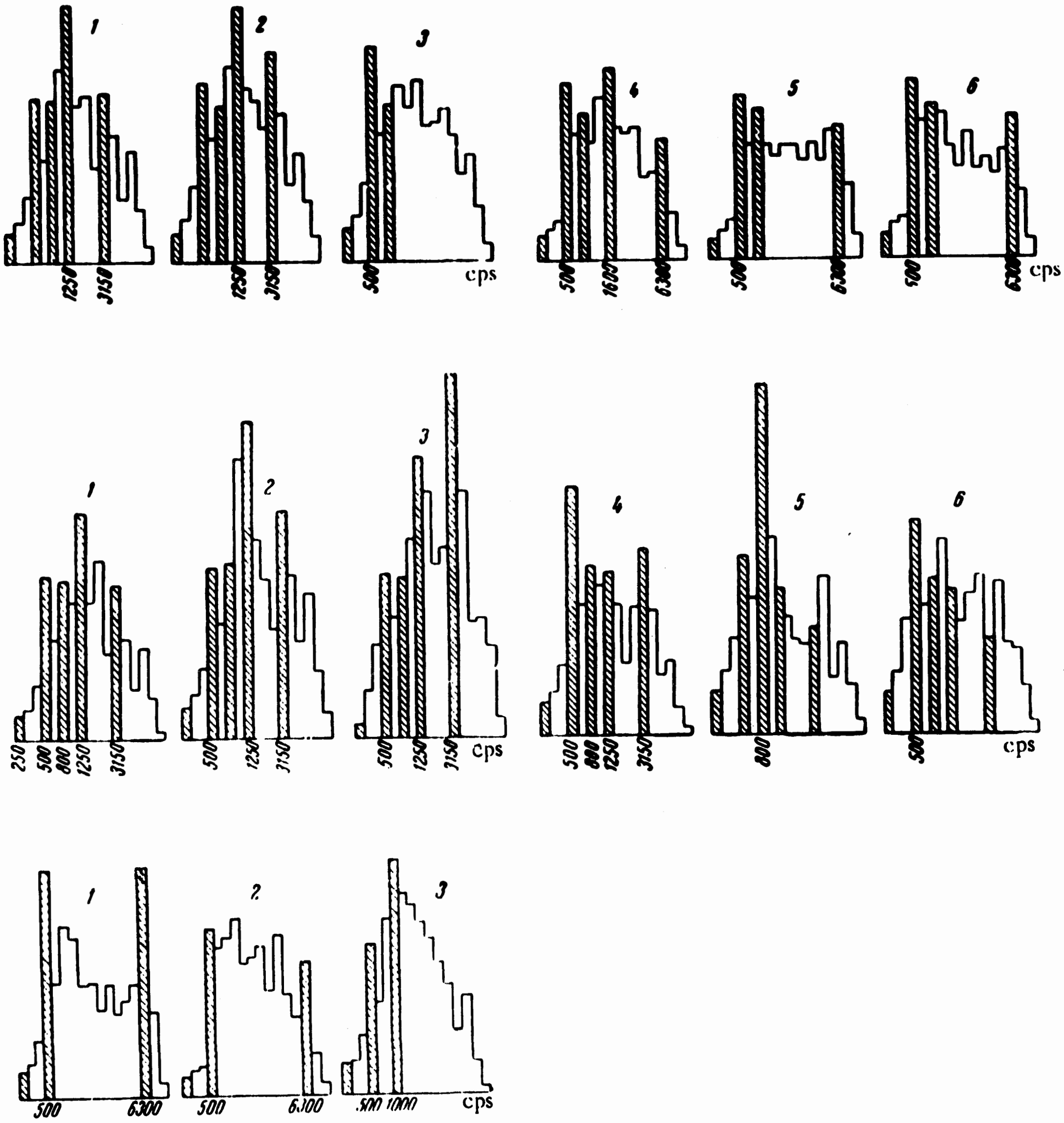



Fig. 10. Impulse response of violins in third octave filter bands. Linear amplitude scale. Upper line: Classical mean (1), Bright (2), Noble (3), Nasal (4), Tight (5) and constricted (6) timbre respectively. Middle line: Stradivarius 1736 (1), Leman 1910 (2), Amati 1629 (3), experimental CE-39 violin 1957 (4), Stradivarius 1723 (5) and Stradivarius 1717 (6). Lower line: Violin of piercing timbre (1), with treble (shrill) quality (2), and with a contra alto tone quality (3). From YANKOVSKII (1966), Soviet Physics-Acoustics.

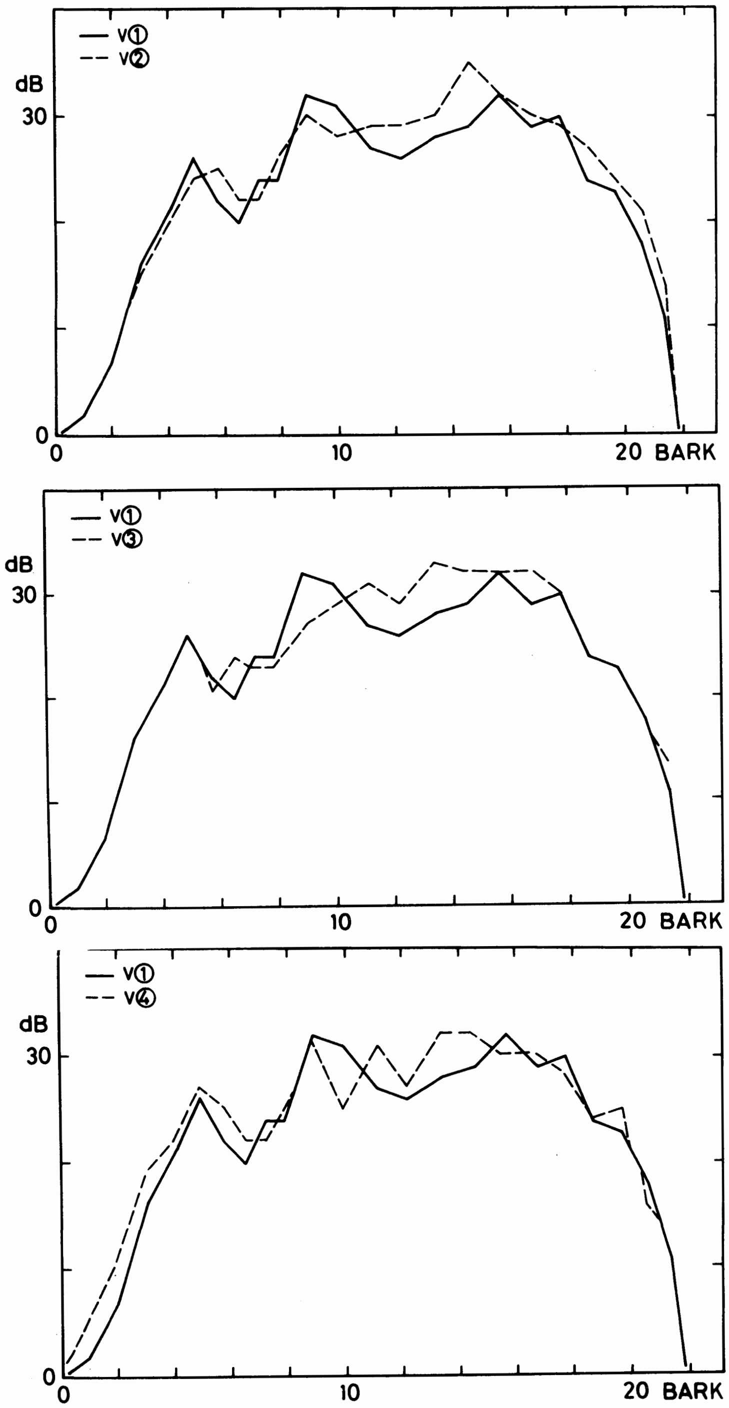



Fig. 11. Long-time-average-spectra of four violins. From JANSSON & SUNDBERG (1975), Acustica.

An improved technique for analyzing the sound of musical instruments has recently been developed in our laboratory. (JANSSON & SUNDBERG, 1975) The technique employs recording in a reverberation chamber, a filterbank, and a computer. The reverberation chamber removes the influence on the results of the microphone position. The filterbank is set to approximate the loudness evaluation of the ear. The analysis is made in three steps. First, three fulltone scales are recorded in succession. Secondly, long-time-average-spectra, LTAS, are made of the three scales and are stored in the computer memory. Finally, LTAS:es of different violins are recalled from the memory for comparisons and manipulations. Such LTAS:es show that different violins can be separated in the analysis, Fig. 11. Does the ear also detect these differences? The answer can be obtained by listening to sound example one: first violin 1 and 2, cf. Fig. 11 a; then violin 1 and 3, cf. Fig. 11 b; and third violin 1 and 4, cf. Fig. 11 c.

Careful analysis has proved that the instruments have the largest influence on the LTAS and the player the least, when the same scales are played in the same way. But even quite different music, as in sound example two, gives LTAS:es of fairly close resemblance, Fig. 12. This means that an LTAS contains information about a specific instrument, which may be sufficient for singling out specific instruments or qualities, even under less rigid test conditions than are used as standard.

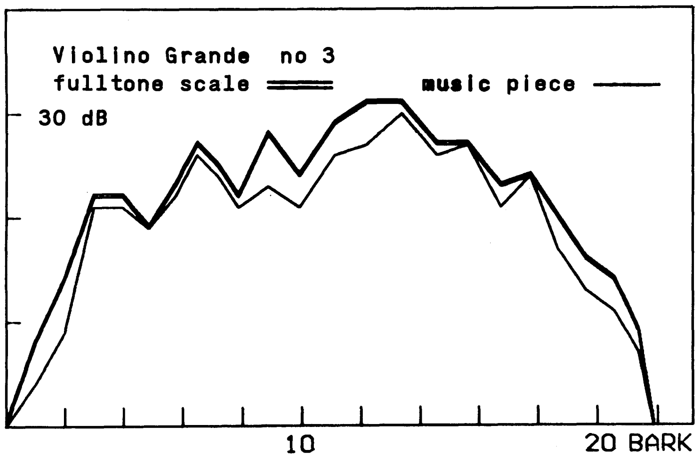



Fig. 12. Long-time-average-spectra of a scale and a music piece played on Violino Grande.

But in addition to the broadband uneveness, the response curves of violins have pronounced zigzag patterns within the broad analysis bands. Investigations of the importance of the zigzag patterns are demonstrated in the first paper in this book (see also MATHEWS & KOHUT, 1973). The "string sound" was picked up from a violin without an acoustical resonance box, was amplified through an electric resonance box and radiated into the room by a loudspeaker. The electronic resonance box could be adjusted to give the amplification curves in Fig. 13. In the experiments it was found that the even amplification curve of Fig. 13 a gave an instrument unresponsive to vibrato and modulations of bow pressure. By introducing a zigzag amplification curve as in Fig. 13 b these shortcomings of the instrument were removed. By making a very pronounced zigzag pattern as in Fig. 13 c and d the musical tones became hollow and uneven.

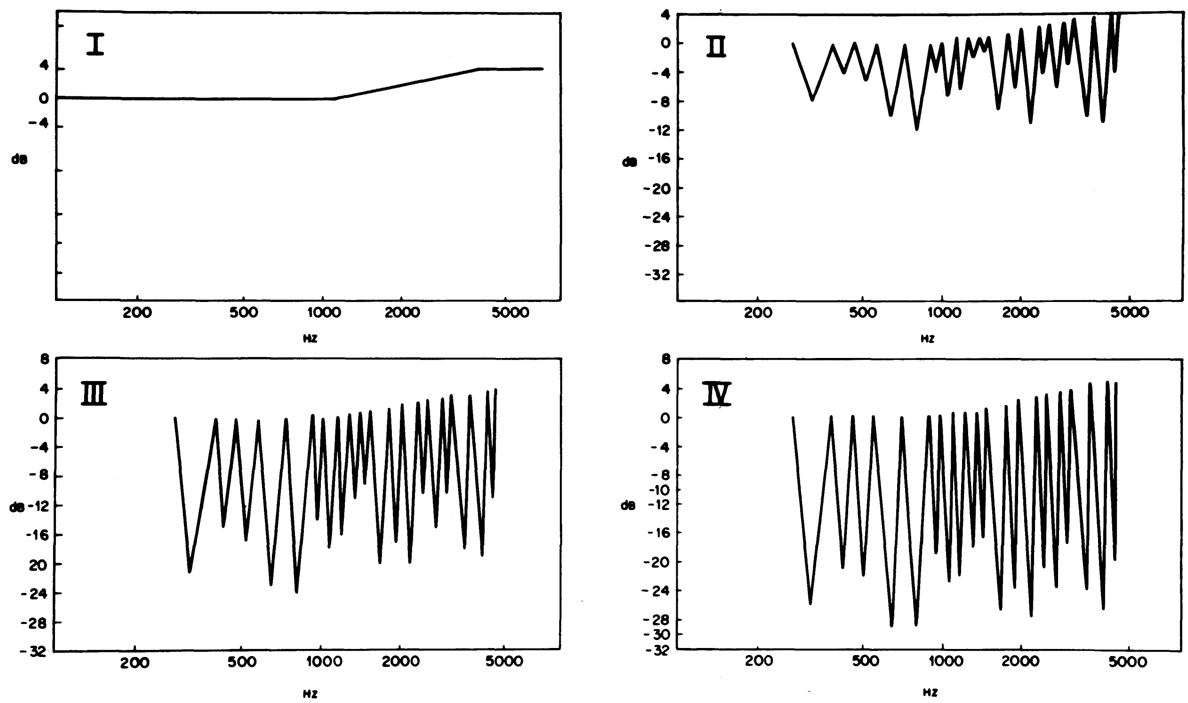



Fig. 13. Peak-to-valley curves for various Q's of electrically simulated resonances at Stradivarius peaks. From MATHEWS & KOHUT (1973), Journal of the Acoustical Society of America.

The conclusion of the above presented information is that the quality criteria for technical products such as amplifiers are not valid for musical instruments. The even response that players find in good instruments is not achieved by physically straight responses. The resonance box should give an amplification with some regions of high and some of low amplification to give desired timbre qualities. Furthermore the amplification curve should include a certain amount of peakiness to make instruments responsive to vibrato.

Let us now turn to the material of the resonance box, which is known to be of great importance. The properties of wood can be summarized in two pairs of measures: these pairs concern the properties along, and the properties in parallel to the wood fibres, respectively. One measure of the pair specifies the elastic properties of the material in terms of the ratio between stiffness and weight $(E/\varrho^3)^{1/2}$, where E is the modulus of elasticity and $\varrho$ the density. It can be shown that the simplest way of compensating for variations in the elastic properties is to tune the first resonance to a specific frequency. If this is done, the differences in the elastic properties affect the sound radiation as sketched in the upper part of Fig. 14. The amplification is simply lowered or increased as much

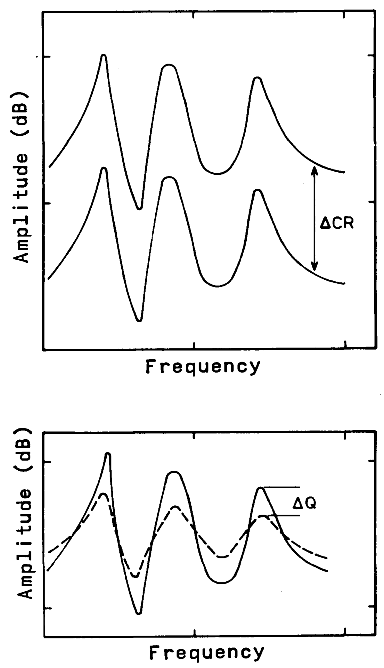



Fig. 14. Effect on the response curve of a change in the elastic properties CR after tuning to maintain the frequency of the first resonance (upper diagram) and a change in the Q-factor (the internal losses of the material).

as the elasticity measure CR is lower or higher than the reference value. (SCHELLING, 1963; JANSSON, 1975) The second measure of the material reflects the internal friction. Variations in the internal friction influence the radiation properties as sketched in the lower part of Fig. 14. They influence mainly the peaks and the dips - low friction (a high Q-factor) results in high peaks and low dips, high friction (a low Q-factor) results in low peaks and shallow dips.

Let me continue by demonstrating the influence on the function of the design, or the shape. The resonance box of the violin consists of a top plate, a back plate and the ribs. Two sound holes are cut in the top plate, the f-holes. The top plate is strengthened by the bass bar and is furthermore supported by the sound post against the back plate. The acoustical function of the sound post has been much disputed.

By means of an optical technique developed at the Institute of Optical Research at KTH, small vibrations can be made visible and can be photographed giving interferograms as in Fig. 15. (JANSSON & al., 1970; JANSSON, 1973a) Let me explain how to interpret the interferogram. The black lines (fringes) on the violin top plate show the pattern of vibration when "photographed". The pattern can be visualized as a typographical map with the black lines connecting points of equal altitude. The black lines of the interferogram show, however, not altitude but magnitude of vibration. The black lines correspond approx. to a vibration of one ten thousandth of a mm, twice as much, three times as much etc. for the first line, the second, the third etc. respectively.

Fig. 15 is a "photograph" of the vibration pattern of the first top plate resonance. In the figure we can first see two vibratory "mountains" (antinodal areas) one wide and high to the left side and a lower and narrower one to the right side. There are no vibrations at the ribs. The large left "mountain" gives mainly the sound we hear from the first resonance. Secondly most black lines start and end at the f-holes, and the vibration amplitude is maximum at the f-holes. This means that the f-holes are effectively cutting the vibrating plate free from the ribs and thus make the plate vibrate more easily. Thirdly at the position of the sound post the plate vibrations are zero (a node). This means that the sound post acts as a fulcrum for a rocking-lever, (the top

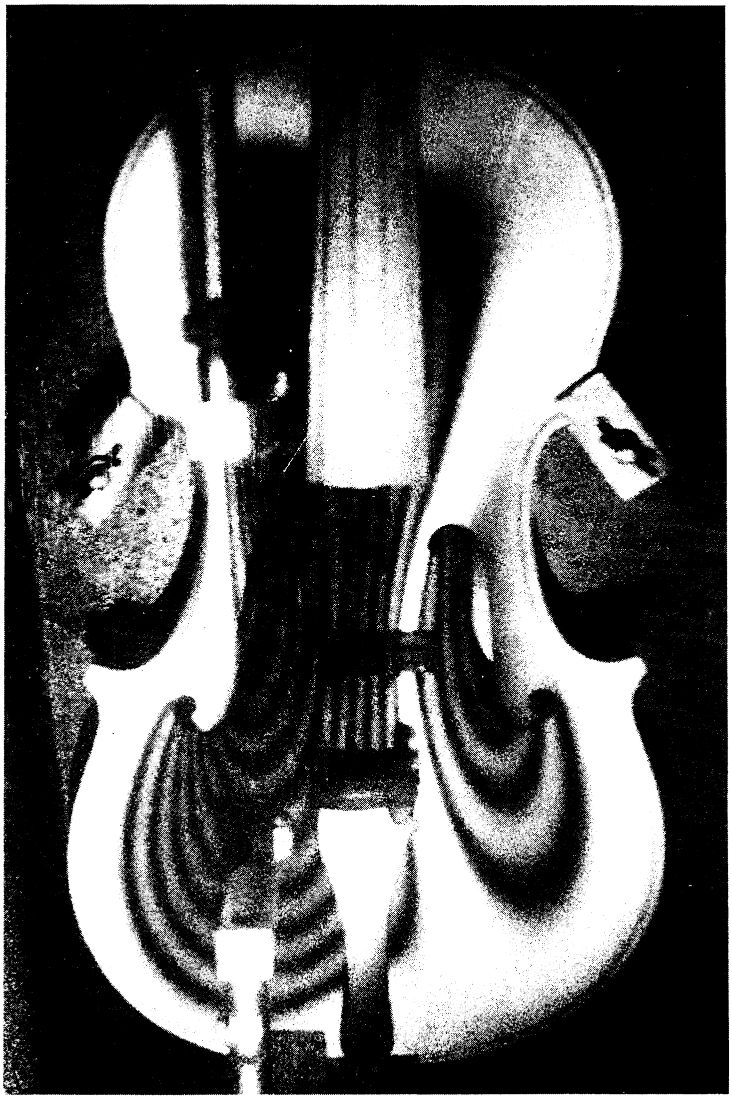


Fig. 15. Interferogram of a violin showing the vibrations of the first top plate resonance. (Jansson, Molin and Sundin unpublished measurements 1970)

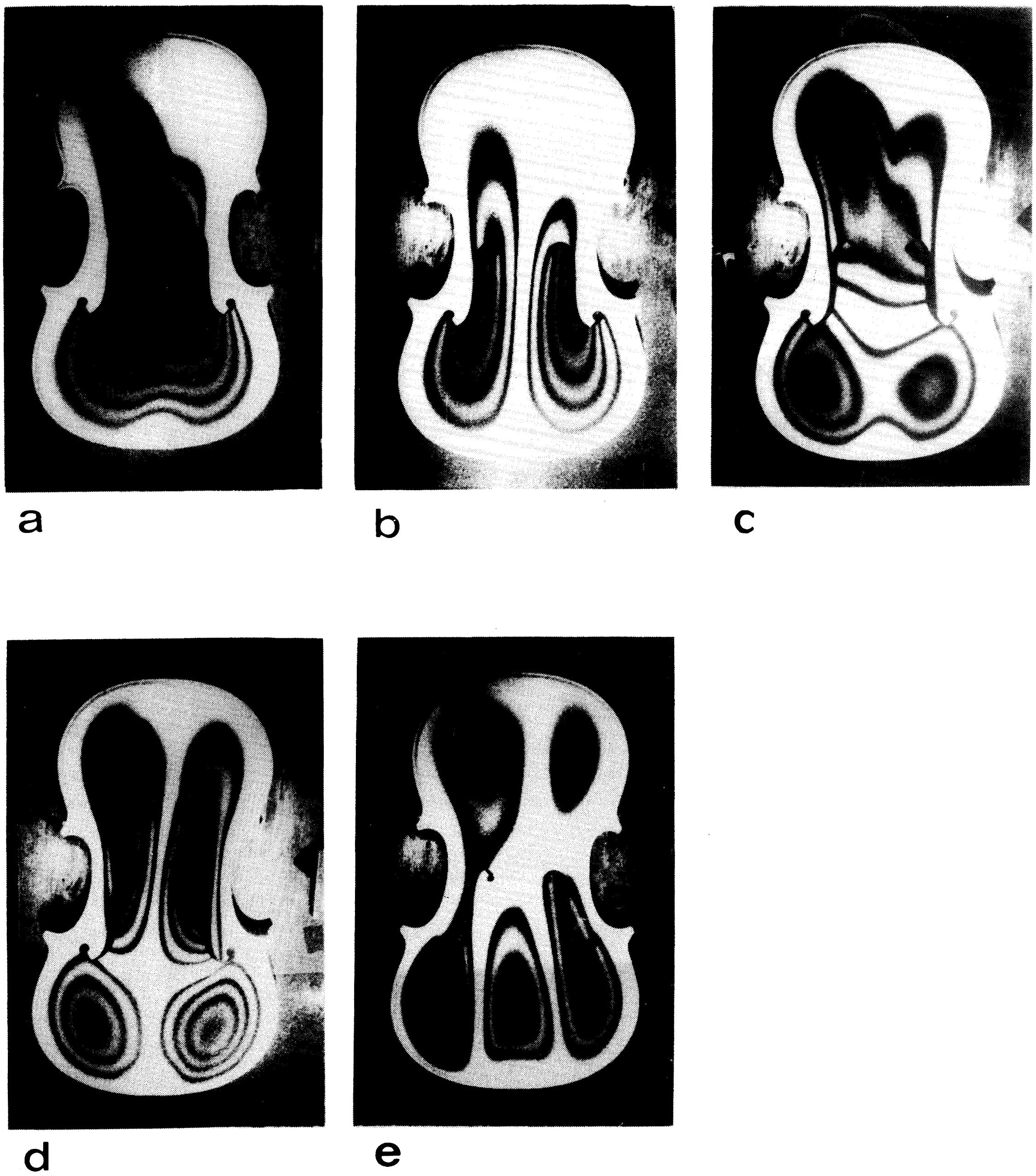



Fig. 16. Interferograms at resonance of a violin top plate with f-holes, bass bar, but no sound post at a) 465 Hz, b) 600 Hz, c) 820 Hz, d) 910 Hz and e) 1090 Hz. From JANSSON & al. (1970), Physica Scripta.

plate with bridge) which is rocking along an axis (the nodal line) on top of the sound post.

But there are several additional resonances in the violin, also in the top plate. Such vibration patterns are displayed in Fig. 16. Note that the number of vibration "mountains" increases with increasing frequency. Note furthermore that the waist of the violin has a clear tendency to divide the vibrations into two areas, one upper and one lower area. A rough estimation indicates that we can expect approx. one resonance per 200 Hz in each plate. This agrees quite well for the frequencies given in Fig. 16.

Let me show another example of the resonance-vibration patterns of a similar instrument, the guitar. Generally we recognize the same patterns in Fig. 17 as in Fig. 16, although the guitar plate and the violin plate are quite differently constructed. The similarity derives in large part from the fact that the edge shape and the edge vibration conditions are similar, thus demonstrating their importance. The vibrations of a guitar top is furthermore influenced by the bracing glued to the inner side of the plate (MEYER, 1974a and b). The plate has also an external bracing, the bridge. A little analysis of the interferograms shows that the plate is rather unwilling to bend perpendicularly to the bridge, i.e. the bridge gives a considerable stiffening effect - see especially Fig. 17 d with the middle "mountain" considerably lower than the two side ones. A rough estimate indicates that we should expext approx. three resonances per 200 Hz, which is in reasonable agreement with the resonance frequencies given in Fig. 17.

The air volumes of the violin and the guitar give two more examples of the importance of the shape. Vibration patterns somewhat similar to those of the plates can be found in the air cavity in spite of the sound holes (in the plates) of the instruments in Fig. 18. (JANSSON, 1976) In the violin as in the plates the waist makes the cavity act as two parts in certain cases, see especially Fig. 18:2 and 3. The resonance A1 exists in all violins depending on the position of the f-holes, as no or little sound radiates through the f-holes. Experiments indicate that interaction exists between plate and air volume resonances. (JANSSON &

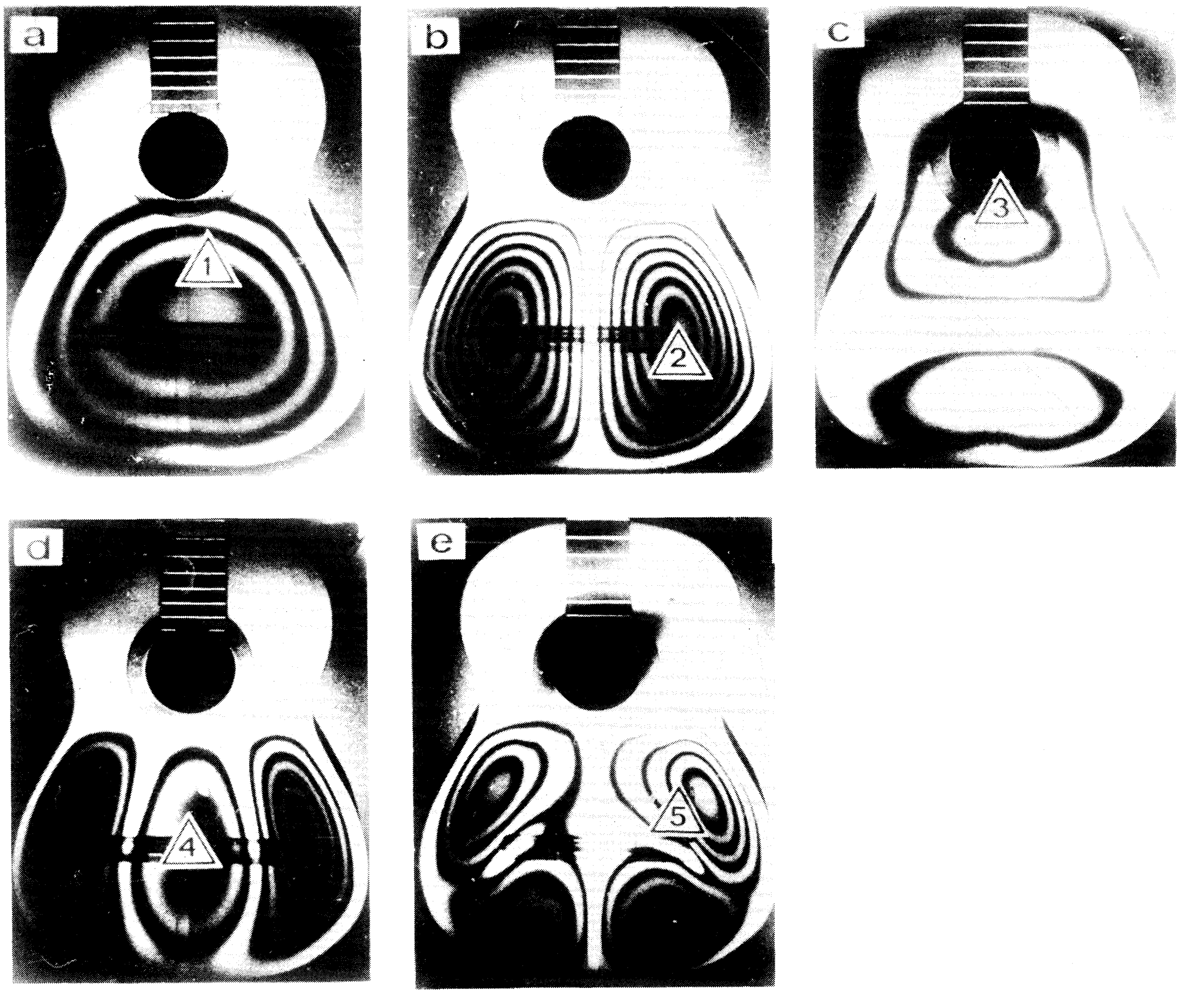



Fig. 17. Time-average interferograms (made by N.-E. Molin and K. Stetson) of a guitar top plate (G. Bolin, Stockholm) at resonance. Driving points (Δ), a) 185 Hz, b) 287 Hz, c) 460 Hz, d) 508 Hz, and e) 645 Hz. From JANSSON (1971), Acustica.

SUNDIN, 1974) The vibration patterns for an air cavity shaped as a guitar are shown in Fig. 19 cf. MEYER (1974b). The patterns are approximately the same as for the violin cavity. The waist of the guitar cavity, being less marked, seems to result in a less clear division into two halves. From Fig. 19:1 it can be seen that the resonance 1. radiates sound because of the position of the sound hole. The total number of resonances in the air volumes is considerable: 25 and 200 below 4 kHz for the violin and the guitar respectively.

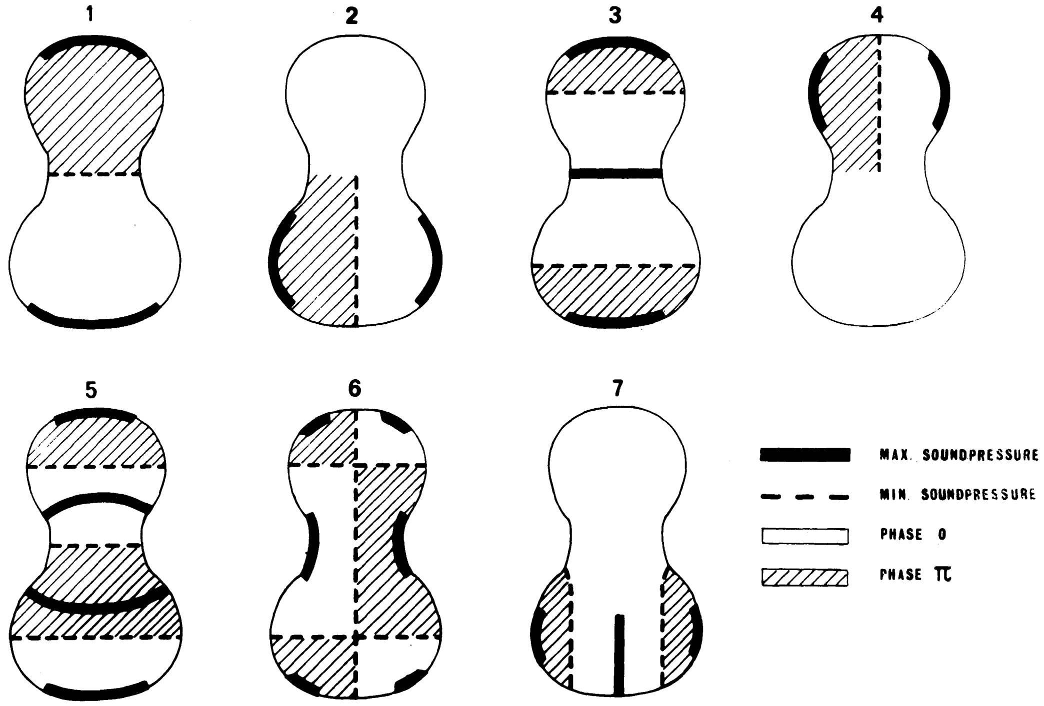



Fig. 18. Vibration patterns of the seven lowest resonances of a violin-shaped flat cavity at 460 Hz(1), 1040 Hz(2), 1130 Hz(3), 1300 Hz(4), 1590 Hz(5), 1800 Hz(6) and 1920 Hz(7). From JANSSON (1972), Report of the 11th Congress of International Musicological Society.

What do the resonance modes tell us then? First, they tell how willing the resonance box is to pick up the string vibrations. If the string of the guitar in Fig. 20 is plucked perpendicular to (up from) the top plate, the top plate obtains an initial deformation of one single vibratory "mountain" as seen in the upper curve. The top plate will thereafter vibrate in and out with the same deformation shape but with steadily decreasing magnitude. If the string is plucked in parallel with the top plate, the top plate obtains an initial deformation of a "mountain" and a trough as in the lower curve, and will thereafter vibrate in and out in the same shape but with decreasing magnitude. A comparison with Fig. 17 shows that the first and second deformations above correspond mainly to vibration modes 1 and 2, respectively. From this one can realize that the direction of plucking affects at least specific notes to a considerable extent.

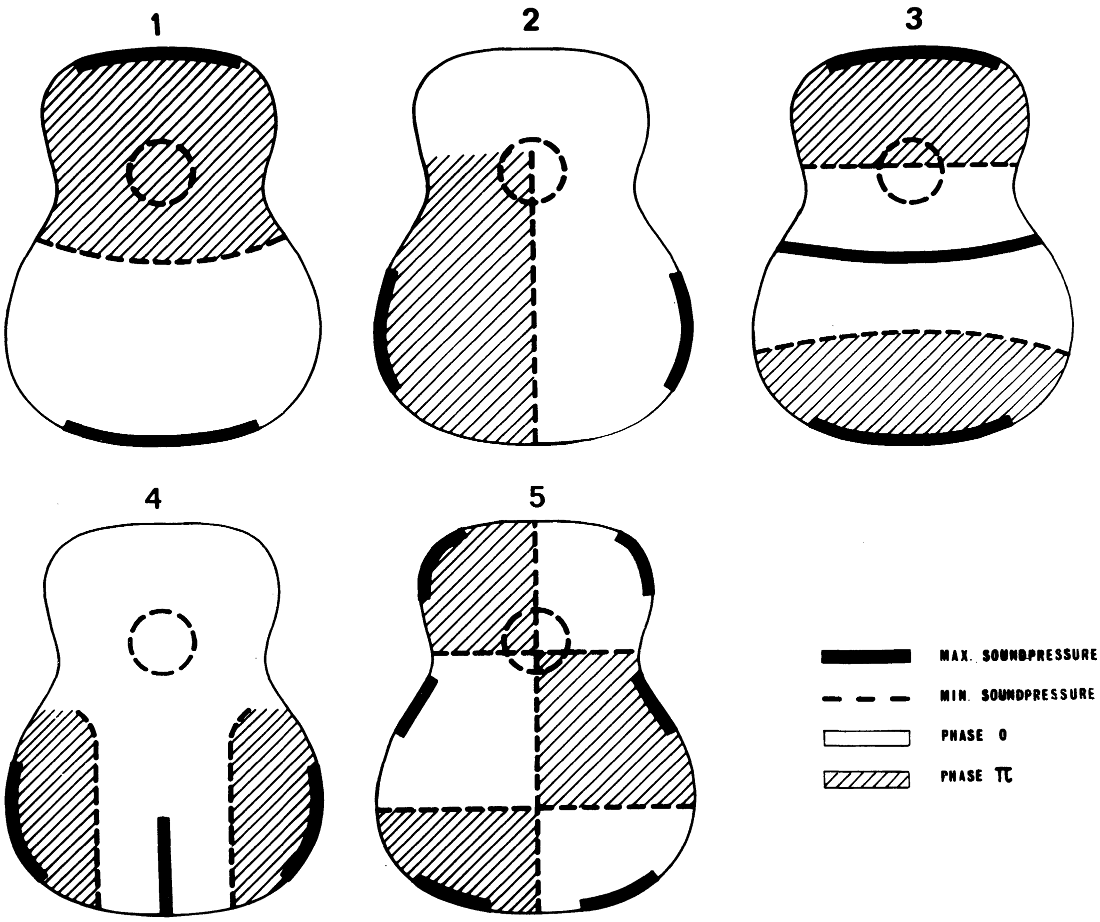



Fig. 19. Vibration patterns of the five lowest resonances of a guitar shaped cavity at 370 Hz(1), 540 Hz(2), 760 Hz(3), 980 Hz(4) and 1000 Hz(5). From JANSSON (1976), Acustica.

Secondly, the first top plate mode mentioned above constitutes a major resonance of the violin (JANSSON, 1973b). The vibration pattern tells a maker where to thin the plate in order to alter its resonance frequency (ÅGREN & STETSON, 1972). The practical importance of the correct tuning of this resonance has been proved by Carleen HUTCHINS (1962; 1967). Hutchins tunes her violin plates to desired resonance frequencies and tests them by means of electroacoustical equipment. (HUTCHINS & FIELDING, 1968) By doing so and by using design rules she can make good instruments routinely. The importance of the tuning of the first top plate mode has been further demonstrated by the construction of a new

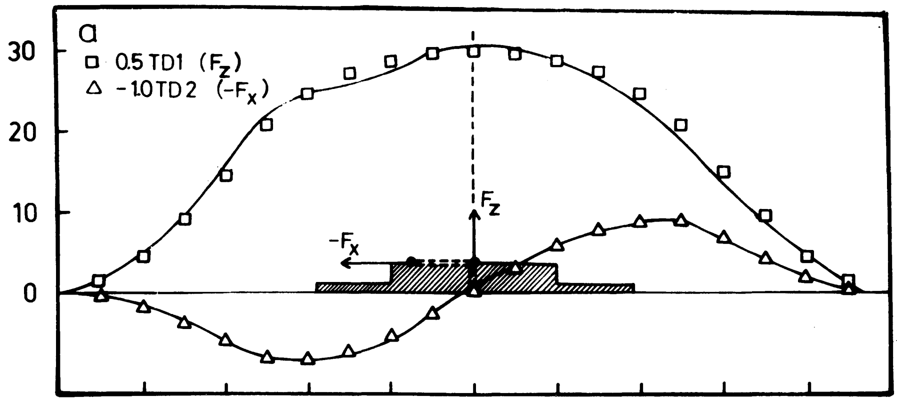



Fig. 20. Recorded and fitted deformations Δz of a guitar top plate along the brigde line (30 approx. equal to .ol mm) with differently applied forces: measured deformations —— and approximated ▫ for a force $F_z$ = 0,49 N and measured deformations —— and approximated Δ for a force $F_x$ = 0,98 N. (JANSSON,1973b)

family of violins. By scaling and designing the new instruments in such a way that the lowest air-mode, i.e. the Helmholtz-mode, and the first top plate mode both fall close in frequency to the middle open strings, a successful family of new instruments has been achieved. (HUTCHINS, 1967) The family consists of an ordinary violin, two smaller and higher tuned violins, and five larger and lower tuned ones down to a big bass violin. This proves that the position of the lowest resonances, i.e. the low frequency response, is very important.

Thirdly, the efficiency of sound radiation and hence the "amplifying" and the spectrum shaping of the resonance box can be calculated from the resonance modes. This, however, is a very complex procedure if we apply the general way to calculate the sound radiation from a vibrating source which is shaped as in these instruments. The resonances show up as constant frequency peaks in the sound radiation properties. (JANSSON, 1971) The top plate is likely to be the dominating sound emitter, but the contribution of the back plate and the ribs may be considerable. (CREMER & LEHRINGER, 1973) The optimum number of peaks in the amplification curve according to MATHEWS & KOHUT (1973) agrees well with the total estimated number of resonances of the top plate.

As indicated above, the musical timbre is less well understood than the pitch. Still some general rules can be given to the instrumentmaker.

Investigations of the acoustics of string instruments have shown that good instruments should have a reasonably even low frequency response, should give a shaping of the "musical-tone-chord" making some tones strong and others weak and should also have an amplification curve of a certain peakiness. Both material and shape are important to the function of the instrument.

For a good low frequency response a careful positioning of the lowest resonances is necessary. The frequencies of the lowest air volume resonance and the first wall resonance should be tuned to a specific relation adapted to the tuning of the instrument. A light and stiff wood is favourable as more sound can be obtained at the frequency of the first wall resonance. Small internal friction makes this resonance pronounced. The shape, size and thickness, influence as follows: Larger air volumes give lower air volume resonance frequencies. Thinner and larger plates give lower resonance frequencies. A less rigid joining of the plates to the sides gives lower resonance frequencies (as much as a factor two for a bar).

It can be shown that the best way to copy the acoustical parameters of an instrument is to scale the thickness of the new plates so that the frequency of the lowest resonance is maintained. By doing so, differences in the elastic properties of the wood are automatically compensated for and a copy of the complete response curve is obtained with only a level shift, cf. Fig. 14.

Good instruments should have certain regions of large response and others of weak to give the desired timbre. These regions are so broad that they contain several resonances. This indicates that the shape and the joining of the different parts determine these peaks. One possible way to obtain such peaks is by shaping the instrument so that resonances in certain regions are more effectively excited by the bridge vibrations. A second possibility is to make the resonances of certain frequency regions more effective as sound radiators. A third possibility is to chose a favourable bridge design, as the bridge affects the string sound transmitted to the resonance box.

Apart from these broad regions of large and weak amplification of the resonance box, the amplification curve of a responsive instrument should exhibit a moderate but still pronounced peakiness. This peakiness is likely to be most affected by the wood material and probably it is directly correlated to the number of resonances in the instrument. A small internal friction increases the magnitude of the peakiness.

The parameters regarding timbre presented above can fairly easily be measured with the acoustic measurement tools of today. The tools may also be simplified and further developed to fit the test procedures presented above. A common difficulty in acoustical measurements is that the room influences the results considerably, if not specially designed rooms are used. This difficulty can be reduced by employing a fixed test set up, which also is convenient for measurements on a larger scale. It should be pointed out, though, that the problems related to timbre qualities are not definitely solved. Parameters not mentioned above may turn out to be of critical importance.

## Outlook

This paper has tried to demonstrate the usefulness of research in the acoustics of music. Most of the material presented is new, and in some cases this has led to difficulties in predicting its practical importance. Still, some general conclusions have been drawn about what properties musical instruments should possess, what parameters can be worked with, and how good instruments can be designed. Some of the material presented is already in use, see e.g. HUTCHINS & FIELDING (1968) and DEKAN (1974).

Another aim of the present paper was to form an introduction and to give a status report in some areas of the acoustics of music. It is then logical to end the paper by pointing out possibilities to further studies in this field. A more thorough but still popular introduction to the acoustics of musical instruments the reader will find in a couple of articles published in the SCIENTIFIC AMERICAN (BENADE, 1960; BENADE, 1973; HUTCHINS, 1962; SCHELLENG, 1975). In "The Acoustical Foundations of Music" BACKUS (1969) offers an introduction to the complete field. In "Introduction to the Physics and Psychophysics of Music" ROEDERER

(1973) reviews our knowledge of hearing and combines it with music. A collection of major papers in acoustics is being edited in the series "Benchmark Papers in Acoustics". The volumes "Musical Acoustics: The Violin Family" and "Musical Acoustics: The Piano and Wind Instruments" provide easy access to important articles in the development of our present knowledge. "Acoustical Aspects of Woodwind Instruments" by NEDERVEEN (1969) is an excellent work containing calculation procedures for intonation. Today the most recent contribution is BENADE's "Fundamentals of Musical Acoustics" (1976) which gives a complete overview of the filed of acoustics of music.

Acknowledgments

In preparing this paper the author had the pleasure to study three chapters on wind instruments in manuscript of Arthur Benade's book "The Fundamentals of Musical Acoustics" which at that time was not yet published. This study forms the backbone for the section "Intonation and wind instruments", which is gratefully acknowledged.

This work was supported by the Swedish Humanistic Research Council and the Swedish Natural Science Research Council.

References:

BACKUS, J. (1969): The Acoustical Foundations of Music, W.W. Norton & Co., Inc., New York.

Benchmark Papers in Acoustics, (series ed. R.B. Lindsay), Dowden, Hutchinson & Ross, Inc.Stroudsbourg:

Vol. 5. Musical Acoustics: The Violin Family. Part I: Violin Family Components (1975) (ed. C.M. Hutchins). Part II: Violin Family Function (1976) (ed. C.M. Hutchins).

Vol. . Piano and Wind Instruments (ed. E.L. Kent) (forthcoming).

BENADE, A.H. (1960): "The Physics of Wood Winds", Scientific American, 203, Oct., pp. 144-154.

BENADE, A.H. (1973): "The Physics of Brasses", Scientific American, 229, July, pp. 24-35.

BENADE, A.H. (1976): Fundamentals of Musical Acoustics, Oxford University Press, New York.

BENADE, A.H. and GANS, D.J. (1968): "Sound Production in Wind Instruments", Ann. of the New York Acad. of Sciences, 155, pp. 247-263.

BENADE, A.H. and JANSSON, E.V. (1974): "On Spherical Waves in Horns with Nonuniform Flare. I: Theory of Radiation, Resonance Frequencies and Mode Conversion", Acustica 31, pp. 79-98. JANSSON, E.V. and BENADE, A.H. (1974): "On Plane and Spherical Waves in Horns with Nonuniform Flare. II: Prediction and Measurements of Resonance Frequencies and Radiation Losses", Acustica, 31, pp. 185-202.

CREMER, L. and LEHRINGER, F. (1973): "Zur Abstrahlung geschlossener Körperoberflächen", Acustica 29, pp. 137-147.

DĚKAN, K. (1974): "The Ionophone in Musical Instruments Research", Hudební Nástroje 11, pp. 49-53 (in Czech).

FRANSSON, F. (1963): "Measurements on Flutes from Different Periods. Part I", STL-QPSR 4/1963*, pp. 12-16.

FRANSSON, F.J. and JANSSON, E.V. (1975): "The STL-Ionophone: Transducer Properties and Construction", J. Acoust.Soc.Am. 58, pp. 910-915.

HUTCHINS, C.M. (1962): "The Physics of Violins", Scientific American, 207, Nov., pp. 78-93.

HUTCHINS, C.M. (1967): "Founding a Family Fiddles", Physics Today, 20, No. 2, pp. 27-38.

HUTCHINS, C.M. and FIELDING, F.L. (1968): "Acoustical Measurements of Violins", Physics Today, 21, No. 7, pp. 34-40.

JANSSON, E. (1966): "Analogies Between Bowed String Instruments and the Human Voice, Source-Filter Models", STL-QPSR 3/1966*, pp. 4-6.

JANSSON, E.V. (1971): "A Study of Acoustical and Hologram Interferometric Measurements of the Top Plate Vibrations of a Guitar", Acustica, 25, pp. 95-100.

JANSSON, E.V. (1972): "On the Acoustics of the Violin", pp. 466-477 in Report of the 11th Congr. of the Internat.Musicological Soc., (ed.H.Glahn, S. Sørensen, P. Ryom), Vol. II, Edition Wilhelm Hansen, Copenhagen (1974).

JANSSON, E.V. (1973a): "An Investigation of a Violin by Laser Speckle Interferometry and Acoustical Measurements", Acustica 29, pp. 21-28.

JANSSON, E.V.(1973b): "Coupling of String Motions to Top Plate Modes in a Guitar", STL-QPSR 4/1973*, pp. 19-38.

JANSSON, E. (1975): "Om materialets, tjocklekens och fastsättningens betydelse för böjsvängningar i och ljudutstrålning från trä", Speech Transmission Laboratory, Technical Report, March 15, 1975.

---

* STL-QPSR stands for the Speech Transmission Laboratory, Qarterly Progress and Status Report.

JANSSON, E.V.(1976): "Acoustical Properties of Complex Cavities. Prediction and Measurements of Resonance Properties of Violin-Shaped and Guitar-Shaped Cavities", Acustica (in the press).

JANSSON, E.V. and SUNDBERG, J. (1975): "Long-Time-Average-Spectra Applied to Analysis of Music. Part I: Method and General Applications", Acustica, 34, pp. 15-19. "Part II: An Analysis of Organ Stops", Acustica, 34, pp. 269-274. "Part III: A Simple Method for Surveyable Analysis of Complex Sound Sources by Means of a Reverberation Chamber", Acustica, 34, pp. 275-280.

JANSSON, E. and SUNDIN, H. (1974): "A Pilot Study on Coupling Between Top Plate and Air Volume Vibrations", Catgut Acoustical Society Newsletter 21, May.

JANSSON, E., MOLIN, N-E. and SUNDIN, H. (1970): "Resonances of a Violin Body Studied by Hologram Interferometry and Acoustical Methods", Physica Scripta, 2, pp. 243-256.

LOTTERMOSER, W. (1968): "Analysen von Geigenklängen aus solistischen Darbietungen", Instrumentenbau-Zeitschrift, Heft 4-5, April/May.

MARTIN, D.W. and WARD, W.D. (1961): "Subjective Evaluation of Musical Scale Temperament in Pianos", J. Acoust.Soc.Am., 33, pp. 582-588.

MATHEWS, M.V. and KOHUT, J. (1973): "Electronic Simulation of Violin Resonances", J. Acoust.Soc.Am., 53, pp. 1620-1626.

MEYER, J. (1974a): "Die Abstimmung der Grundresonanzen von Gitarren", Das Musikinstrument, 23, pp. 179-186.

MEYER, J. (1974b): "Das Resonanzverhalten von Gitarren bei mittleren Frequenzen", Das Musikinstrument, 23, pp. 1095-1102.

NEDERVEEN, C.J. (1969): Acoustical Aspects of Woodwind Instruments, Frits Knuf, Amsterdam.

ROEDERER, J.G. (1973): Introduction to the Physics and Psychophysics of Music, The English Universities Press, Ltd., London, and Springer Verlag, New York, Heidelberg, Berlin.

SCHELLENG, J.C. (1963): "The Violin as a Circuit", J. Acoust.Soc.Am., 35, pp. 326-338 and p. 1291.

SCHELLENG, J.C. (1975): "The Physics of the Bowed String", Scientific American, 230, Jan., pp. 87-95.

SUNDBERG, J. (1967): "The Scale of Musical Instruments", Svensk Tidskrift för musikforskning, 49, pp. 119-133.

SUNDBERG, J. and TJERNLUND, P. (1971): "Real Time Notation of Performed Melodies by Means of a Computer", Paper 21S6 in the Proc. of the VIIth Int. Congr. on Acoustics, Budapest.

WOGRAM, K. (1972): "Ein Beitrag zur Ermittlung der Stimmung von Blechblasinstrumenten", Dr. Thesis, TU Braunschweig (Verlag das Musikinstrument, Frankfurt a. Main).

WORMAN, W.E. (1971): "Self-Sustained Non-Linear Oscillations of Medium Amplitude in Clarinet-Like Systems", Dr. Thesis, Case Western Reserve University, Cleveland, USA.

YANKOVSKII, B.A. (1966): "Methods for the Objective Appraisal of Violin Tone Quality", Soviet Physics-Acoustics, 11, pp. 231-245.

ÅGREN, C.H. and STETSON, K.A. (1972): "Measuring the Resonances of Treble Viol Plates by Hologram Interferometry and Designing an Improved Instrument", J. Acoust.Soc.Am., 51, pp. 1971-1983.

# PART I: ACOUSTICAL CRITERIA AND ACOUSTICAL QUALITIES OF CONCERT HALLS

by
Vilhelm Lassen Jordan, Roskilde, Denmark

## PART I: DEVELOPMENT OF ACOUSTICAL CRITERIA

### Introduction

One of the few early physical criteria applied to assess the quality of a concert hall was reverberation time, defined by the slope of the statistical reverberation process as the time interval corresponding to a level difference of 60 dB.

The last 10-20 years have seen a development of criteria more concerned with the short time interval succeeding the emittance of a very short sound pulse. This first interval is associated with the "direct sound" from the source and the "early (distinct) reflections" from the boundaries of the hall, followed by the rapid transition of a growing number of reflections into the statistical process of reverberation.

It has grown increasingly evident, that the acoustical properties of any big hall is more intimately connected with the early transient rather than with the later statistical process and that quantitative criteria, if applicable to subjective assessments, must also be associated with this early transient process. Many research workers making individual approaches to this problem have independently developed several different criteria more or less interrelated with each other. Without being exhaustive, we shall expose a number of these criteria.

### A "family" of criteria

In principle, from this "family", the individual members could all be deduced by applying the impulse method (even if some of them were proposed prior to the introduction of this method). (cf. SCHROEDER, 1965) The first criterion was "Deutlichkeit", D, as defined by Thiele (1953):

$$D = \int_0^{50} p^2 dt \Big/ \int_0^{\infty} p^2 dt$$

The basic idea was that the useful sound energy is the one arriving in the first 50 milliseconds. This, of course, is closely associated with the concept of articulation for speech and only vaguely with musical sounds.

Another early criterion associated more directly with the quality of musical sounds was "rise time", defined as the time interval within which half the sound energy (of the complete process) arrives. This definition is associated with the idea that the rise period of musical sounds must not be too long and that the arrival of 50 % of the energy indicates the rapidity of the room response. (JORDAN, 1959)

The energy (related) measures were further developed by BERANEK & SCHULTZ (1965) and independently also by REICHARDT & SCHMIDT (1966):

Reverberant energy level/early energy level =

$$= 10 \log \int_{50}^{\infty} p^2 dt \Big/ \int_{0}^{50} p^2 dt \qquad \text{(Beranek, Schultz)}$$

"Hallabstand" = Direct energy level/reverberant energy level (Reichardt)

The early decay measures, likewise introduced in relation to musical quality, are:

initial reverberation time (interval 0-15 dB or 160 msec), introduced by SCHROEDER & al. (1965),

early decay time (EDT, interval 0-10 dB), introduced by JORDAN (1968).

Criteria related to direction or shape

The preceeding section discusses exclusively criteria which do not distinguish between directions of early reflections. It has, however, for quite a few years been known that lateral reflections create more impression of "reverberance" than non-lateral reflections. If you apply this experience to an energy measure like reverberant/early energy (or to the reciprocal measure) you may as SCHROEDER & al. did (1966) define a

directional distribution factor as:
early/reverberant energy (lateral)/early/reverberant energy (non-lateral). Other criteria are more concerened with the subjective perception of "spatial responsiveness" (MARSHALL, 1967-68) or "apparent source width" (KEET, 1968). The latter is correlated with an objective criterion, namely: fraction of incoherent lateral energy within 50 msec. This has been proven by Keet.

Still another proposal to involve the room shape (and at the same time the hearing conditions of the orchestra members) has been suggested by JORDAN (1968). The definition of inversion index assumes that the same criterion (EDT or rise time) has been measured and averaged over a certain number of positions on the stage and in the audience area (of a concert hall). The inversion index, I.I., then, is defined as:

I.I. = EDT average, auditorium/EDT average, stage, or
I.I. = rise time average, auditorium/rise time average, stage.

The numerical value of I.I. is judged preferable when it is larger than or equal to 1.0, based on the assumption that conditions in the stage area should reach stationary values earlier than in the audience area.

Correlation between subjective impressions and objective criteria

The above mentioned criteria are only examples showing an increasing number and complexity of these criteria. The question of correlation between the objective measures and subjective impressions of acoustical quality therefore becomes more and more urgent. This question has been pursued now for more than a decade and with growing intensity.

By applying synthetic sound fields simulating the "direct sound" and the "first reflections", using loudspeaker signals in an anechoic room, it has been possible to establish certain basic relations.REICHARDT & SCHMIDT(1966) e.g. used 2 speakers in a frontal position with undelayed sound and 4 speakers in a circle with delayed, reverberant sound, to establish a correlation between "Räumlichkeit" and level difference between direct and reverberant sound (Fig. 1). Having established this, they went further and added "side reflections" or "above reflections" with different delays. One of the recent comprehensive results of these studies

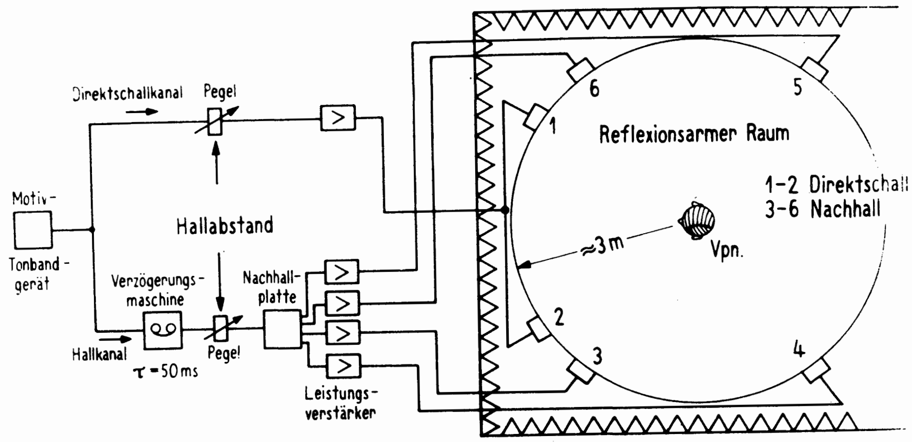



Fig. 1. Synthetic sound field. From REICHARDT & SCHMIDT (1966), Acustica.

is given as a table indicating the effect of single reflections (depending on delay and orientation), increasing either "reverberance" or "direct sound" (REICHARDT & al. 1975):

Principal Effect of Single Reflections

| Group | t,time delay (msec.) | angle to direct sound | increase of |
|---|---|---|---|
| I | $t < 25$ | any | dir.sound |
| II | $25 \leq t \leq 80$ | $<40^{o}$ | dir.sound |
| III | $25 \leq t \leq 80$ | $>40^{o}$ | reverberance |
| IV | $t > 80$ | any | reverberance |

Contrary to earlier results this grouping does not distinguish between lateral and non-lateral reflections and this may prove to be an oversimplification. It may well be that the important group III needs further analysis. That lateral reflections in this group contributes more to "reverberance" than non-lateral cannot be disputed.

One of the results of this research was also the proposal of yet another objective criterion (related to the subjective impression of "Durchsichtigkeit") namely clarity, C:

$$C = 10 \log \int_0^{80} p^2 dt \Big/ \int_{80}^{\infty} p^2 dt \qquad \text{(dB)}$$

relating the energy arriving in the first 80 msec. to the energy arriving later. Musical samples have indicated that the value of C = 0 is adequate for certain Mozart motifs whereas other styles require lower values of C.

Incidentally, the value of C = 0 obviously corresponds to a "rise time" of 80 msec. since the arrival of 50 % of the energy corresponds to a level (of the total energy of -3 dB).

Other approaches to the problem of correlation between subjective impressions and objective criteria have been made in recent years in several institutes. Two prominent examples must be mentioned, one from Göttingen and one from West Berlin. (SCHROEDER & al., 1974; GOTTLOB, 1973; PLENGE & al., 1975)

Both research groups have applied a new method of registration (of musical programmes) over artificial heads, to get signals corresponding closely to the actual signals at the ears of a listener. When replaying the Göttingen group applies two speakers in an anechoic room (with compensation filters to eliminate "false" signals) whereas the Berlin group use headphones. The Göttingen group applies "dry music" radiated over loudspeakers (in a number of concert halls) and the Berlin group followed the Berlin Philharmonic Orchestra on a tour to 6 different West-German concert halls. The integrated pulse method was used in all cases to provide any of the objective criteria (of the "family" group).

When trying to correlate subjective impressions with objective criteria the Göttingen group relied almost exclusively on pair comparisons whereas the Berlin group worked with a binary scale of (16) different subjective judgements.

The Berlin group claims that differences in assessments of interpretation (of the same musical programmes) are small compared to differences in assessments of acoustical qualities.

Preliminary conclusions of the two research tasks seem to indicate that the number of acoustical qualities which may be distinguished and correlated with objective criteria are 3 or 4. Taking into consideration the work of both groups and also the work of Reichardt and his coworkers, it is possible, on a preliminary basis, to establish a list of qualities and related criteria:

| Quality | Criteria |
|---|---|
| Volume of the sound | Loudness level |
| Reverberance | EDT,<br>Reverberant/early energy,<br>Point of gravity time |
| "Durchsichtigkeit" | Clarity,<br>"Rise Time" |
| Tonal Quality | Frequency dependence of EDT<br>or other criteria |

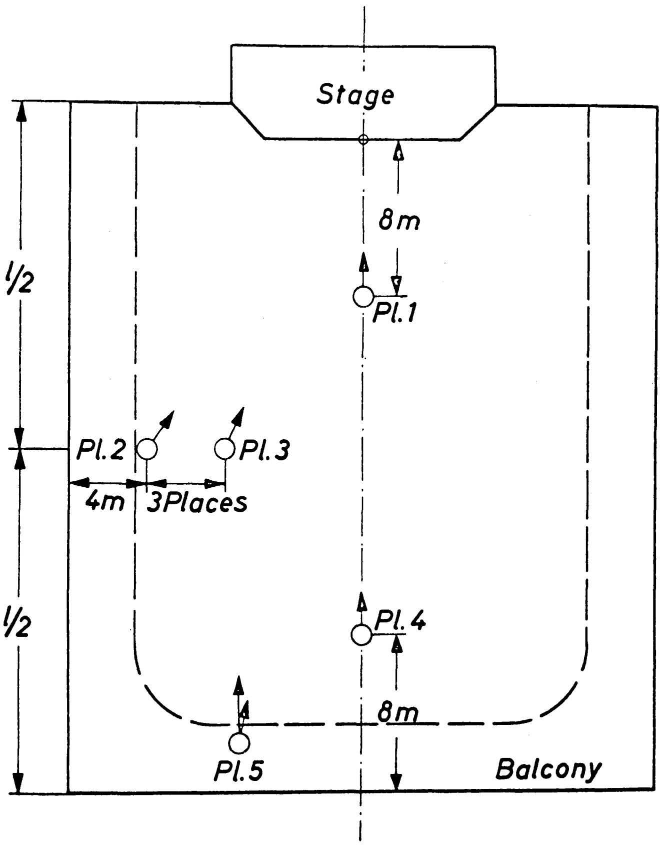



Fig. 2. Listening and measuring positions (schematically) applying to (6) West-German concert halls.

Analysis of I.I. and EDT from some published results

None of the qualities include a comprehensive evaluation of the listening properties of the stage compared with the listening properties in the audience area. It has already been indicated that this could be linked with the concept of inversion index as calculated from measurements of the same criterion in both areas.

Although all measurements (published by the two research groups mentioned above) only include measurements in the audience area, it is possible to deduct some values of inversion index from the results of the Berlin group. They used (5) different locations (as indicated in principle in Fig. 2) and as a first approximation we may use location 1 as "stage location".

With reservation for the limited number of locations and individual values we can then calculate I.I. applying either EDT or "rise time" (denoted as TR). The values have been normalized to "expectation values" (values expected by completely statistical processes).

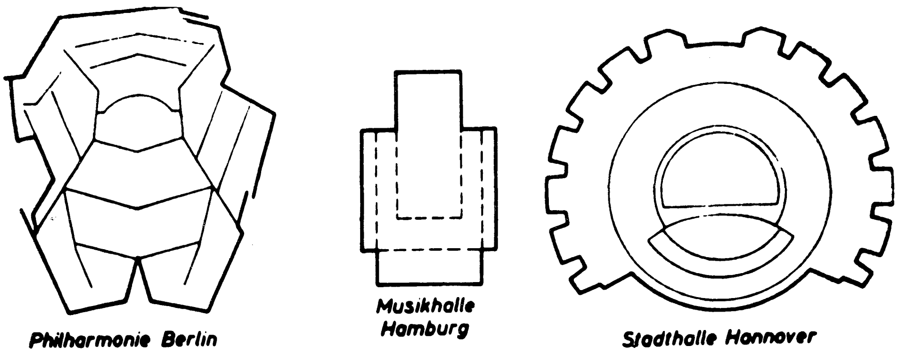
The variations of a criterion like EDT over the audience area of a concert hall is also used as an indication of homogenity for that particular hall. These variations as expressed by

$$100 \cdot \Delta EDT, max/EDT, average \qquad (\%)$$

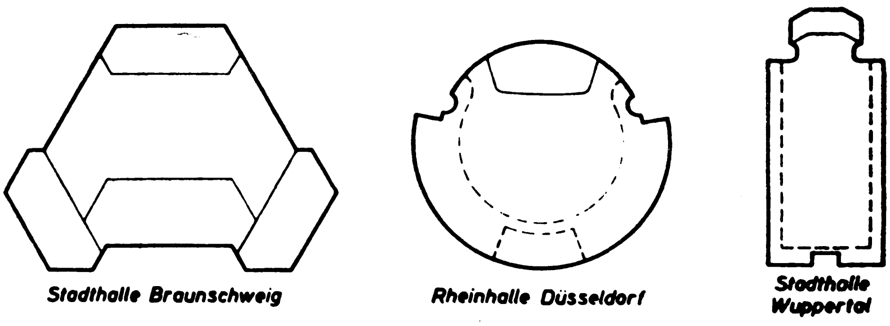
have likewise been calculated from the Berlin results.

Fig. 3 shows the outcome of these calculations for the (6) different concert halls. It is remarkable that I.I. calculated from TR (Rise Time) values show larger variations than those calculated from EDT values. It is also remarkable that the values of I.I. (especially calculated from TR values) indicate the strong influence of the gross shape on this criterion. Arena shape or round shape show smaller values than rectangular shape.

Fig. 4 shows a single instance of the subjective appraisals of the six halls by a group of individuals. There is a clear tendency for some of the halls to divide the groups (according to differences of taste?). On




| | | | |
|---|---|---|---|
| I. I. (EDT) | 0.93 | 0.96 | 0.90 |
| I. I. (TR) | 0.57 | 1.41 | 0.95 |
| EDT, variation % | 17.8 | 8.6 | 14.0 |




| | | | |
|---|---|---|---|
| I. I. (EDT) | 1.13 | 0.83 | 1.03 |
| I. I. (TR) | 1.11 | 0.73 | 1.31 |
| EDT, variation % | 21.0 | 27.0 | 11.0 |

| Raum: | Form: | Plätze: | Volumen: | $T_{500-1000 Hz}$ |
|---|---|---|---|---|
| Berliner Philharmonie | Arena | 2200 | ≈ 25000 m³ | 2,0 s |
| Musikhalle Hamburg | Rechteck | 1980 | ≈ 11600 m³ | 2,2 s |
| Stadthalle Hannover | Rundbau | 3660 | ≈ 34000 m³ | 2,0 s |
| Stadthalle Braunschweig | Sechseck | 2166 | ≈ 19000 m³ | 1,9 s |
| Rheinhalle Düsseldorf | Rundbau | 1842 | ≈ 33000 m³ | 2,5 s |
| Stadthalle Wuppertal | Rechteck | 1614 | ≈ 25000 m³ | 2,7 s |

Fig. 3. Calculated values of I.I. (from measured values of EDT and TR) and EDT variation in the (6) concert halls.

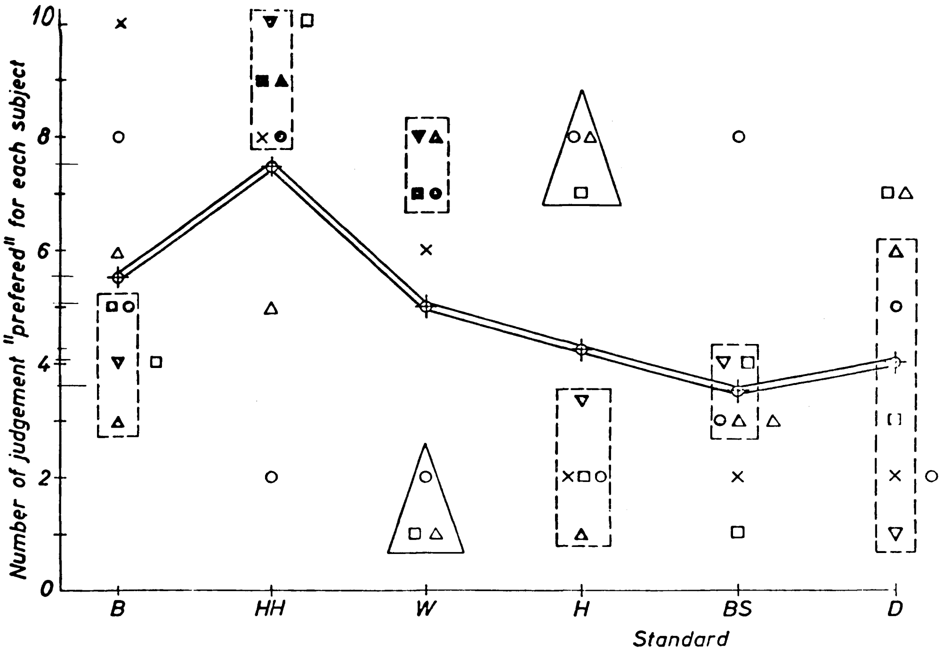



Fig. 4. Average (⊕) and single results of a subjective evaluation of the (6) concert halls at location on center in the back. From PLENGE & al. (1975), Journal of the Acoustical Society of America.

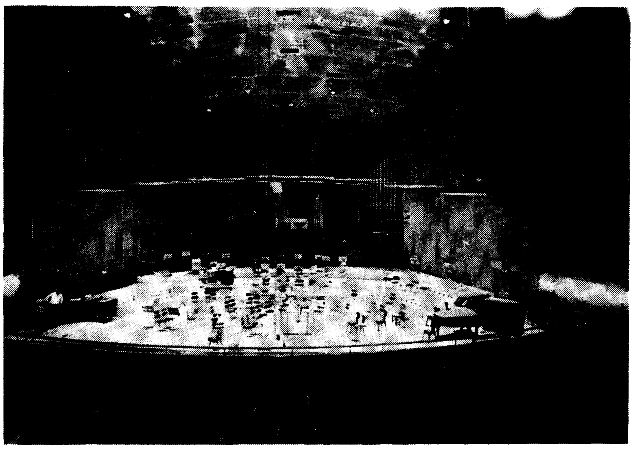


Fig. 5. Concert Studio of "Radiohuset", Copenhagen, view towards the stage.

the other hand there is also at least one example of a hall where all individuals agree on the good quality. This is a rectangular hall with high inversion index and low variation of EDT.

Recent research on correlation of quality and criterion

Quite recently the Göttingen Institute has published some new results which, although they are preliminary, contribute essentially to the understanding of the relative importance of the different objective criteria. (EYSHOLT & al., 1975; GOTTLOB & al., 1975)

By questioning a group of individuals about the relative "reverberance" ("Halligkeit") of a number of samples it was possible to establish a function showing the correlation coefficient of objective criteria with the subjective reverberance.

This function showed that criteria relating to the time interval between 60 and 160 msec. have the maximum of correlation with "reverberance" (close to a correlation coefficient of 0.90). This becomes even more critical if we limit the RT to values between 1.7 and 2.3 sec.

In another context the question was asked whether it is possible (at all) to distinguish between "Räumlichkeit" and "Halligkeit". The answer at least preliminary is no. There is a correlation coefficient of 0.76 between subjective assessments of the two (assumed) qualities.

Also interesting is an investigation of the criteria which correlates best with a consensus scale. The energy of the reverberant field correlates much better than RT, but even better is the correlation if you include the early energy of the lateral reflections. The correlation coefficient then becomes 0.80. This result, although preliminary, might be compared with the previous mentioned grouping of reflections in which no distinction between lateral and non-lateral reflections was made.

Without concluding too definitely it appears as if "clarity" may apply to non-directional concepts, but that "reverberance" must include early lateral energy and therefore call for measuring methods including directional criteria.

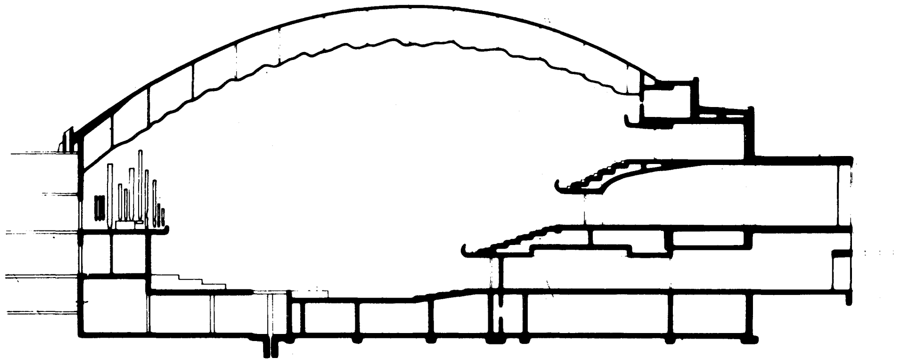


Fig. 6. Longitudinal section of the Concert Studio of "Radiohuset", Copenhagen.

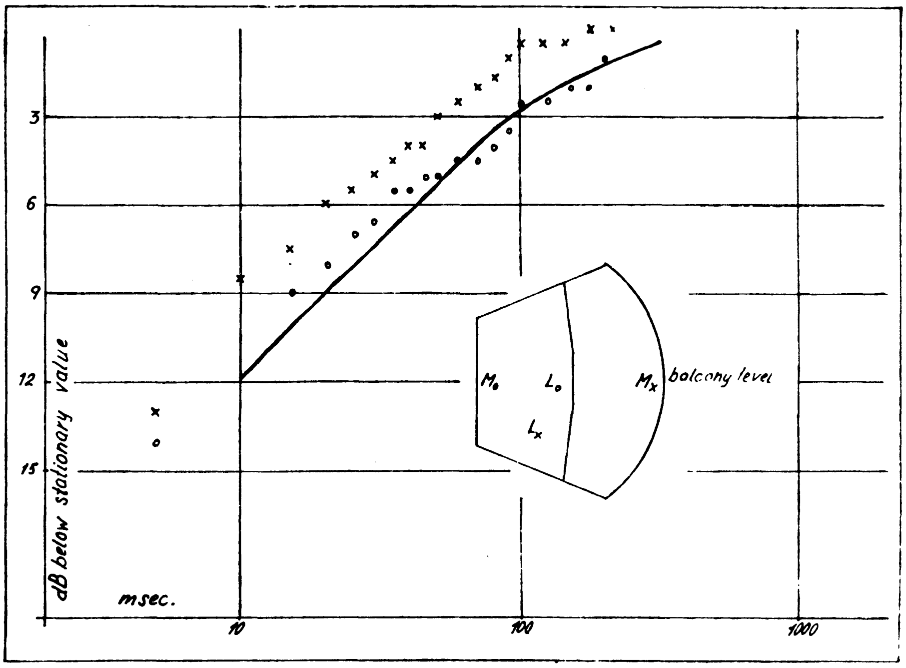



Fig. 7. Typical values of rise time including plan of the Concert Studio of "Radiohuset", Copenhagen.

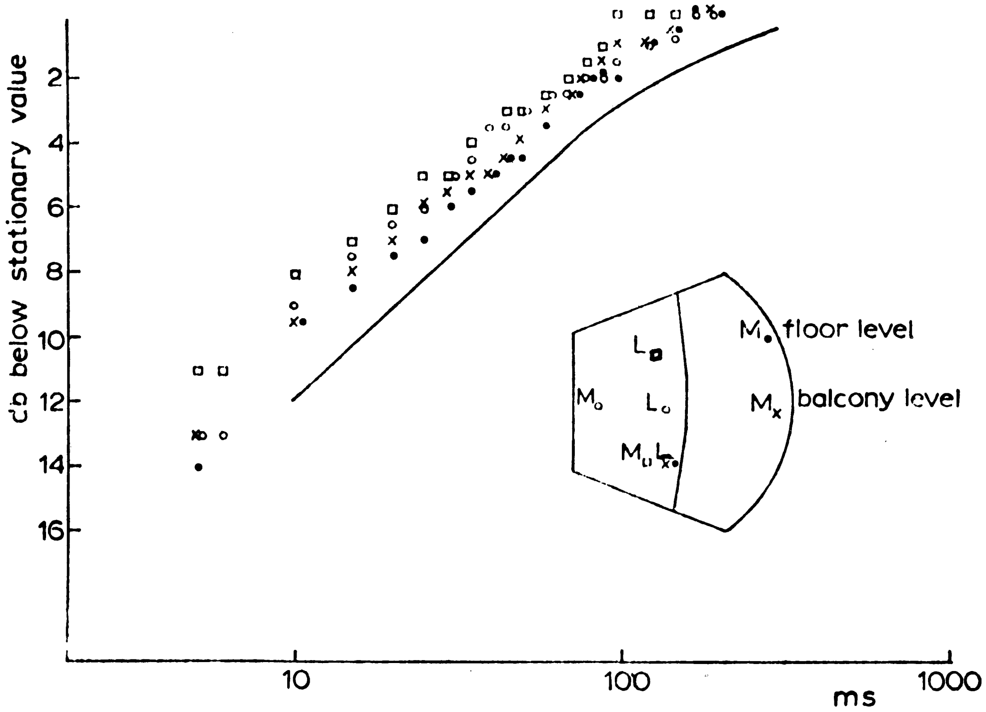



Fig. 8. Concert Studio of "Radiohuset", Copenhagen, values of rise time when reflectors are present. From JORDAN (1959), Proceedings of the Third International Congress of Acoustics.

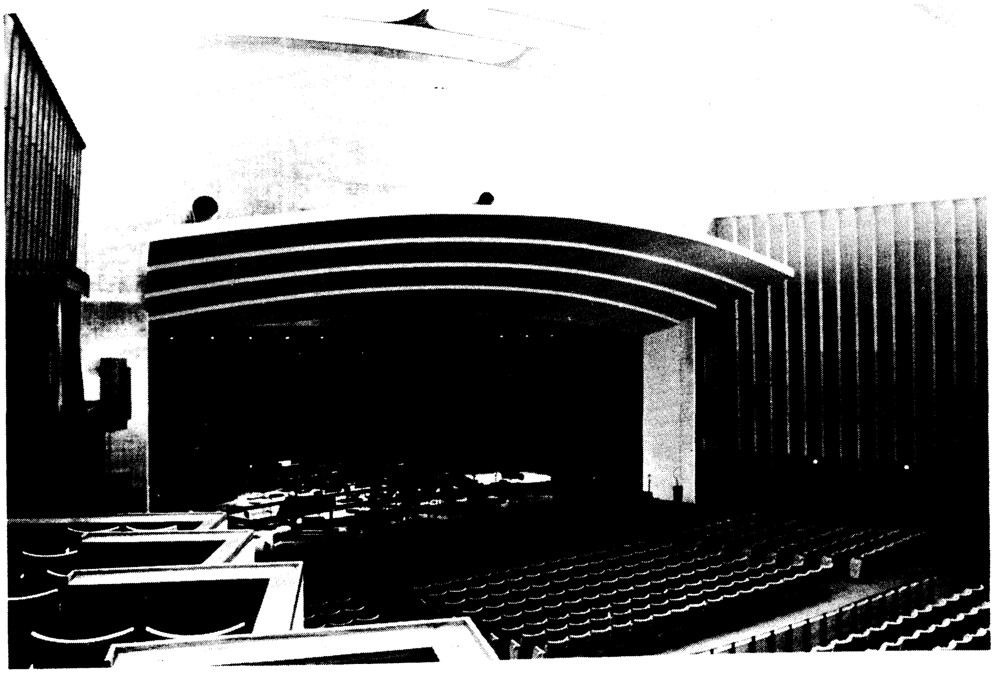


Fig. 9. Concert Hall of Tivoli, Copenhagen, view towards the stage.

# PART II: EXPERIENCES FROM VARIOUS HALLS INAUGURATED IN THE POSTWAR PERIOD

## Introduction

Originally, the classical concert halls were not tested acoustically, since objective criteria and measuring methods belong to this century. Any important new hall inaugurated in the postwar period, however, has been tested acoustically either before or after inauguration, or both. With one exception, the selection of halls is taken from the authors own personal sphere of experience. The exception is Philharmonic Hall in New York (recently renamed Avery Fisher Hall) which is included on account of its specific problems and the extensive, published reports of the investigations.

The year of inauguration of each hall is stated in parenthesis.

The methods of testing have been developed parallel to the applied objective criteria and, as part I has shown, this development has not been finalized yet.

## Concert Studio of "Radiohuset", Copenhagen (1945)

This studio is a good example of a hall where the acoustical problems did not show up at the inauguration but revealed themselves gradually.

Today, it may seem almost evident, when looking on a view towards the stage (Fig. 5) or on a longitudinal section (Fig. 6), that the very open and fan shaped stage (Fig. 7) could create special problems for the orchestra (and the conductor).

Fig. 5 also shows how these problems were tackled: by introducing reflectors suspended at a certain height over the podium. The subjective effect for the orchestra members was beyond discussion and to check this effect objectively the criterion "rise time" was introduced and measured for the first time, (JORDAN, 1959; see also Fig. 7 and 8, and Table 1)

The introduction of the reflectors has especially improved conditions on the stage but has not solved all problems in this auditorium (e.g. a certain lack of reverberance).

Table 1

Values of rise time of the "Radiohuset" Concert Studio, Copenhagen

| Source location (on platform) | Microphone location | no reflectors rise time (msec) | reflectors above stage rise time (msec) |
|---|---|---|---|
| left | platform right | 70 | 48 |
| center | platform rear | 105 | 55 |
| right | audience area (6'row) | 60 | 70 |
| right | audience area (balcony-first row) | 50 | 60 |

Concert Hall of Tivoli, Copenhagen (1956)

This hall was designed precisely in the period where the problems of the Concert Studio of "Radiohuset" were investigated. The design of the stage in the Tivolihall was thus influenced by the experiences with the Concert Studio.

Fig. 9 gives an impression of the stage which to both sides and above is carefully framed to provide early reflections for the benefit of the orchestra. Measurements of rise time were conducted by the same method as in the Concert Studio and the results showed considerably lower values of rise time in the stage area so that inversion was avoided.

Avery Fischer Hall (earlier called Philharmonic Hall) New York (1962)

It is well known that this hall right from the beginning was exposed to unfavorable comments, expecially that it was deficient in reverberance at medium and high frequencies (although reverberation time had been measured to around 2 sec). The specific arrangement of suspended "clouds", not only above the stage area but extending over two thirds of the length of the hall (Fig. 10), had the purpose of creating early reflections. These reflections (only occurring at medium and high frequencies) from above did not increase the impression of reverberance but rather increased the impression of direct sound.

This phenomenon was disclosed by several measurements in the hall undertaken partly by BBN (SCHULTZ, 1965) and partly by Bell-lab. (SCHROEDER & al. (1966) One of the criteria applied was the ratio of early to reverberant sound (or the reciprocal value). The values of this criterionare shown in Fig. 11 and 12. When the reflectors were moved up to a position just below the ceiling, conditions were improved (Fig. 12). This became even more evident by the measurements of directional distribution actor (ratio of early/reverb. energy for lateral and non-lateral directions (Fig. 13).

The Avery Fischer Hall has went through guite a number of changes and has not been accepted as satisfactory yet. A total reconstruction of the interior has been announced to take place.

University Hall, Reykjavik (1965)

The design of this hall was influenced by the dual purpose (concert/ cinema) which necessitated means of varying reverberation time (by movable panels on the sawtoothed sidewalls). Fig. 14 and 15 show plan and longitudinal section. The stage has overhead reflector panels but originally it was only partly closed at the sides and at the rear.

An investigation, using the integrated pulse method and measuring both EDT and RT values, especially on the stage, made it possible to compare these values for different octave bands. The tendency towards EDT values being lower than RT values was neutralized when more efficient screening of the stage (at the sides and towards the rear) was undertaken (Fig. 16).

The New Metropolitan Opera, New York (1966)

A new development in the design of concert halls and opera houses becomes apparent in the sixties. It is no longer considered satisfactory to rely on precalculations of reverberation time (from drawings and

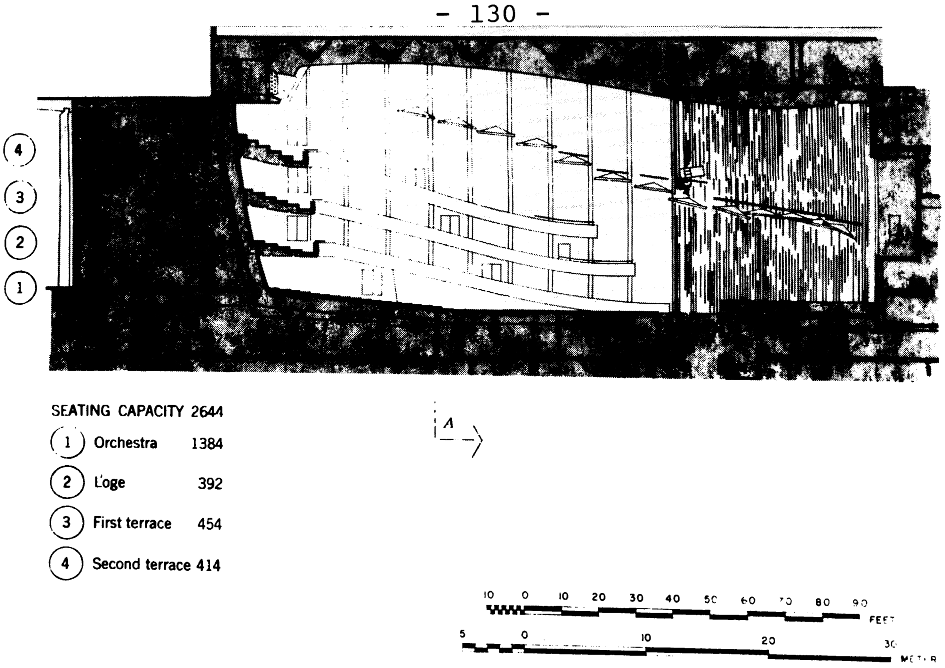



Fig. 10. Philharmonic Hall, New York, longitudinal section. From BERANEK (1962): Music, Architecture and Acoustics, John Wiley & Sons.

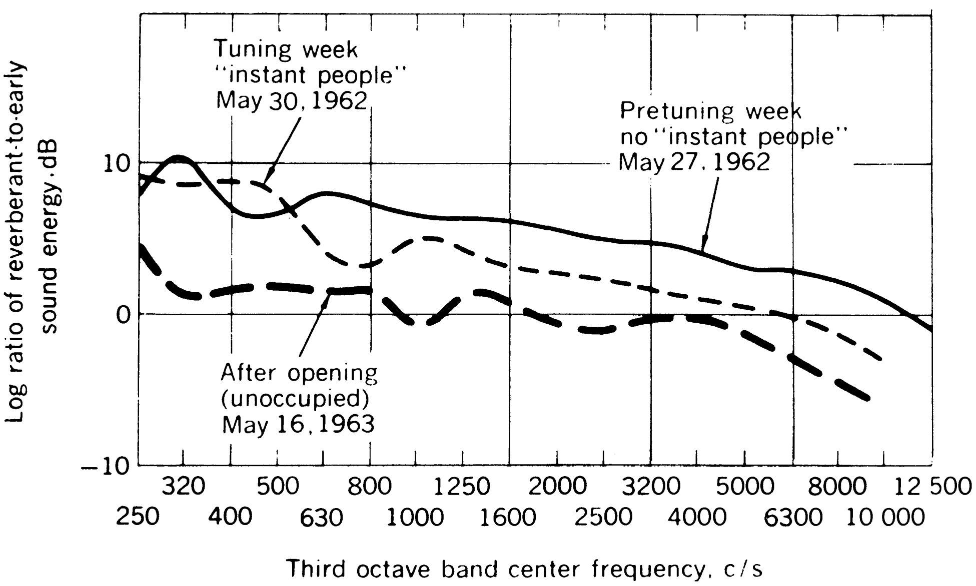



Fig. 11. Ratio of reverberant to early sound energy for the New York Philharmonic Hall. From SCHULTZ (1965), IEEE Spectrum.

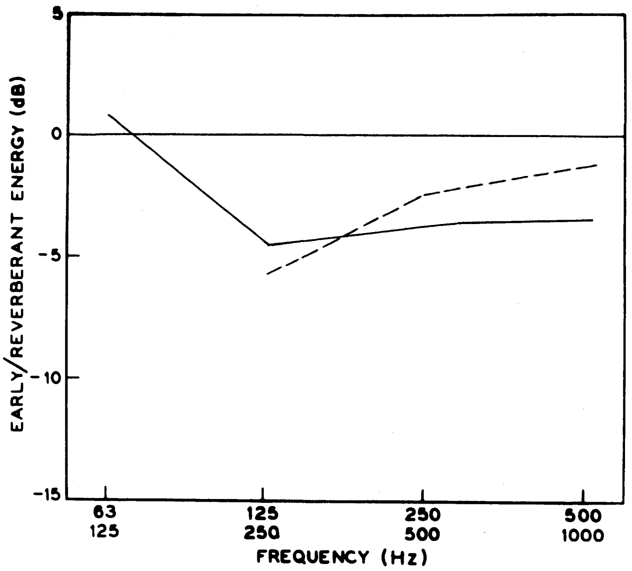



Fig. 12. Ratio of early to reverberant sound energy for the New York Philharmonic Hall before ( - - - - ) and after ( —— ) raising the reflectors to a position just below the ceiling. From SCHROEDER & al. (1966), Journal of the Acoustical Society of America.

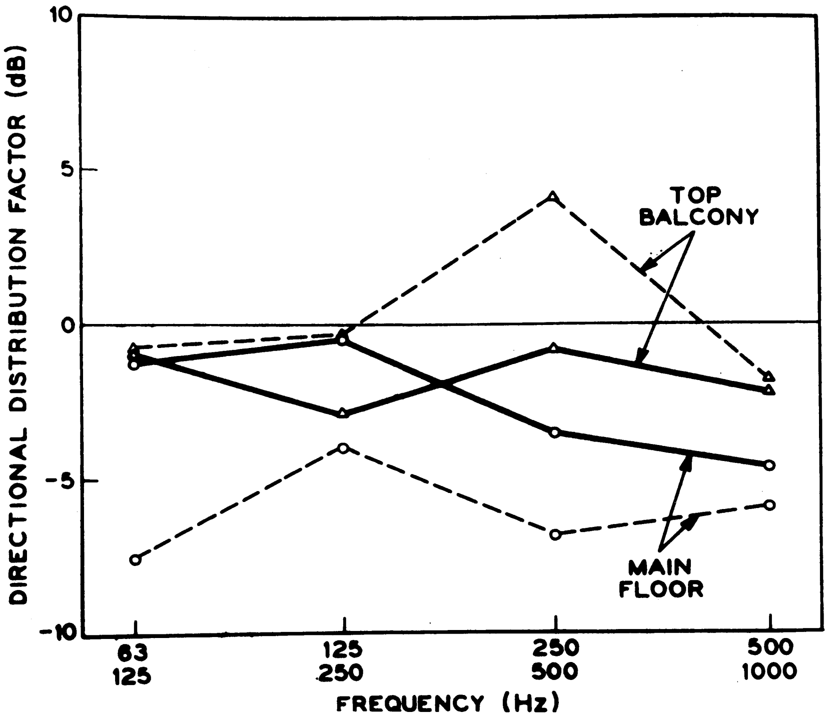



Fig. 13. Directional distribution factor for the New York Philharmonic Hall before ( - - - ) and after ( —— ) raising the reflectors to a position just below the ceiling. From SCHROEDER & al. (1966), Journal of the Acoustical Society of America.

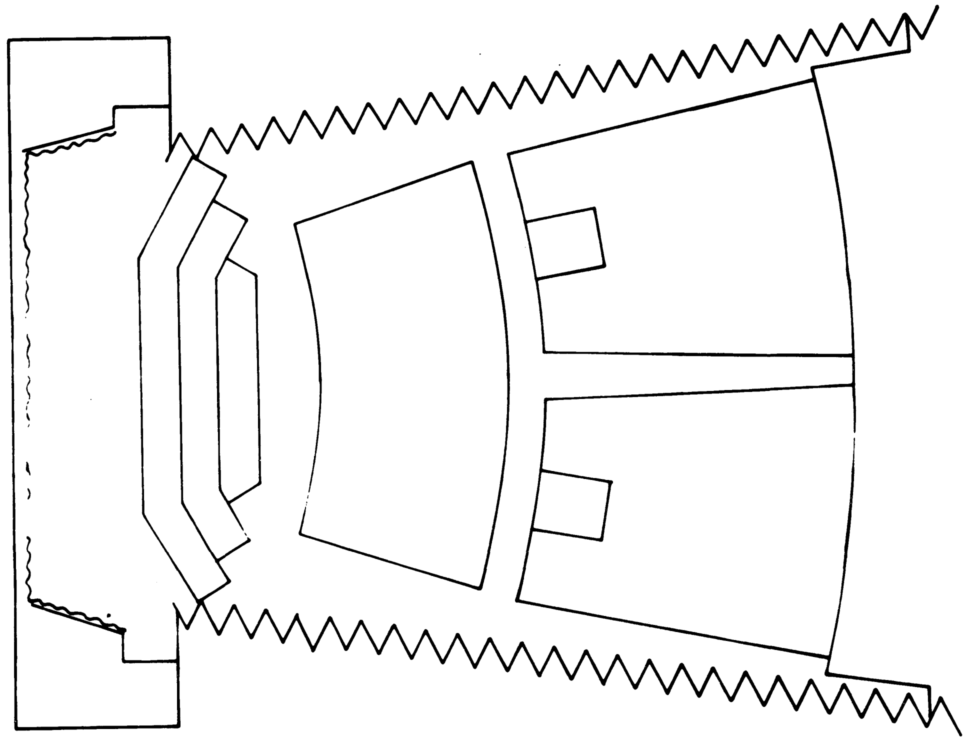


Fig. 14. University Hall, Reykjavik, plan. From JORDAN (1970), Journal of the Acoustical Society of America.

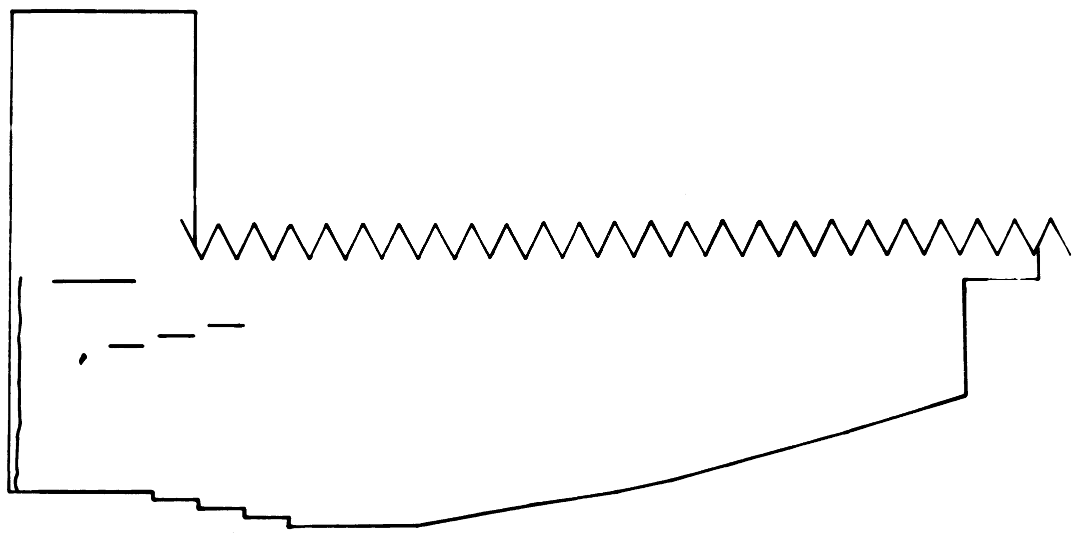


Fig. 15. Longitudinal section of the Reykjavik University Hall. From JORDAN (1970), Journal of the Acoustical Society of America.

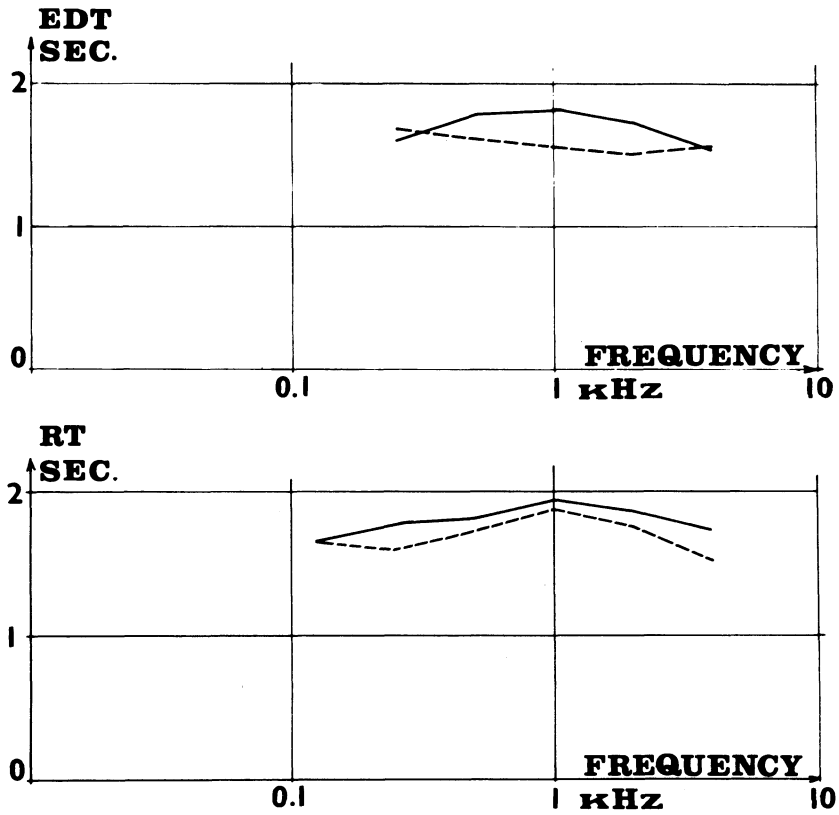



Fig. 16. EDT and RT values as functions of frequency with (solid line) and without (dashed line) enclosed stage for the Reykjavik University Hall. From JORDAN (1970), Journal of the Acoustical Society of America.

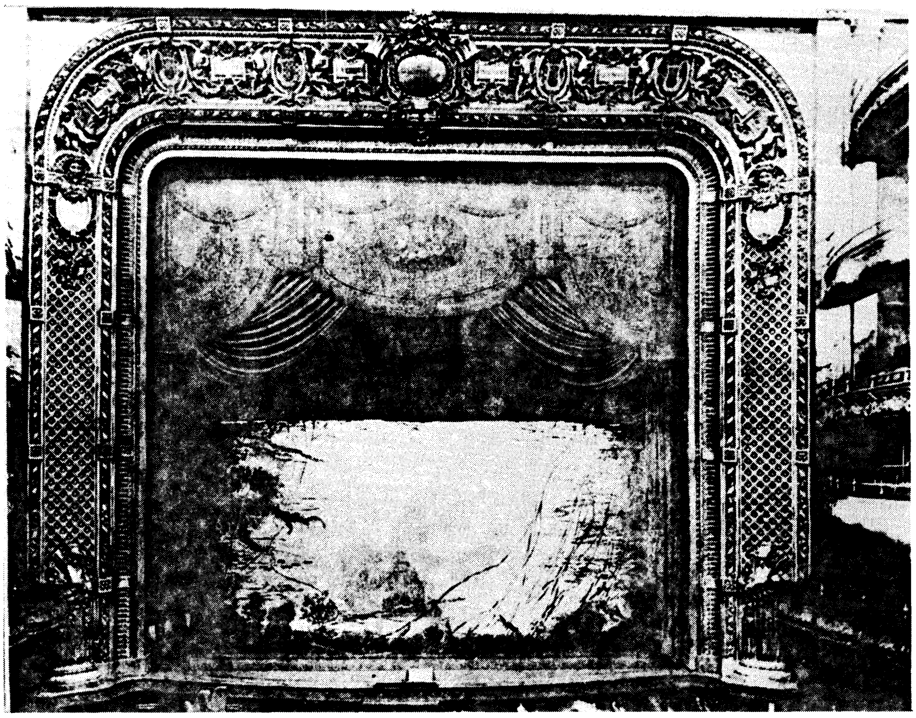


Fig. 17. Old Metropolitan Opera, New York, view towards the stage. From BERANEK (1962), John Wiley & Sons.

knowledge of applied materials, possibly supplemented by design of sound rays in sections of the hall). It is considered useful to investigate the general and detailed shape of the hall by constructing physical models in rather large scale (between 1:8 and 1:20), and to test these models preferably by the integrated pulse method.

In a number of cases the author has applied models in 1:10 scale and has specifically been interested in measuring EDT values for several positions in a hall. Apart from calculating inversion index from such measurements it is possible to compare EDT values with RT values and to evaluate the results, on the assumption that EDT values should not be shorter than RT values and also should not vary too much with frequency or location. One of the projects, where this method was applied extensively, was the new Metropolitan Opera of New York. (JORDAN, 1970)

In the early design period it was decided to shape the proscenium as a specific transition zone (from stage to auditorium) by applying large vertical splays (joined by a horizontal member) to both sides of the proscenium. This zone was thought to create many early reflections, especially lateral reflections, which would be valuable to the acoustics (both for the singers and for the orchestra).

This framing of the proscenium is actually an old traditional design (examples: San Carlos in Naples and Teatro Colon in Buenos Ayres). It is also applied in the Paris Opera but not in the old Metropolitan Opera of New York (Fig. 17). A photo of the New Metropolitan Opera is shown in Fig. 18 and a line drawing in Fig. 19. By the testing in the model it was found that EDT values in several locations and at different frequencies were very close to RT values. By a test (with audience) in the finished theater it was checked that this correspondence between EDT and RT values also existed in reality. Fig. 20 shows the (octave) values of EDT and RT in the finished hall.

A comparison of the RT vs. frequency characteristic of the Old and the New Metropolitan Opera shows a marked difference between the values, especially at high frequencies. No doubt, the new hall has a more balanced frequency characteristic (Fig. 21).

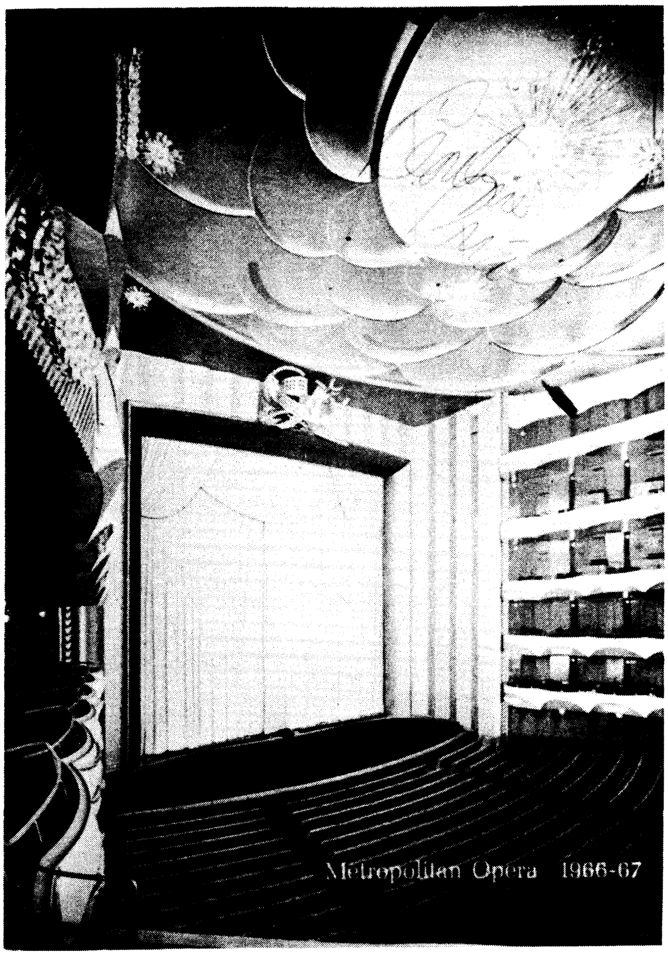



Fig. 18. New Metropolitan Opera, New York, view towards the stage.

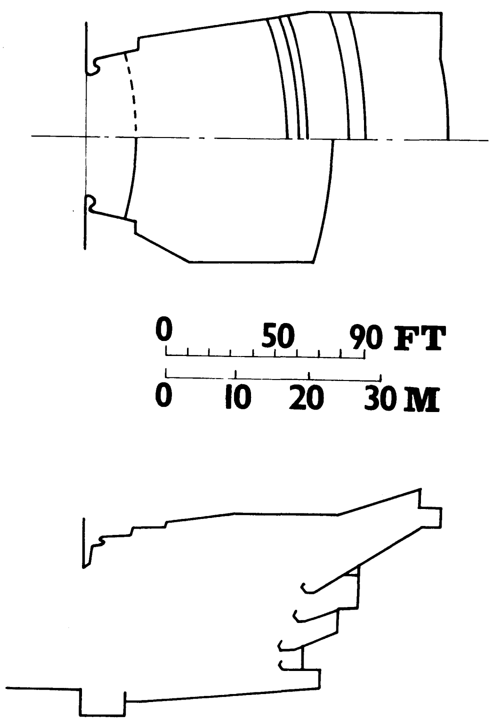



Fig. 19. Plan and longitudinal section of the New Metropolitan Opera, New York. From JORDAN (1970), Journal of the Acoustical Society of America.

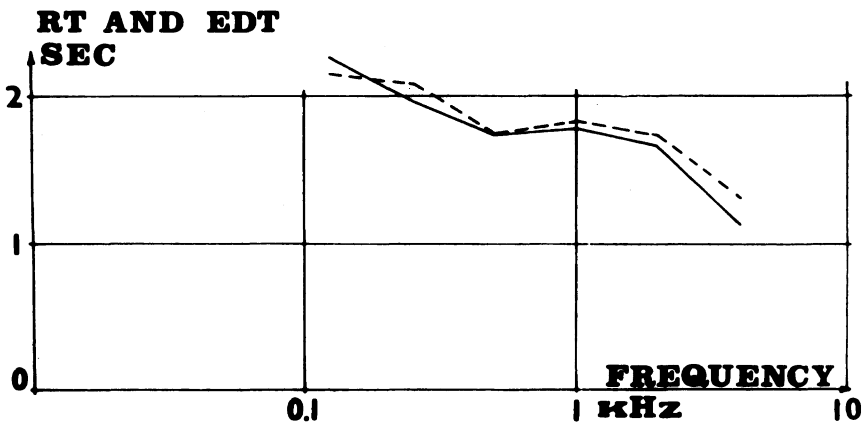



Fig. 20. RT (solid line) and EDT (dashed line) values as functions of frequency for the New Metropolitan Opera, New York. From JORDAN (1970), Journal of the Acoustical Society of America.

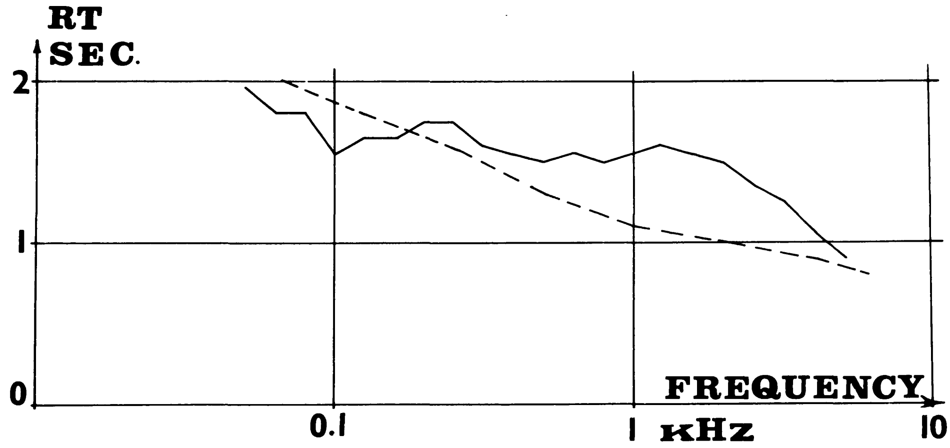



Fig. 21. Comparison of RT values as functions of frequency for the Old (dashed line) and the New (solid line) Metropolitan Opera in New York. From JORDAN (1970), Journal of the Acoustical Society of America.

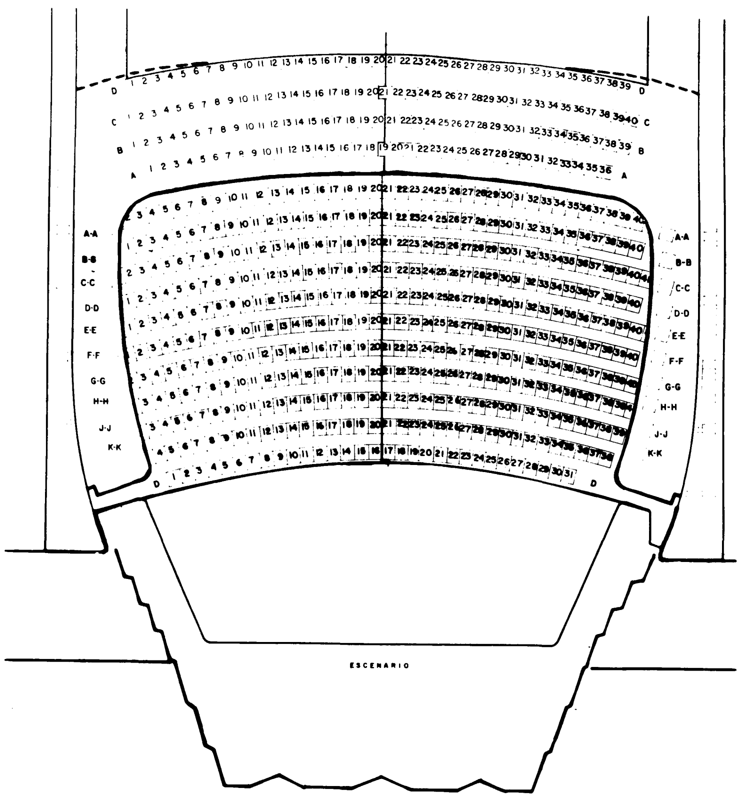



Fig. 22. National Theater, Nicaragua, plan.

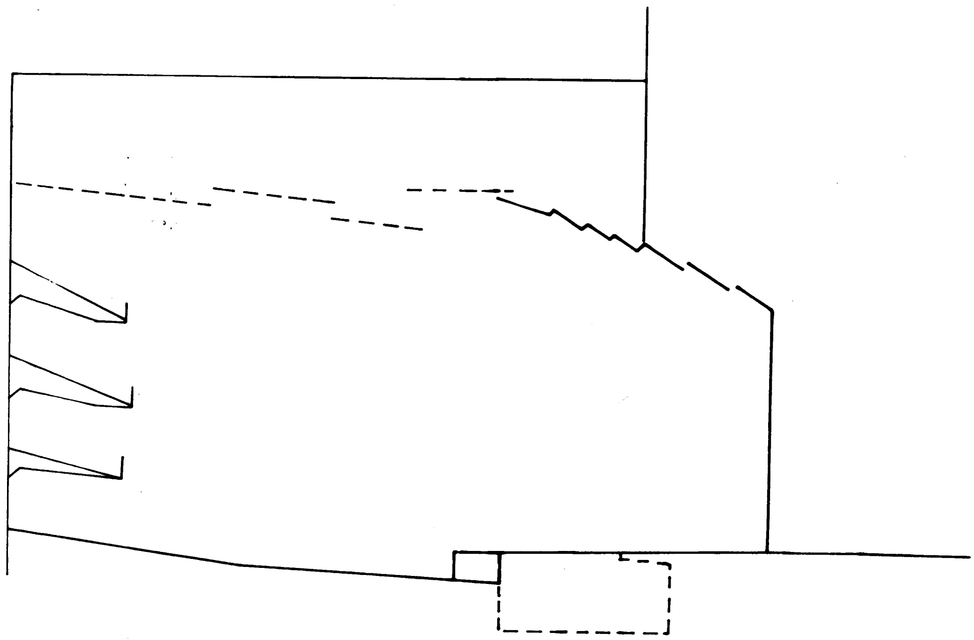


Fig. 23. Longitudinal section of the National Theater, Nicaragua.

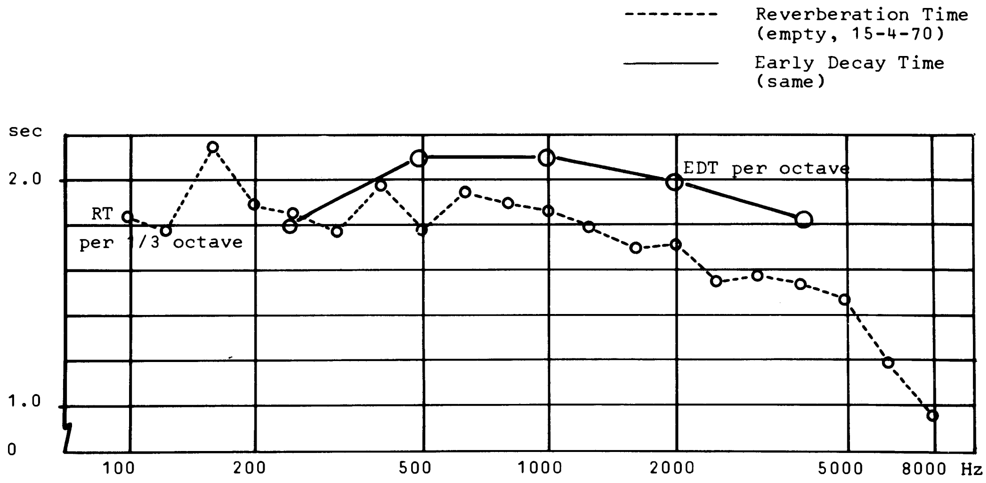



Fig. 24. EDT and RT values as functions of frequency of the National Theater, Nicaragua.

National Theater, Nicaragua (1969)

This theater, Fig. 22, 23, is mentioned as an example of a combination of a proscenium theater with a concert hall. The shape of the main auditorium is rectangular supplemented with an "orchestra shell" applied at concerts. The ceiling is designed specifically to provide early, reverberant energy by a sound transparent metal ceiling and an empty space above, partly subdivided by a system of vertical, suspended sound reflecting panels.

Measurements of EDT show, in a wide frequency range, values somewhat above RT values (Fig. 24). In different locations, throughout the auditorium, measured EDT values showed a maximum variation of only 12 % and also measurements of peak values of a short pulse (pistol shot) showed very small variations (only 4 %), when measured with a sound level meter (peak reading, Table 2).

Table 2

National Theater, Nicaragua

Values of EDT (sec) and impulse response (dB)

| Location | EDT (sec) | Impulse response (dB) |
|---|---|---|
| IV S - III S | 1.79 | 104 |
| I S - II S | 2.04 | 103.2 |
| VI S - V S | 2.02 | 105.7 |
| VI S - VII S | 2.01 | 104.5 |
| VI S - I S | 2.11 | 112.7 |
| Average on stage | 2.00 | |
| I S - I A | 2.14 | 103.8 |
| I S - II A | 1.90 | 103.8 |
| I S - III A | 1.88 | 101.5 |
| I S - IV A | 2.06 | 103.2 |
| I S - VI A | 2.02 | 105.2 |
| I S - VII A | 2.00 | 103.4 |
| I S - VIIIA | 2.04 | 103.4 |
| I S - IX A | 2.11 | 103.7 |
| Average in auditorium | 2.02 | |

Values of EDT are measured for 2 kHz octave band.

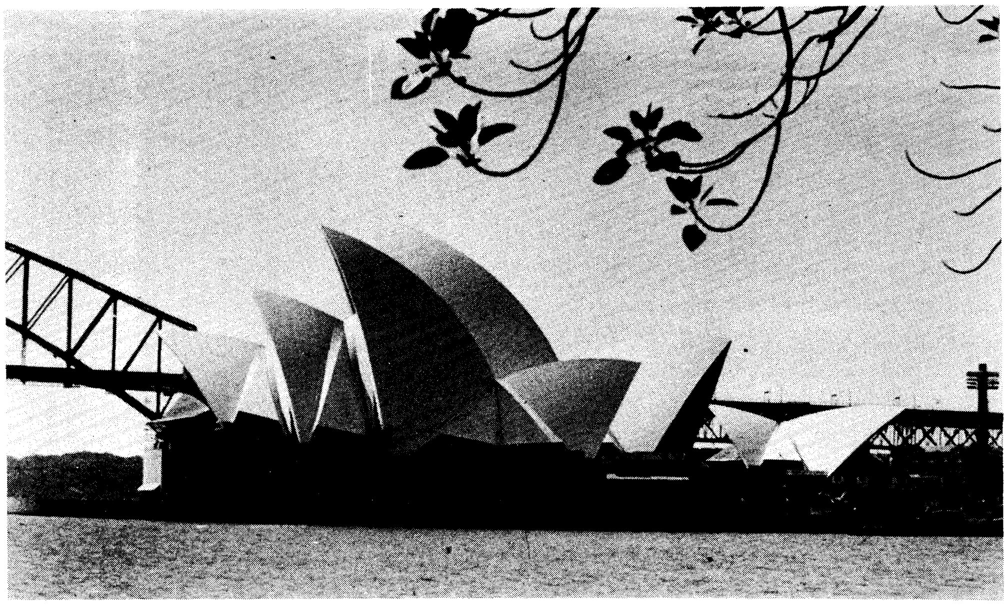


Fig. 25. Distant view of the Sydney Opera House.

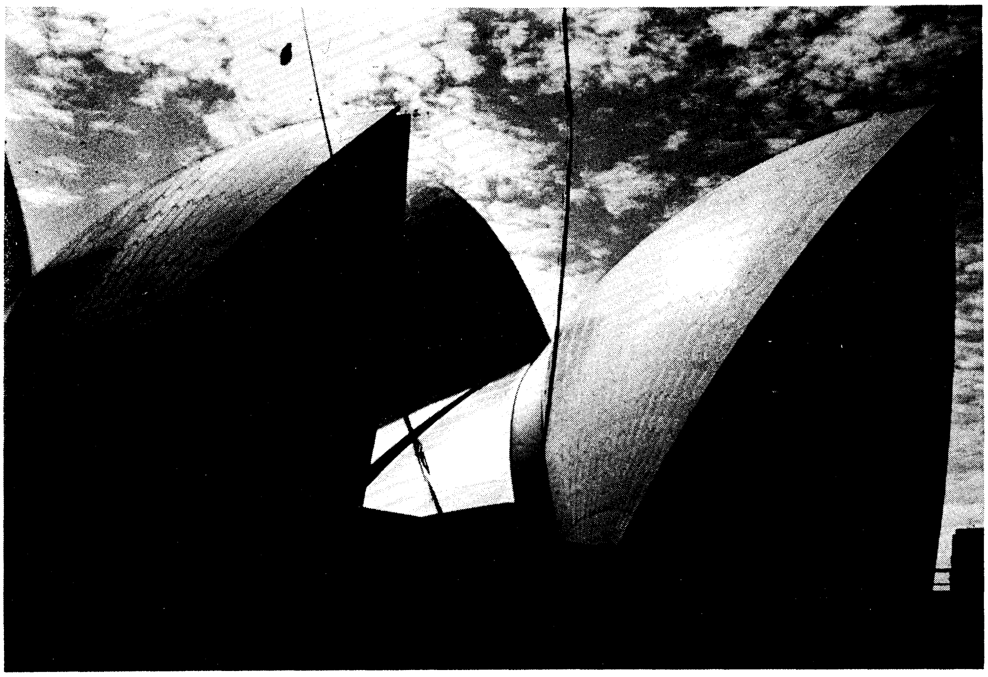


Fig. 26. Close view of the shells of the Sydney Opera House.

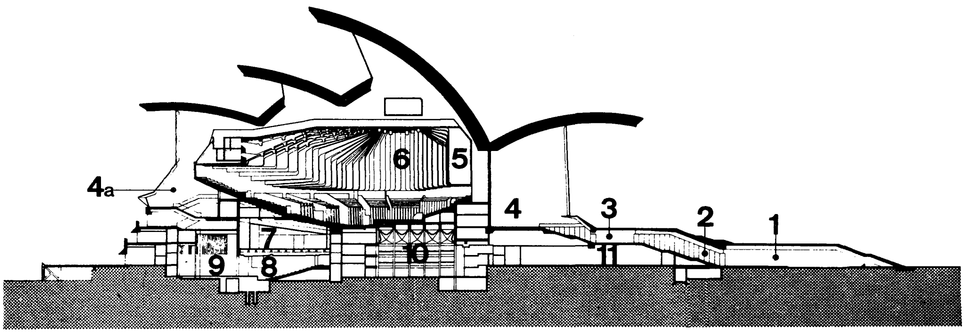



Fig. 27. Longitudinal section of the Concert Hall in the Sydney Opera House. 1: Car concourse, 2: staircase to foyer, 3: foyer, box office, cloak rooms, 4: Concert Hall foyer, 4a: promenade lounge, 5: organ loft, 6: Concert Hall, 7: rehearsal room, 8: drama theater, 9: drama theater stage, 10: rehearsal and recording hall, 11: cinema and exhibition hall foyer.

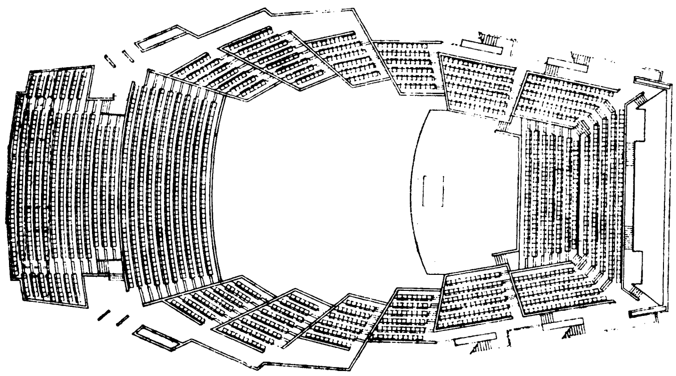


Fig. 28. Plan of the Concert Hall in the Sydney Opera House.

The Sydney Opera House (1973)

If you take a look at the exterior of this building complex whether from a distance (Fig. 25) or at a close position (Fig. 26) it is difficult to imagine that the outer shapes in anyway should have any intimate relationship to the interior shaping of the auditoria and to their acoustical functions.

Actually the original conceptions of the architect were based on sculptural and estethic phantasies, no doubt of superior quality but creating a number of practical problems which, however all were solved, ultimately. All the interiors had to be completely redesigned according to a revised programme for the applications of the Opera House. The largest hall (under the larger shell system) was redesigned to fill the needs of Sydney for a large concert hall, (Fig. 27, 28), whereas the next largest hall (under the smaller shells) was reshaped to be an Opera Theater. The complex also contains an Orchestra Studio, a Drama Theatre and a Chamber Music Room, but only the largest halls will be further described here.(JORDAN, 1973)

One of the problems of the Concert Hall was due to the fact that cross sections out of necessity become narrow with increasing height (due to the outer shells). Originally, the ceiling shape (in longitudinal sections), was thought of as consisting of large convex spans. The sidewalls were divided into steps so that in cross section both side walls and ceiling participated in these steps (Fig. 29). In planning the outlay of the seating it was decided to bring the stage somewhat forward and to have one part of the audience sitting behind the orchestra (Fig. 28), thereby reducing the average distance from the orchestra to the seats.

This design was built as a scale model in 1:10 and was tested acoustically by the pulse method (measurements of RT and EDT).

The results of the model test were not entirely satisfactory especially the values of EDT in the stalls were too low. It was decided to change the design, retaining the outlay in plan, but reshaping the side walls by placing the side boxes so that the side walls could have large vertical surfaces. The idea was to make side reflections more dominant and ceiling reflections less dominant.

The measurements of EDT showed definite improvements and also showed less variation from one location to another.

In the final design the ceiling (with increased ceiling height) had a large number of boxshaped diffusors which were also acting as airinlets. The circular reflectors over the stage (convex on both surfaces) were changed into toroidal shapes (Fig. 30, 31).

The pulse method used for the model investigations was also applied in the final testing of the finished concert hall (with audience). A comparison of the results for the model and the hall is shown in Table 3. Inversion Index, EDT values (in percent of RT values) and maximum variation of EDT values are shown. The optimal height of the reflectors is at a medium position 10 m above stage floor.

The Concert Hall, (Fig. 32, 33), has had a remarkably favorable reception amongst musicians, singers and audience.

The next largest hall, the Opera Theatre, has also been tested as a model in several versions. The first design had a relatively flat, low ceiling (Fig. 34). The results of the measurements are shown in Table 4 after the same principle as for the Concert Hall. In this case there are two different definitions of inversion index one applying to the stage, the other to the pit. The final design (Fig. 35) showed the best balance between stage and pit (the smallest variation in I.I.). Like the Concert Hall the Opera Theatre has had a good reception. However, the limited space in the pit and the prominent overhang created some problems for the musicians (too heavy sound, compensated by adding extra absorption in the pit).

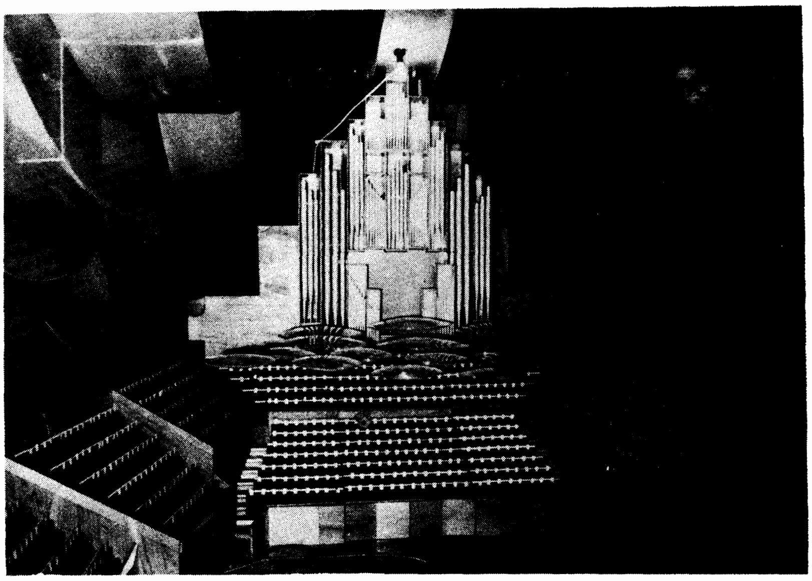


Fig. 29. 1:10 model of the first design of the Concert Hall in the Sydney Opera House, view towards the stage.

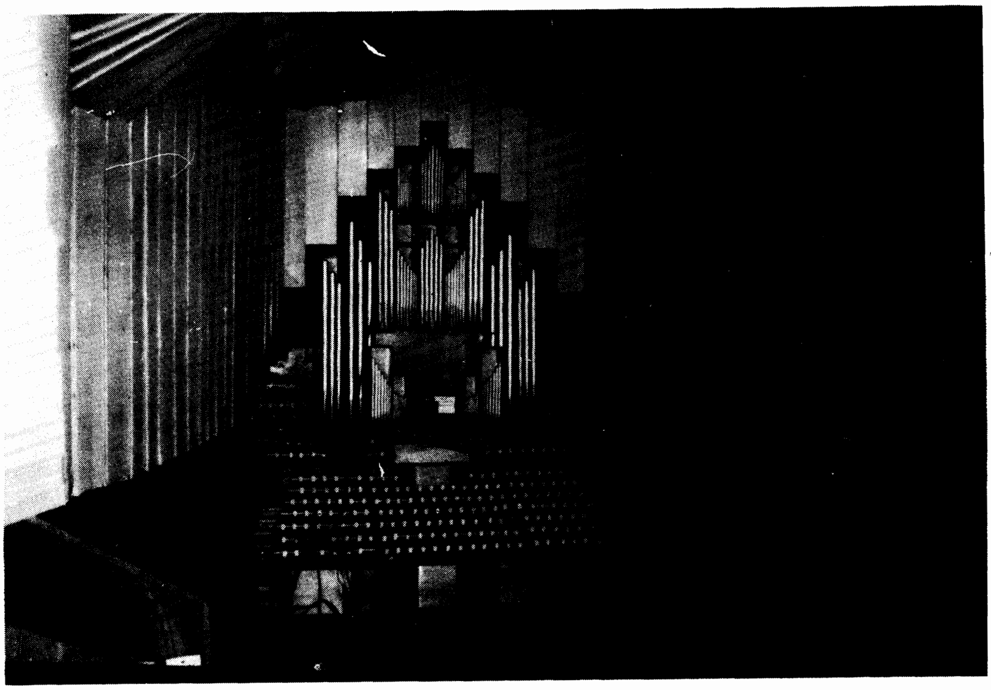


Fig. 30. 1:10 model of the final design of the Concert Hall in the Sydney Opera House, view towards the stage.

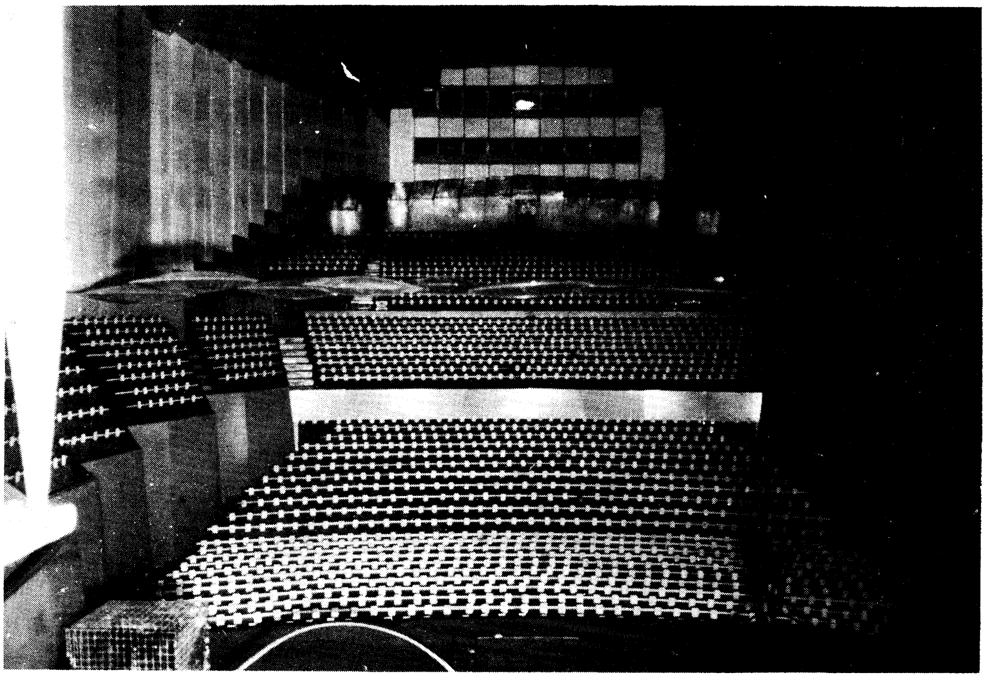


Fig. 31. Same as Fig. 30, view towards the audience.

Fig. 32. View towards the ceiling of the Concert Hall in the Sydney Opera House.

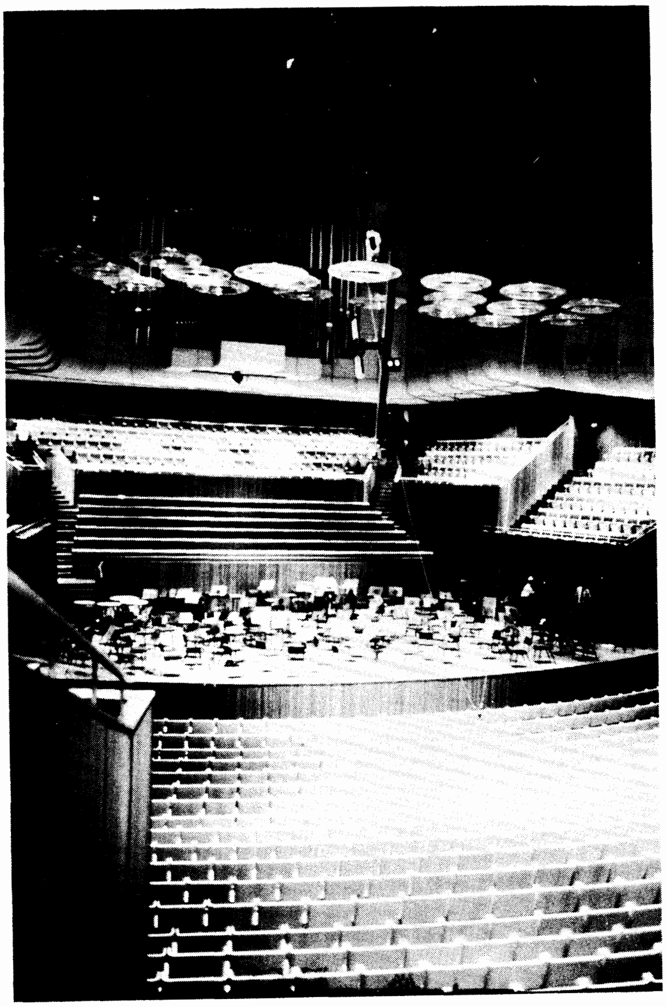


Fig. 33. Same as Fig. 32, view towards the stage.

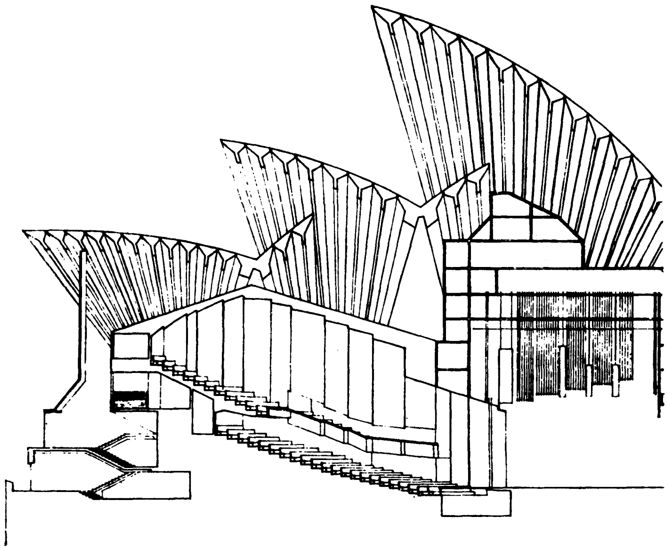


Fig. 34. Longitudinal section, first design of the Opera Theater in the Sydney Opera House.

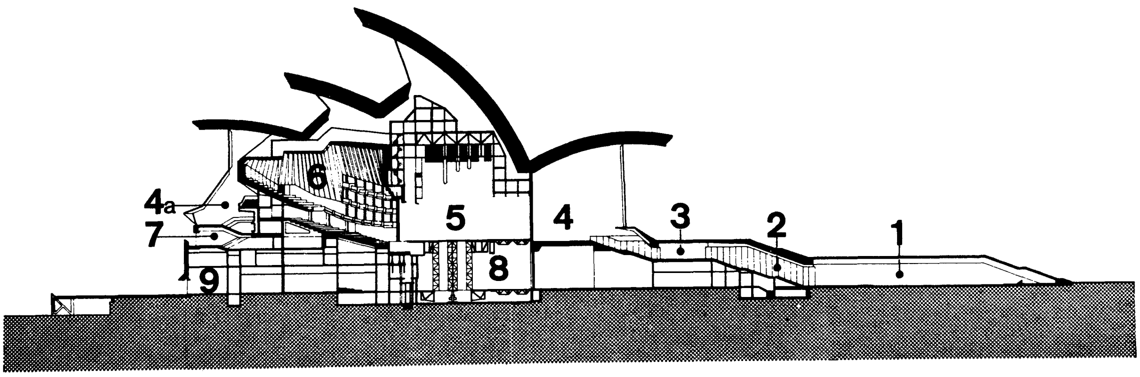



Fig. 35. Same as in Fig. 34, final design. 1 - 4a, see legend to Fig. 27, 5: Opera Theater stage, 6: Opera Theater, 7: Opera Theater lounge, 8: below-stage area, 9: broadwalk restaurant.

Table 3

Sydney Opera House

The Concert Hall

| Criterion and conditions | Model (2-16 kHz) | Hall (250-2000 Hz) empty | cap.audience |
|---|---|---|---|
| Inversion index | | | |
| (Ideal: I.I. $\geqq$ 1.0) | | | |
| Reflectors at soffit level | 1.11 | 1.14 | 0.96 |
| Reflectors just below crown | 1.05 | 1.10 | 1.06 |
| Reflectors 34' above stage | - | - | 1.05 (NB) |
| Percentage "good" EDT values | | | |
| (Ideal: = 100 %) | | | |
| Reflectors at soffit level | 94 | 96 | 85 |
| Reflectors just below crown | 83 | 100 | 85 |
| Reflectors 34' above stage | - | - | 94 (NB) |
| Percentage overall variation of EDT values | | | |
| (Ideal: = 0 %) | | | |
| Reflectors at soffit level | 20 | 13 | 17.5 |
| Reflectors just below crown | 21 | 13.5 | 17.5 |
| Reflectors 34'above stage | - | - | 10 (NB) |

NB: means frequency range only 500-2000 hz

Table 4

Sydney Opera House

The Opera Theatre

| | Model | | Hall | | | |
|---|---|---|---|---|---|---|
| | | | empty | | cap.audience | |
| Source at: | Pit | Stage | Pit | Stage | Pit | Stage |
| Criterion: | | | | | | |
| Inversion Index (Ideal $\geqq$ 1.0) | 1.22 | 1.13 | 1.52 | 1.09 | 1.22 | 1.10 |
| Percentage "good" EDT values (Ideal: 100 %) | 100 | 100 | 100 | 100 | 100 | 96 |
| Percentage overall variation (Ideal: 0. %) | 38 | 30 | 21 | 15 | 9 | 19 |

The Sydney Opera House was completely finished and inaugurated by the fall of 1973. Some critical remarks on the limited stage space of the Opera Theatre does not appear to be justified, all facts considered. The problem of staging "Grand Opera" requiring unusual stage space (like "Aida") in the Opera House, was recently solved simply by using the Concert Hall (having the singers on stage and the orchestra in front of the stage and applying a permanent stage picture). The experiment apparently was a success, also acoustically.

## Criteria and practical design

A recent period in the history of room acoustics has been reviewed in this article partly by exposing the development of acoustical and musical criteria, partly by relating practical experience and tests in a number of finished halls. It is a period which has been marked by many investigations into the problem of what actually defines the "acoustics" of large halls and by many proposals of objective criteria to be correlated with subjective acoustical qualities.

In this decade several groups, by systematical research, are trying to establish this correlation between a limited number of acoustical qualities and some corresponding objective criteria.

During this process which may not be completed in the seventies, the problem of varying musical taste is becoming apparent. Some music lovers may prefer clarity and others reverberance. It is also evident that variation of qualities may be larger within one room than from one room to another.

It remains to be seen whether in practice concert halls exist which have such qualities that "local variations" become less important. In the opinion of the author such halls will be characterized by certain basic properties:

(1) Good balance between stage area and audience area, with rapid approach to statistical conditions, especially in the stage area.

(2) Optimum of both clarity and reverberance

(3) Great volume of the sound

(4) Tonal balance

Maybe divergences in the judgement of quality become less apparent in such halls and in this connection it may be enlightening to take another look at Fig. 4. No doubt the hall, HH (Musikhalle Hamburg) is approaching this condition. Another look at Fig. 3 show furthermore that this hall has the classical, rectangular shape, the least variation in EDT values and that inversion index (calculated from TR = rise time) is the highest of the six halls. One example does not prove the case but all the same it gives a useful indication in tracing a more comprehensive relationship.

References:

ATAL, B.S., SCHROEDER, M.R., and SESSLER, G.M. (1965): "Subjective Reverberation Time and its Relation to Sound Decay", paper G32 in Proc. of the 5th Int.Congr. on Acoustics, Liège, Vol. 1b.

BERANEK, L.L. (1962): Music, Architecture and Acoustics. John Wiley & Sons, Inc., New York.

BERANEK, L.L. and SCHULTZ, T.J. (1965): "Some Recent Experiences in the Design and Testing of Concert Halls with Suspended Panel Arrays", Acustica, 15, pp. 307-316.

EYSHOLDT, U., GOTTLOB, D., SIEBRASSE, K.F., and SCHROEDER, M.R. (1975): "Räumlichkeit und Halligkeit", Deutsche Arbeitsgemeinschaft für Akustik, pp. 471.

GOTTLOB, D. (1973): "Vergleich objektiver akustischer Parameter mit Ergebnissen subjektiver Untersuchungen an Konzertsälen", Diss. Göttingen.

GOTTLOB, D., SIEBRASSE, K.F., and SCHROEDER, M.R. (1975): "Neuere Ergebnisse zur Akustik von Konzertsälen", Deutsche Arbeitsgemeinschaft für Akustik, pp. 467.

JORDAN, V.L. (1959): "The Building-up Process of Sound Pulses in a Room and its Relation to Concert Hall Quality", pp. 922-925 in Proc. of the Third Int. Congr. on Acoustics, Stuttgart (ed. L. Cremer), Elsevier Publ. Co., Amsterdam 1961.

JORDAN, V.L. (1968): "Einige Bemerkungen über Anhall und Anfangsnachhall in Musikräumen", J. of Appl. Acoustics, 1, pp. 29-36.

JORDAN, V.L. (1970): "Acoustical Criteria for Auditoriums", J. Acoust. Soc.Am., 47, pp. 408-412.

JORDAN, V.L. (1973): "Acoustical Design Considerations of the Sydney Opera House", J. and Proc. of the Royal Society of New South Wales, 106, parts 1 & 2, pp. 33-53.

MARSHALL, A.H. (1967): "A Note on the Importance of Room-Cross Sections in Concert Halls", J. Sound and Vib., 5, pp. 100-112.

PLENGE, G., LEHMANN, P., WETTSCHURECK, R., and WILKENS, H.W. (1975): "New Methods in Architectural Investigations to Evaluate the Acoustic Qualities of Concert Halls", J.Acoust.Soc.Am., 57, pp. 1292-1299.

REICHARDT, W. and SCHMIDT, W. (1966): "Die hörbaren Stufen des Raumeindrucks bei Musik", Acustica, 17, pp. 175-179.

REICHARDT, W., ABDEL ALIM, O., and SCHMIDT, W. (1975): "Definition und Messgrundlage eines objektiven Masses zur Ermittlung der Grenze zwischen brauchbarer und unbrauchbarer Durchsichtigkeit bei Musikdarbietung", Acustica, 32, pp. 126-137.

SCHROEDER, M.R. (1965): "New Method of Measuring Reverberation Time", J.Acoust.Soc.Am., 37, pp. 409-412.

SCHROEDER, M.R., ATAL, B.S., SESSLER, G.M., and WEST, J.E. (1966): "Acoustical Measurements in Philharmonic Hall", J.Acoust.Soc.Am., 40, pp. 434-440.

SCHROEDER, M.R., GOTTLOB, D., and SIEBRASSE, K.F. (1974): "Comparative Study of European Concert Halls", J.Acoust.Soc.Am., 56, pp. 1195-1201.

SCHULTZ, T.J. (1965): "Acoustics of the Concert Hall", IEEE Spectrum, June, pp. 56-67.

THIELE, R. (1953): "Richtungsverteilung und Zeitfolge der Schallrückwürfe in Räumen", Acustica, 3, pp. 291-302.

de VILLIERS KEET, W. (1968): "The Influence of Early Lateral Reflections on the Spatial Impression", paper E-2-4 in Reports of the 6th Int. Congr. on Acoustics, Tokyo, Vol. III.

# ON SOUND RADIATION OF MUSICAL INSTRUMENTS

by

Erik V. Jansson, Center for Speech Communication Research and Musical Acoustics, Royal Institute of Technology (KTH), Stockholm, Sweden.

## Introduction

The sounding music is made up of complex acoustical signals. Furthermore it is radiated from the musical instruments in a complex way. Therefore no simple way exists to describe the acoustical signal of the sounding music in detail or how it is radiated. But even fairly coarse measures can be meaningful and of great help. Therefore such measures are presented in this report. More detailed information can be found in the references listed.

First in this report data are presented on dynamic and frequency ranges both of the hearing and of music played on single instruments. Secondly, time measures relevant for the hearing and for the starting transients of played notes are given. Thirdly directional characteristics of some musical instruments are presented. The report ends with a demonstration of how a violin sounds from different directions and presents corresponding long-time-average-spectra.

## Dynamic and frequency ranges

The human ear sets the limits for the sounds that we can hear and enjoy, see Fig. 1. It sets a lower dynamic limit, the threshold of audibility (Hörschwelle) and an upper dynamic limit at which the sound elicits pain, the threshold of pain (Schwelle der Schmerzempfindung). Furthermore it sets a lower frequency limit approximately at 16 Hz and a higher frequency limit at approximately 16 000 Hz for the sounds we can hear. Within these four limits we find the sounds that we communicate with and listen to, as speech and music (Sprache und Musik). A more detailed description of the dynamic ranges are given in Fig. 2, in which equal loudness curves and corresponding phone measures are plotted. It can be seen that the different families of instruments all have approximately the same magnitude of dynamic ranges, 30 dB, but the absolute values of

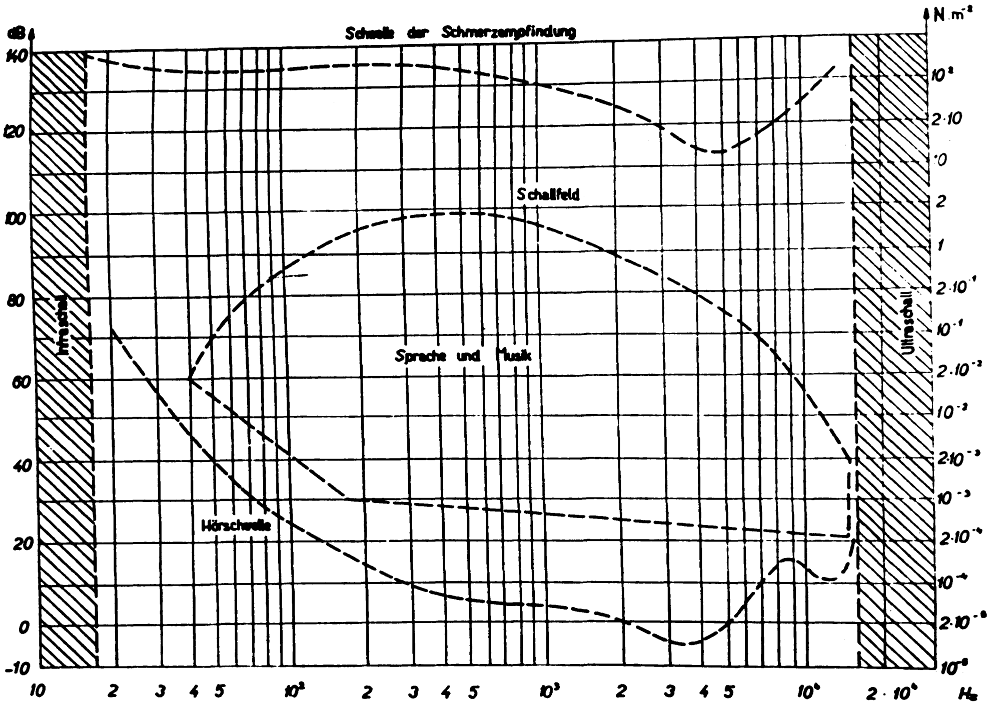



Fig. 1. Dynamic and frequency ranges of the hearing. From BURGHAUSER & SPELDA (1971). Akustische Grundlagen des Orchestriercus, Gustav Bosse Verlag.

the dynamic limits are slightly different. Among the orchestral instruments the string instruments are the weakest, the woodwinds approximately 10 dB louder and the brass winds still 10 dB louder. The guitar sound is approximately 5 dB weaker than the orchestra strings. The total range of the single instruments is approximately 50 dB, which means that a dynamic range of 50 dB is minimum for true recordings of the different instruments. (BURGHAUSER & SPELDA, 1971)

The frequency ranges are set by the lowest fundamental frequencies and the partials of the highest frequencies encountered in playing. The lower limits are given in Fig. 3 for the trombone (B), the bassoon (W) and the double bass (S). The upper frequency limit is slightly less than 10 000 Hz and varies within 2000 Hz for the different instrumental groups. (BURGHAUSER & SPELDA, 1971; MEYER, 1972; JANSSON, 1973) These variations are large in absolute measures, but moderate as compared to the sensitivity of the hearing in this frequency range. Thus for the instruments listed it seems appropriate to use recording equipment with a frequency range of 40 to 10 000 Hz.

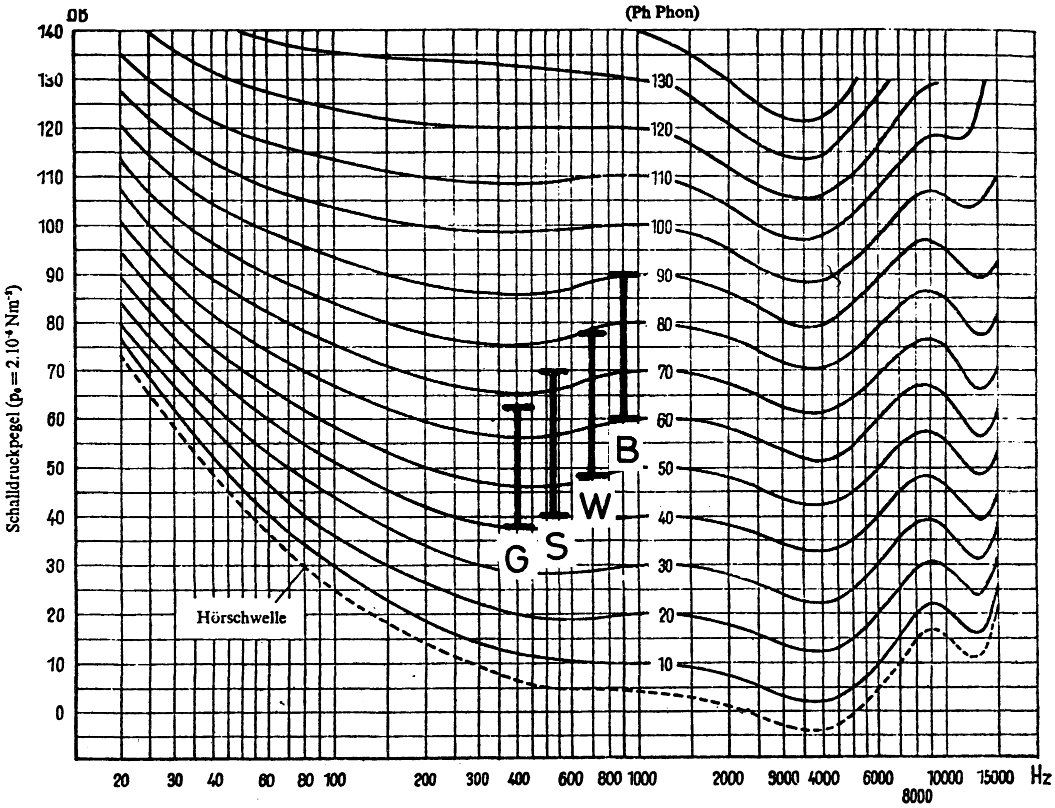



Fig. 2. Equal loudness contours with sketched dynamic ranges of the guïtar (G) and of the instrument families of the symphony orchestra; strings (S), woodwinds (W) and brass instruments (B). Adapted from BURGHAUSER & SPELDA (1971).

Transients

The musical sounds have also beginnings and ends. The beginnings are especially important and some measures of specific interest are plotted in Fig. 4. The durations and the closing transients may also be important. A sound must have a certain duration to give an accurate pitch sensation. For tone pulses shorter than 0.025 seconds the pitch is slightly lowered. (DOUGHTY & GARNER, 1970) A duration of 0.25 seconds gives optimum pitch detection, no better is obtained by longer duration. (ZWICKER & FELDTKELLER, 1967) A played tone must furthermore have a certain duration to give full loudness. This duration is given by the time constant for loudness, which is approximately 0.1 second. (ZWICKER & FELDTKELLER, 1967) The starting transients and the durations of played tones are within the time ranges presented, see Fig. 4. (MEYER, 1972; MELKA, 1970; JANSSON, 1973)

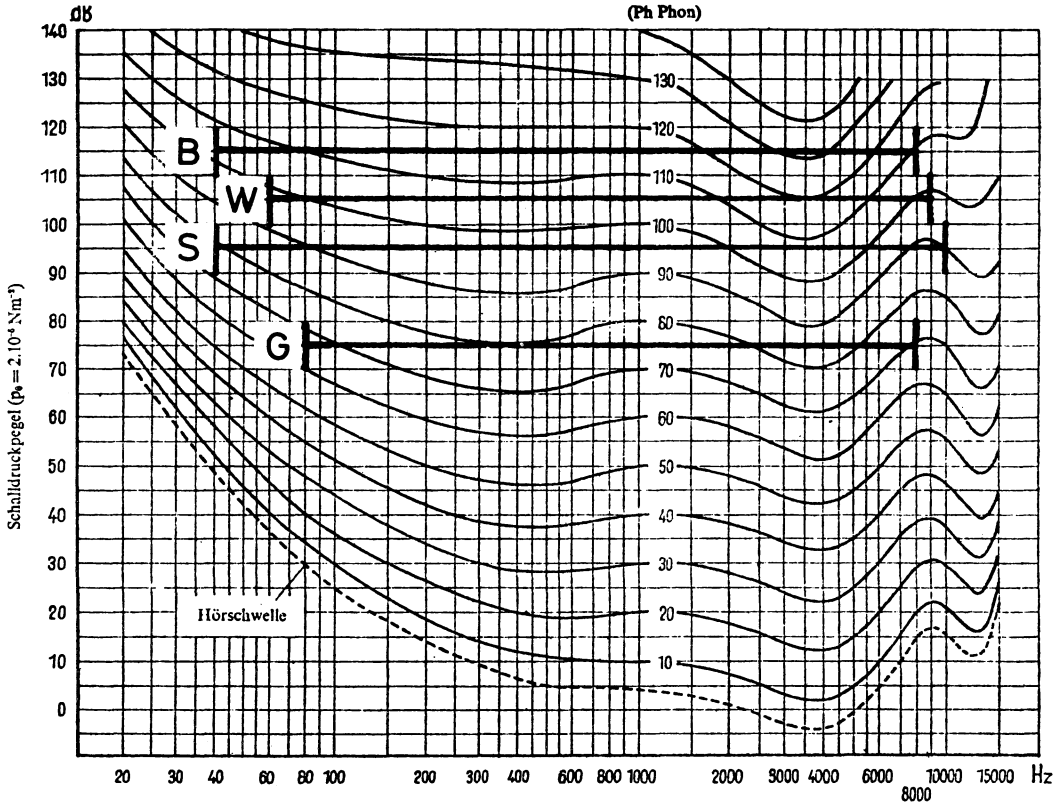



Fig. 3. Equal loudness contours with sketched frequency ranges of the guitar (G) and of the instrument families of the symphony orchestra; strings (S), woodwinds (W) and brass instruments (B). Adapted from BURGHAUSER & SPELDA (1971), MEYER (1972), and JANSSON (1973).

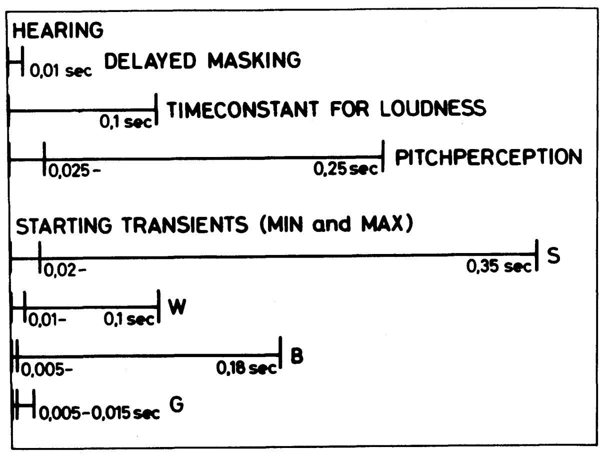



Fig. 4. Perceptually important durations, and time constants of starting transients of the guitar (G) and of the instrument families of the symphony orchestra; strings (S), woodwinds (W) and brass instruments (B).

The starting transients can be extended within the time ranges for hearing. This means that the starting transients are important for perception, as is well recognized. True recordings of starting transients need no extension of the frequency limits given earlier.

## Sound radiation

The musical tones have still another important relation to the musical instrument. Different instruments have different directional characteristics, thus giving different sound qualities in different positions of a room. Typical examples of main radiation directions are given in Figs. 5, 6 and 7. (MEYER, 1972) The trombone, Fig. 5, has a wide radiation lobe at low frequencies, the width of which decreases with increasing frequency, with only one exception. The widened lobe at 650 Hz is likely to stem from interaction between a longitudinal and a radial mode of the trombone bell. (BENADE & JANSSON, 1974) The same gross features, a wide radiation lobe at low frequencies and a narrow at high are generally true for all musical instruments, which is supported by the Figs. 6 and 7 for the oboe and the violin. At intermediate frequencies the side holes of the oboe radiate sound, thus giving the more complex radiation pattern. At high frequencies most energy radiates through the bell as a tube with side holes acts as an acoustical high pass filter. (BENADE, 1976.) The violin seems to radiate most sound energy perpendicular to the top plate at high frequencies, Fig. 7.

Let me finally demonstrate how a violin sounds from different directions and present the corresponding acoustical analysis. The long-time-average-spectra of a fulltone scale recorded from four different directions in an anechoic chamber are presented in Fig. 8. (JANSSON, 1975.) The corresponding first four tones of each scale are given as sound examples. For easy comparison the spectra and the played scales are presented in pairs in the order, 1) directions one and two - Fig. 8 a and sound example one a, 2) directions one and three - Fig. 8 b and sound example one b, 3) directions one and four - Fig. 8 c and sound example one c. All spectra display a peak at 5 Bark (500 Hz), followed by a pronounced dip, i.e. the violin is an approximately omnidirectional sound radiator in this frequency range in agreement with Fig. 7. The continuations of the curves of Fig. 8 display grossly the same contours, but there are considerable

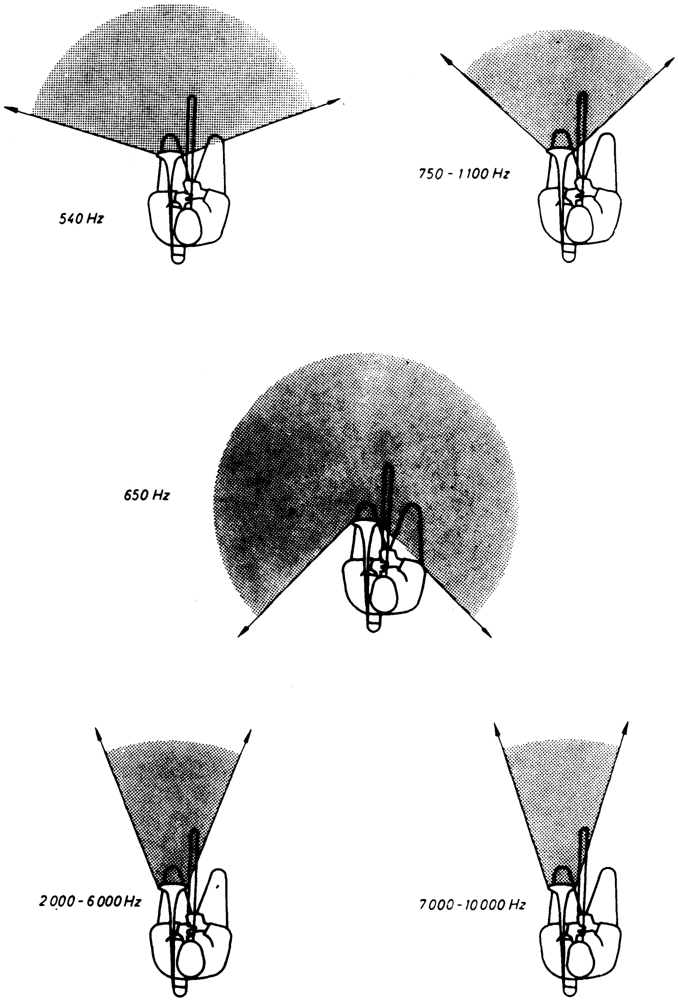



Fig. 5. Main radiation directions (0 ... -3 dB) of the trombone. From MEYER (1972): Akustik und musikalische Auffürungspraxis, Verlag Das Musikinstrument.

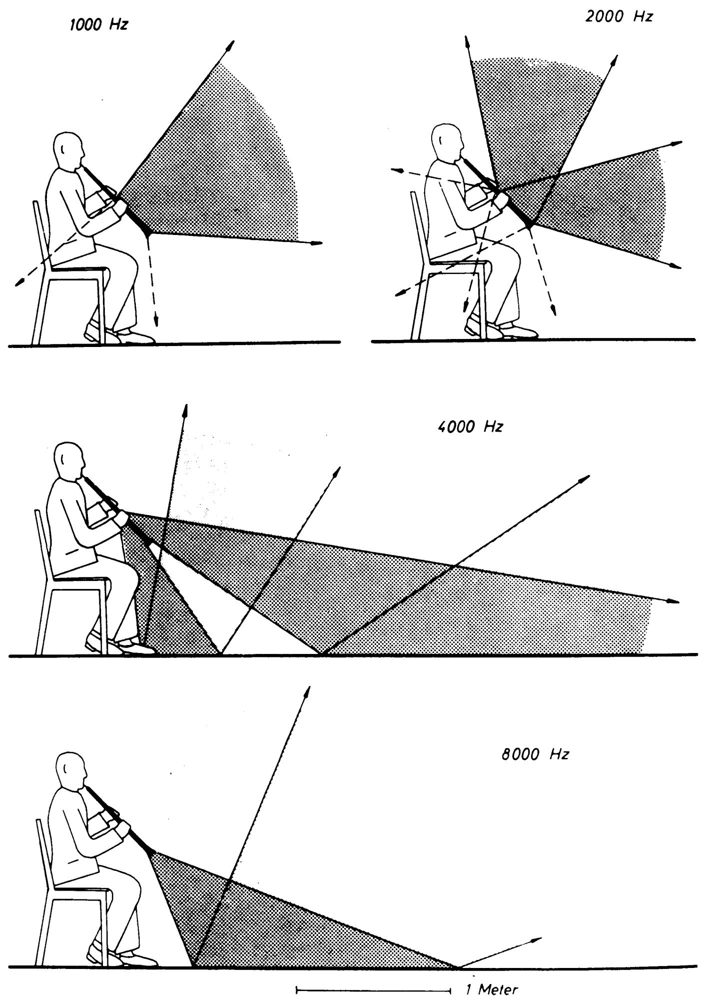



Fig. 6. Main radiation directions (0 ... -3 dB) of the oboe. From MEYER (1972): Akustik und musikalische Auffürungspraxis, Verlag Das Musikinstrument.

single discrepancies as can be expected from Fig. 7. The average level differences of the long-time-average-spectra for direction one, two, three and four compared to the average over five perpendicular directions are 0, +3, -2 and -7 dB respectively. These measures show that this violin is loudest from the direction of its neck (direction two) and weakest in the opposite direction (direction four). Furthermore the direction perpendicular to the top plate (direction one) gives the sound most closely corresponding to the average. The discrepancy measures given should be detectable as loudness differences as the just noticable difference in intensity is 1.5 dB. (FLANAGAN, 1965) Thus the main radiation directions corresponding to Figs. 7 and 8 are slightly different. This shows that no single position of a microphone can record all properties of a played tone or a musical instrument.

## Concluding remarks

Let me thus end this presentation by briefly summarizing my points presented above. Musical sounds and the radiation of musical sounds result in complex acoustical signals. The observable measures provided show that dynamic ranges, frequency ranges and time constants of starting transients of the musical sounds are within the working ranges of hearing. The radiation properties tend to be simple at low and high frequencies but complex for intermediate frequencies. The sound of the violin, at least, varies with direction in such a way that it is detectable by the ear even for played notes of low fundamental frequency.

## Acknowledgements

This work was supported by the Swedish Humanistic Research Council and the Swedish Natural Science Research Council.

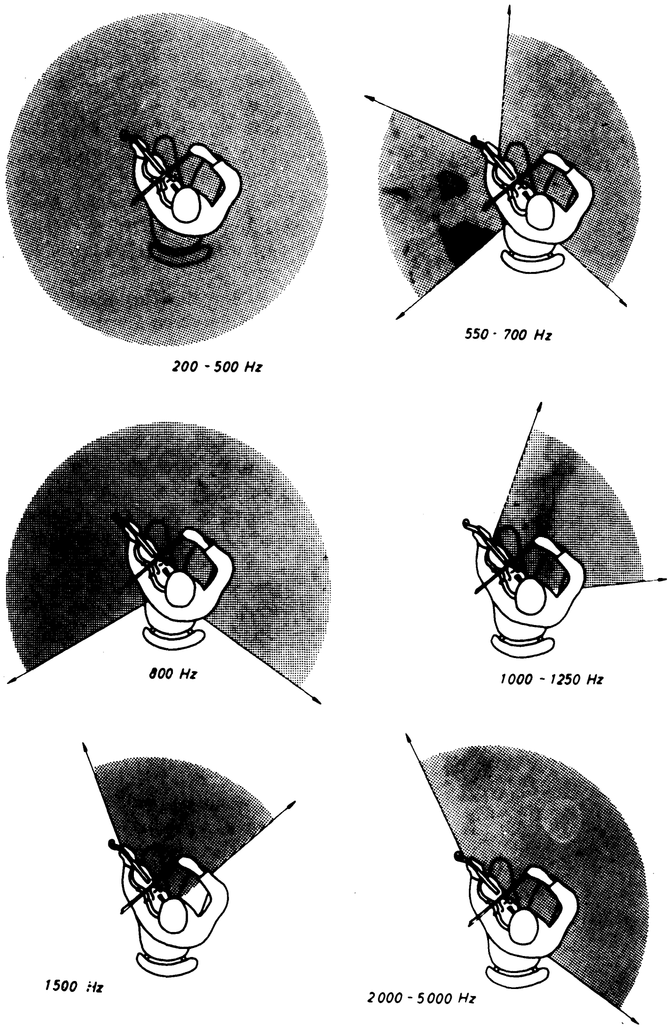



Fig. 7. Main radiation directions (0 ... -3 dB) of the violin. From MEYER (1972): Akustik und musikalische Auffürungs-praxis, Verlag Das Musikinstrument.

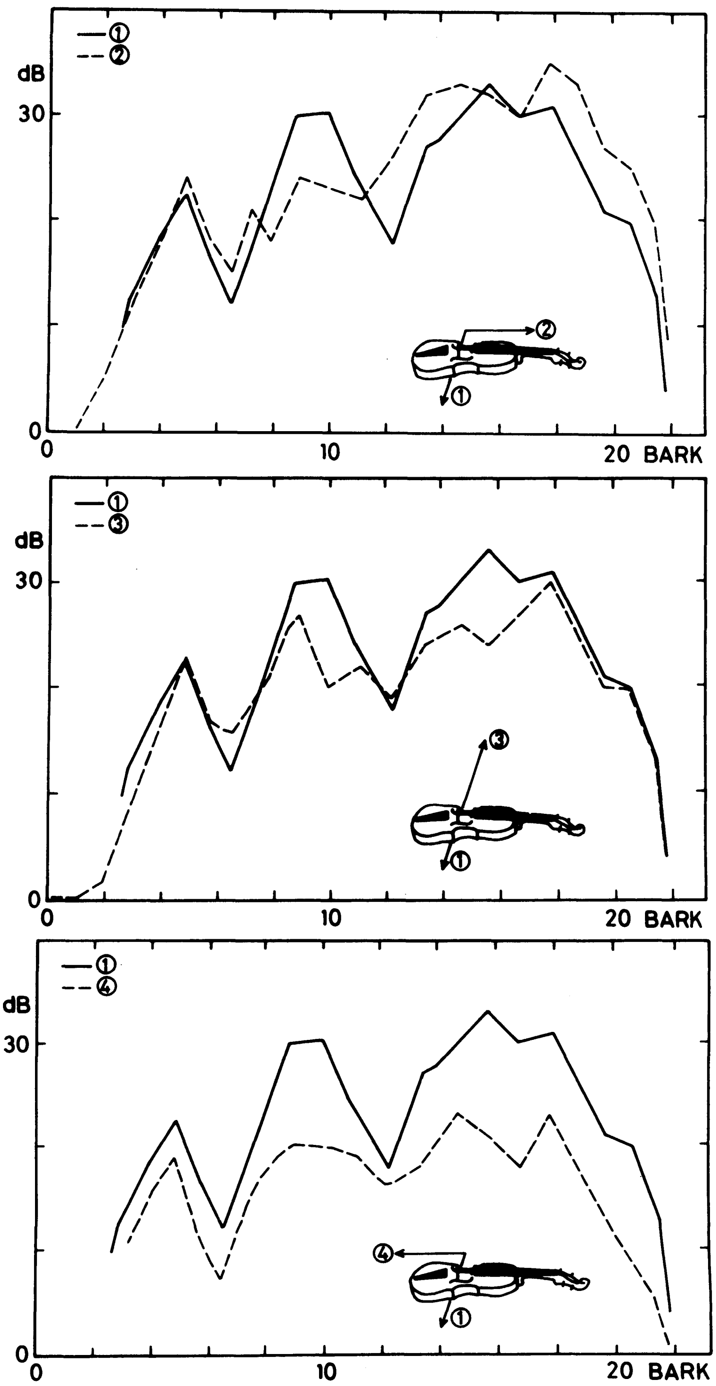



Fig. 8. Long-time-average-spectra of a three octave fulltone scale recorded in four perpendicular directions as marked in the diagrams. The frequency scales are divided in steps equal to the critical bands of hearing, Bark, cf. ZWICKER & FELDTKELLER (1967). Adapted from JANSSON (1976), Acustica.

References:

BENADE, A. (1976): Foundations of Musical Acoustics, Oxford University Press, New York.

BENADE, A. and JANSSON, E. (1974): "On Plane and Spherical Waves in Horns with Nonuniform Flare. I. Theory of Radiation, Resonance Frequencies, and Mode Conversion", Acustica, 31, pp. 79-98.

BURGHAUSER, J. and SPELDA, A. (1971): Akustische Grundlagen des Orchestrierens (transl. by A. Zanger), Gustav Bosse Verlag, Regensburg.

DOUGHTY, J. and GARNER, W. (1948): "Pitch Characteristics of Short Tones. II. Pitch as Function of Tonal Duration", J. Exp. Psychol. 38, pp. 478-494, as reviewed by D. Ward in "Musical Perception", p. 428 in Foundations of Modern Auditory Theory, Vol. I (ed. by J. Tobias), Academic Press, New York 1970.

FLANAGAN, J.L. (1965): Speech Analysis, Synthesis and Perception, Springer Verlag, Berlin, Heidelberg, and New York.

JANSSON, E. (1973): unpublished measurements on the guitar.

JANSSON, E. (1976): "Long-Time-Average-Spectra Applied to Analysis of Music. Part III. A Simple Method for Surveyable Analysis of Complex Sound Sources by Means of a Reverberation Chamber", Acustica, 34, pp. 275-280.

MELKA, A. (1970): "Messungen der Klangeinsatzdauer bei Musikinstrumenten", Acustica, 23, pp. 108-115.

MEYER, J. (1972): Akustik und musikalische Aufführungspraxis, Verlag Das Musikinstrument, Frankfurt a. Main.

ZWICKER, E. and FELDTKELLER, R. (1967): Das Ohr als Nachrichtenempfänger, S. Hirzel Verlag, Stuttgart (2nd edition).

# THE ROOM, THE INSTRUMENT, AND THE MICROPHONE

by
Ulf Rosenberg, Rikskonserter,
Stockholm, Sweden

## Introduction

In the following I will try to give an account of some of the problems associated with the placing of musician and microphone in documentary recordings. Such recordings aim at a realistic acoustic clarity with the highest possible fidelity, and are normally used in connection with e.g. older music. The problems associated with these recordings are of basic importance to most other types of recordings. The documentary recording method may thus serve as a kind of base, or reference, for all other types of recording.

Assuming a documentary recording of a particular composition played by a specific ensemble is to be made, then in a simplified form the problem can be considered as composed of three parts: the choice of recording room, the physical arrangement of the musicians, and the selection and positioning of microphones.

Before making a recording it is essential to know how the recording is to be used. Recordings are generally made for purposes of reproduction in an average living room. But if a recording is intended for reproduction in a theatre for example, quite different demands are placed on the information on the recording room acoustics encoded in the recording. Acoustically, control rooms normally differ significantly from an average living room. Consequently the tone colour perceived in these two acoustical environments are markedly different. Actually this difference may be an important reason for disagreement between technicians and musicians in connection with recordings. It is generally easier for the technician to imagine how colouration will be affected by a changed reproduction environment than for the musician.

## The recording room

Obviously the choice of room is important for a recording. This fact cannot always be taken advantage of, however, sometimes for reasons of dire necessity. Often just one concert room is available. But the reason often appears to be a lack of understanding of the problems involved, or an excessive faith in the power of modern techniques to compensate for serious deficiencies in hall acoustics by means of e.g. synthetic echo.

Earlier, the choise of room for a performance was counted as highly significant as is well known from the history of music. The music was composed with consideration to the specific acoustic conditions prevailing during the performance; it was written with the idea of performning it in a cathedral, a palace state-room, outdoors perhaps, to give a few examples. Even the instruments were chosen with similar considerations. A musician was considered a bad craftsman if he gave no regard to these points of view. Historical guidelines of this kind should of course be allowed to exert an influence on the choice of recording room. Another historical aspect supporting the same conclusion is to be found in the interaction between the concert forms during various epochs and the development of the musical instruments. The development of musical instruments is not a question of technical development in the first place, but rather a matter of adapting the instruments, so that in an interplay with the concert hall they would come to sound "as beautiful as possible". The family of baroque string instruments disappeared with the growth of modern concert life around the turn of the century 1700/1800, and was replaced by the present family of string instruments. Two acoustical reasons for this are conceivable. First this new type of concert hall necessitated instruments with a greater dynamic level. Second the higher partials became harder to hear, among other things because high frequency absorption can no longer be neglected in the new and larger concert halls. Thus while the baroque instruments were deficient with respect to the lowest partials and rich in high partials, the new instruments generated stronger low partials and weaker high partials. In a documentary recording it is essential that such differences in colouration between instruments of differing epochs are preserved. Hence the choice of recording room becomes an important consideration. It is not particularly encouraging, to hear e.g. a baroque flute sounding as an ocarina on a recording. If during the baroque period the ideals of tone colour had been determined by the sound of the ocarina, Fredrik the Great would most probably have played ocarina as a consequence.

A further reason for carefully heeding the choice of recording room has been expressed by Nicolaus Harnoncourt (leader of the 'Concentus Musicus' in Vienna), among others. The essence of what he has to say is that the first thing to be taken into account for achieving a fine recording of music is to allow the musicians to play their music in a suitable acoustic environment, i.e. an environment fair to the music under consideration. Otherwise the musical result will be jeopardized, since it is scarcely credible that a musician should be able to achieve his best work in a defective and uninspiring surrounding. In this respect, recording studios often leave a great deal to be desired.

<u>Positioning of instruments</u>

When the recording room has been chosen, two variables are left which can in principle influence the eventual outcome of the recording. These are the arrangement of the instrumentalists in the room, and the choice and positioning of the microphones. A further possibility, although seldom available, is a rearrangement of the acoustic properties of the room.

With due consideration to the possibilities available, the main task when positioning musical instruments is to attempt to optimize any interactions between the various acoustically reflective surfaces in the room and the different musical instruments. In principle each surface in the room functions as a more or less effective 'acoustic mirror'. If the mirror reflects well, most of the incident sound is reflected, and vice versa. Sound thus behaves in a similar way to light in many respects although in one important respect there is a big difference. The wave length variability of visible light corresponds to slightly less than one octave, and these wave lengths are usually short as compared to other dimentions involved. The range of audible sounds extends through almost eleven octaves and the wave lengths of low notes are often of the same magnitude as room dimensions, or the dimensions of sound radiating surfaces of musical instruments, e.g. the plates of a violin. This means for instance that the low notes of an instrument radiate with reasonable similarity in all directions, while high notes radiate with a pattern which is similar only in certain directions. The difference is roughly illustrated when the light radiation pattern of a naked lamp is compared with the light radiation produced by placing the lamp in the reflector of a spotlight. This dependence on wave length in acoustics has other consequences too. A surface is acoustically reflective only as long as the dimensions of the surface are roughly the same, or greater, than the wave length of the incoming sound wave. Stucco and sculptures, and fixtures and fittings of a similar nature reflect high frequencies, contributing actively therefore to the acoustic properties of a room. On the other hand, the bass register remains relatively unaffected.

What has all this to do with recordings and the placing of musicians? It is the interaction between the sounds radiated from the instruments and the room where the music is played, that determines the overall effect. The most important room surfaces are of course, the floor, walls and ceiling. Reflecting surfaces in the close vicinity of the instruments contribute in general to increase the pregnancy and sharpness of the sound picture by producing reflections which reach the listener, or the microphone, at almost the same time as the sound coming directly from the instrument. The shorter the distance between an instrument and a reflecting surface, the sharper the sound picture becomes. Distant reflecting surfaces contribute to produce reverberation. Achieving a good final result demands a proper mixture of near and distant reflectors, which among other things, requires a reasonably uniform distribution in time of the resulting reflections. This is one of the reasons why it is generally advantageous to place an ensemble at the narrow end of a rectangular room, rather than in the middle of one of the long walls. With the latter arrangement practically all the distant reflections arrive simultaneously. The positioning of instruments in a hall must be carried out of course, so as to achieve a final tone colour result which is "as good as possible". As this is an aesthetic task it is impossible to provide general rules. But in any case, knowledge about the physical behaviour of sound is a valuable help in planning the positioning of the instruments.

What then in principle are the various parameters affecting the end result? A schematical answer may be the following.

a) The radiation lobes, or pattern of radiation, indicate the manner in which sound radiates from the instrument in different directions and at various frequencies. For instance, it is of vital importance that the instruments are positioned in such a way that the lobes cover the audience, the microphones, and useful reflecting surfaces.

A mirror placed in the shadow is unable to reflect the sun. In the same way an acoustically reflecting surface is unlikely to provide reflections unless some sounds impinge on it.

b) The tonal balance between different instruments. The crucial matter during a recording is, first and foremost, the distance between the various instruments and the microphones. But the sound radiation properties of the instrument enter as an important factor even here. An instrument far away from the microphone can become quite dominant if the microphone is placed in a lobe. It is often possible for example, to record both choir and orchestra using the so called single microphone technique, if the microphone is mounted at the same height or a trifle above mouth level of the choir. Inversely the sound of an instrument in the immediate vicinity can be muffled if the microphone is placed beyond the radiation lobe of the instrument.

c) To some extent the room acoustics can be altered by rearranging the positioning of the instruments. This applies not only to recording but also to the sound actually heard in the hall. If the ensemble is placed close to any reflecting surfaces, usually a wall, the timbre becomes more distinct. An opposite effect appears if the ensemble is a long way from a reflecting surface. The opportunities for altering the room acoustic become progressively more limited with larger ensembles: it is impossible to position all members of a symphony orchestra close up to the podium walls.

d) It is quite enough of a problem to figure out an arrangment of the instruments which facilitates ensemble playing. Nicolaus Harnoncourt's statement, referred to earlier, concerning the choice of a place for a performance can be recalled. The same arguments are valid here.

e) Concert halls often have peculiarities which place special requirements on the positioning of instruments. The hall can for example have standing waves at certain bass frequencies. A sign of this is that the tones on either side of these bass frequencies sound weak, while tones coincident with the standing wave frequencies will boom out strongly, but indistinctly. This effect can be accentuated to a greater or lesser extent by the positioning of the bass instruments.

Microphone type and positioning

Within the scope of this paper, the remaining problem in connection with recording is the choice of correct microphone position. With regard to choice of microphone, the only matter that will be discussed here is the selection of suitable directional characteristics for different recording situations. Microphones have lobes as the instruments have; the directional features of microphones are specified in terms of lobes. The three usual types are - omnidirectional, eight pattern, and cardioid. It would take us too far afield to enter into details of situations determining the choice of a particular type of microphone. Simplifying one can say that a directional(eight pattern or cardioid) supplies more direct sound from the instrument(s), while an omnidirectional microphone, being equally sensitive in all directions, provides more information about the room. Simplifying again (and more this time!), the same resulting tone colour can be achieved by use of either an omnidirectional or a directional microphone, by choosing appropriate distances between the instrument and the microphone. With a directional microphone, an increased distance will reduce the direct sound and increase the room information. With regard to tonal balance, an important property of a microphone is that the frequency characteristic should be as straight and linear as possible, irrespective of the direction. This requirement is equally important for all the types of microphone mentioned above. For reasons dependent on the physics of the matter however, it is extremely hard (unfortunately) to achieve good frequency linearity for microphones with a cardioid characteristic. When a directional microphone is required, as in recording with distant microphone or in a room with a 'cathedral acoustic', a eight pattern microphone is far superior as far as colouration is concerned, to a cardioid of comparable performance. It is only to be regretted that the eight pattern microphone seems to be considered out of date these days.

It is difficult within the space available to provide any useful information as far as the actual positioning of microphones is concerned. The general principles mentioned earlier are applicable here too of course, not the least of which are those deriving from the physical nature of sound. A useful rule is always to strive to introduce as little technical equipment as possible. To follow this rule can be quite hard for many technicians suffering from an excessive enthusiasm for technical equipment. This means that the microphone positions should be chosen to obtain a good resultant tonal balance with just one microphone per recorded channel, e.g. two microphones for a two channel stereo recording. This is not always possible of course, but it can be done in far more situations than are usual today. Only when it appears quite impossible, e.g. for reasons of time, to achieve a good tonal balance with one microphone per channel, there is any real basis for switching in support microphones. It is not out of place here to reiterate the importance of the control room. This should be as acoustically similar as possible to the intended reproduction environment in order to facilitate assessments of tonal colouration and timbre.

Final considerations

To do as I have done here, separating the problem areas into groups and then attempting the discover solutions ordered according to this scheme, is something that naturally is seldom possible in practice. The practical problems interweave and it is necessary to compromise until solutions are found which weigh all the various aspects together, those mentioned above plus a few more caused by circumstances. I hope, however, that what has been put forward will contribute to increase understanding of the interaction between physics and music. The intention underlying the title of this paper is just that. An example illustrating the effects on the sound obtained by various positioning of microphones and musicians is given in the sound examples.

Sound examples

A couple of years ago the Stockholm Concert Hall was rebuilt. In order to find out what happened to the sound quality perceived in a couple of listening places in the audience and on an ordinary recording when the orchestra was arranged in different ways on the stage a whole day was spent on experiments. For documentary purposes the entire session was recorded on tape. The recording was made in A/B-stereo using omnidirectional microphones. One recording was made with a pair of microphones placed on the stage as in an ordinary recording. In the other recording two similar microphones were placed and rather far from the stage in two good listening positions in the stall.

To illustrate the effect of merely rearranging the orchestra on the stage four excerpts from the recordings are presented in the sound examples. In <u>Sound exaple one and three</u> the orchestra was arranged as shown in Fig. 1 representing a slightly modified version of a normal orchestra arrangement. In <u>Sound example two and four</u> the orchestra was arranged as in Fig. 2. This arrangement was rather inconvenient to the musicians, as the stage is not intended for this arrangement.

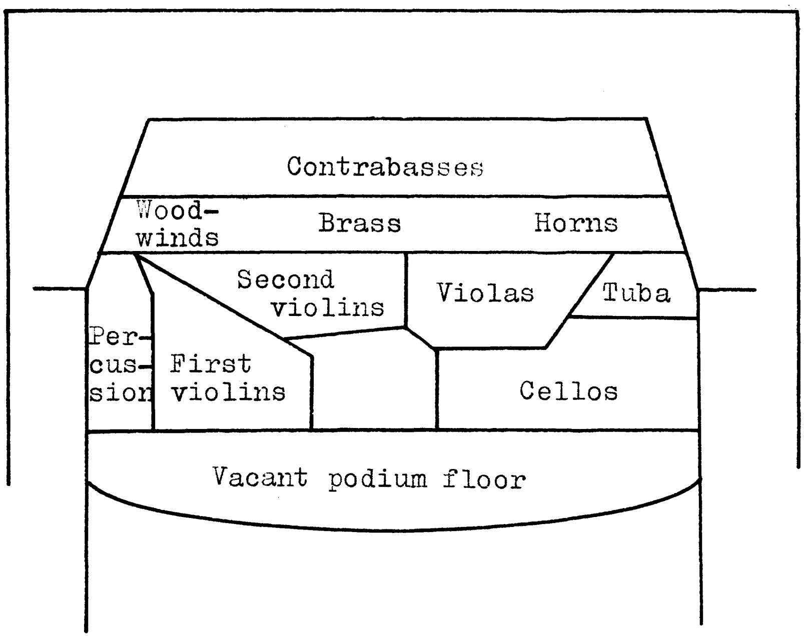



Fig. 1. Schematic view of the arrangement of the orchestra used in sound examples one and two.

It was chosen for stricly acoustic reasons, cf. section Positioning of instruments above. The purpose was to give each group of instruments as much assistance from reflecting walls as was required for each individual group. Examples one and two were both recorded with the microphones on the stage. This positioning of the microphones was not optimal for neither of the two arrangements of the orchestra (but rather for a third arrangement which however is not included in the sound examples). This is particularly true for the arrangement shown in Fig. 2.
<u>Sound examples three and four</u> were recorded simultaneously with examples one and two, respectively, but the microphones were placed in the stall in stead of on the stage, as mentioned.

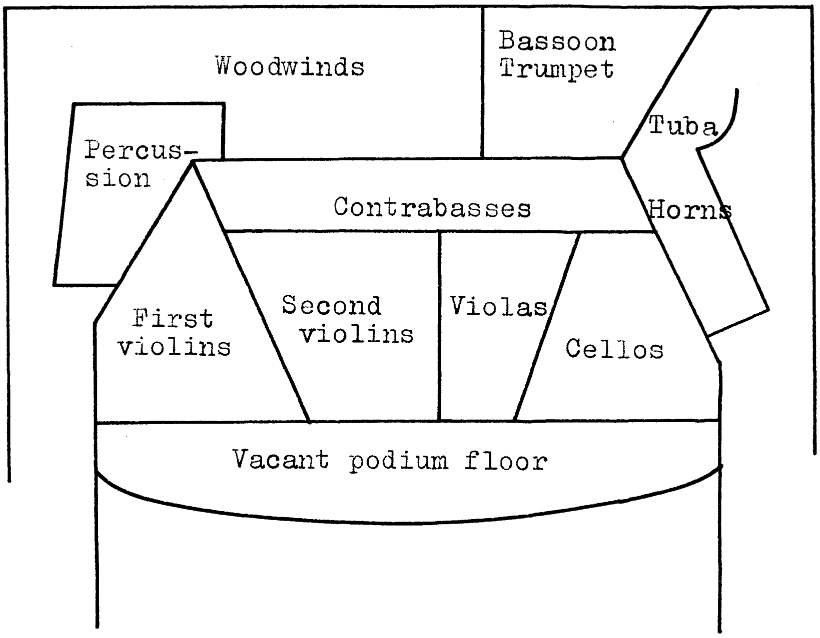



Fig. 2. Schematic view of the arrangement of the orchestra used in sound examples three and four.

Before listening to these examples some more information should be provided. First the examples should not be judged from an esthetical point of view. The musicians found it very hard to play, as the sound from their colleagues' instruments sounded highly unfamiliar. Moreover they were instructed to play exactly in the same way and at the same dynamic level each time. Rather one should listen for the differences in the colour and the level of the sound of the various groups of instruments. Moreover, example one and three should be compared with example two and four, respectively. The difference between examples three and four was greater in reality than can be heard from the recording. It was even greater than the difference between examples one and two. Finally we must apologize for the hum in example four which was due to a cable accident.

# MUSIC AND THE GRAMOPHONE RECORD

## Some viewpoints on how to increase the sound quality

by

Håkan Sjögren, The Institute for Technical Audiology, Karolinska Institutet, Stockholm, Sweden.

## Introduction

When we listen to recorded music over loudspeakers, we enjoy the result of a rather complicated procedure. The system includes many links, starting from the microphone in the recording room to the loudspeaker in the listening room. The number of stages in the process is often as high as 10-15, typically consisting of:

microphones
mixer amplifiers
master tape recorder, probably including a Dolby system
mixing procedure with two or more tape recorders
tape recorder in the cutting room
cutter
stampers
play-back cartridge and tone arm
play-back amplifier
loudspeaker

Besides, of course, we have influence from the acoustics in the recording and listening rooms and from actions and decisions taken by the recording engineer.

The performance of each link in the chain is governed by international standards which specify nominal values and tolerances. It is, however, not possible to use tolerances as wide as those given in the standards because these tolerances are added, and in the worst case will give an end result which is very uncertain. Some recording engineers try to compensate for an error in one link with a deliberate error of reverse sign in the next link.

The reason why the tolerances specified are so wide might be the lack of knowledge on the perception of various errors such as deviations in frequency response, non-linear distortion of different kinds, transient distortion and so on. Therefore more fundamental research in this field is needed.

In this paper I want to discuss only one of the major problems in high-fidelity reproduction: The art of the cutting of gramophone records.

The gramophone record is a mechanical medium, where movements of the groove correspond to an electrical signal which is fed to the cutter. The groove amplitudes are very small, between 0.1 mm and 0.00001 mm. It is rather peculiar that such a medium can indeed be used for sound reproduction.

Engineers and acousticians have been working three decades to refine the technology in this field. The quality today is high, but consumers very often complain about distortion, noise, clicks and the like. With to-day's technology the average sound quality of a gramophone record could be much higher, if the monitoring of the production was more intense.

The main problem within the industry is the process of making the stampers. There are possible solutions to those problems, but I refrain from discussing them here. It would, however, be easier to get high quality pressings if the cut itself is optimum. Then one could be sure that if distortion is found in the record it is not due to a groove amplitude which the play-back cartridge is not able to follow.

Obviously it is desirable to cut the music at the highest possible level in order to get a high signal to noise ratio when playing the record. Two factors which determine the maximum level are:

1. The dynamic force from the walls of the groove on the play-back needle must be less than the pick-up bearing weight.
2. The sound distortion should be less than what is perceivable by listeners.

Dynamic forces

Groove amplitude

The pick-up bearing weight $M_B$ is a mass which is located at the stylus end of the tone arm. Gravity produces a vertical stylus force of

$M_Bg$ ($N/m^2$) where g is the acceleration due to gravity. This force preloads the elastic element of the stylus support, giving a maximum vertical displacement amplitude of the groove modulation of $X_s$ (m). If the compliance of the stylus suspension is called C, we have

$$M_B \geq \frac{X_s}{C}$$

For a given cartridge $M_B$, g and C are known. This gives a criterion on $X_s$, maximum groove amplitude. In practice a limit of 100$\mu$ could be used for cartridges of high quality.

Groove acceleration

If the equivalent mass of the stylus and its elastic support is called $M_S$ and the acceleration of the groove is $a_S$, the following relation obviously holds.

$$M_Bg \geq M_S \cdot a_S$$

This gives a criterion on $a_S$, maximum groove acceleration. A limit of 1000 g ($m/sec^2$) could be used.

Groove velocity

The cartridge is connected to a preamplifier with a standardized deemphasis.

The output of the cartridge is proportional to the groove velocity. Maximum input voltage of the amplifier will give a criterion on maximum velocity. A maximum value of 0,3 (m/sec) can be used.

It follows from the foregoing that it is necessary to measure and indicate amplitude, velocity, and acceleration to plan for the cutting of a record. This is indeed very seldom done in the industry. A common way of measuring the level is to use a peak-indicating instrument connected to the line input of the cutter amplifier. This meter will show an approximation of the groove amplitude and can therefore only be used for popmusic, which has most of its energy in the bass region.

Distortion

W.D. Lewis and F.V. Hunt published the theory of tracing distortion in 1941. (LEWIS & HUNT, 1941) It can be deducted from their work, that distortion components are a function of groove acceleration, a measure we already know of. The theory was extended by CORRINGTON (1949). The results of Corrington can be transformed into values for second and third harmonics, given in table the below.

| | Stereo | Mono |
|---|---|---|
| Second harmonic | $\dfrac{r}{8\pi^2B^2\Omega^2} \cdot a_s$ | 0 |
| Third harmonic | $\dfrac{3r^2}{128\pi^4B^4\Omega^4} \cdot a_s$ | $\dfrac{3r^2}{256\pi^4B^4\Omega^4} \cdot a_s^2$ |

where
$r$ radius of stylus
$B$ groove radius
$\Omega$ rotation speed of the turntable
$a_s$ groove acceleration

There is some information in the literature on the detection of non-linear distortion, see for instance GABRIELSSON & al. (1967) and GABRIELSSON & SJÖGREN (1972). For some kinds of program material 10 % distortion is hardly possible to detect, for other types 1 % is easily heard. There is a need for more studies in this field.

If a limit of 10 % distortion is set, the relation between groove acceleration and groove diameter can be plotted, see Fig. 1.

Records are normally cut giving accelerations of more than 500 g at all groove diameters. This should give 2nd harmonic distortion of far more than 10 %. This can be heard in the sound examples. You will hear music performed by a pop-group cut with (1) full frequency range 50-15000 Hz, (2) high-pass filtered at 10 kHz, (3) full frequency range, (4) high-pass filtered at 5 kHz, and (5) full frequency range. Notice that the high-frequency components are severely distorted - 2nd harmonics and corresponding intermodulation. These distortions are normally masked in the ear by the original programme material.

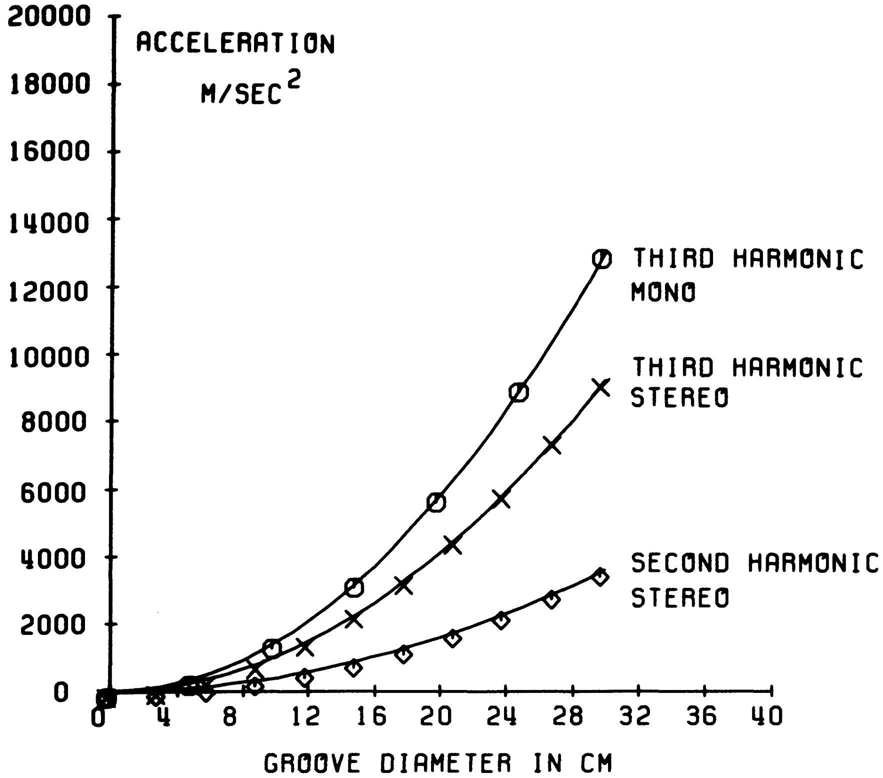



Fig. 1. Maximum recording level at 10 % harmonic distortion.

References:

LEWIS, W.D. and HUNT, F.V.: (1941) "A Theory of Tracing Distorsion in Sound Reproduction from Phonograph Records", J. Acoust. Soc. Am., 12, pp. 348-365.

CORRINGTON, M.S.: (1949) "Tracing Distortion in Phonograph Records", RCA Review, June 1949, pp. 241-253.

GABRIELSSON, A., NYBERG, P.O., SJÖGREN, H. and SVENSSON, L.: (1967) "Detection of Amplitude Distortion by Normal Hearing and Hearing Impaired Subjects". Report No. 83 from Technical Audiology, Karolinska Institutet, Stockholm, Sweden.

GABRIELSSON, A. and SJÖGREN, H.: (1972) "Detection of Amplitude Distortion in Flute and Clarinet Spectra", J. Acoust. Soc. Am., 52, pp. 471-483.

# PANEL DISCUSSION

(Moderator: Hans Åstrand, Royal Academy of Music, Stockholm, Sweden)

In an introduction the chairman pointed out that the short final hour left for the musicians to argue about their problems, after a long days interesting lectures and demonstrations by acousticians, was by no means a depreciation of their importance in the efforts to solve them. On the contrary it was implied in the programme of the symposium that it was to be regarded only as an initial contact between acousticians and musicians, to be followed by the inauguration of a Committee For the Acoustics of Music under the auspices the Royal Academy of Music, where all matters concerning music and rooms will be considered, investigated and at best take the form of advice and admonition.

Three short introductory contributions were given by Eric Ericson, professor at the Music Conservatory of Stockholm, (Musikhögskolan), choir conductor; Erik Lundkvist, organist at the Gustav Vasa Church, Stockholm; and Siegfried Naumann, composer, conductor and teacher at the Music Conservatory of Stockholm.

Eric Ericson's report took the form of a "cry of distress" and a request for help. His experience of concert "halls" of all kinds in Sweden often inspired despair at the murder of all joy in singing that they involve, and he stressed the importance of good acoustics for music and music making, giving examples of all the efforts needed to find the best place for the choir, to adapt their singing to an unknown hall, etc.

Underlining the importance of the announced acoustic Committee, he proposed some kind of "rescue corps" that could depart to make the best possible of the bad halls at least temporarily (removing/adding tapestries, carpets, finding a better place for the podium, etc.), giving preliminary advice about acoustic shells etc. during the years preceding the construction of a new hall.

Erik Lundkvist also uttered a cry of distress especially at the thought of the numerous new churches that are being built now without due planning of the acoustic properties. He compared his experiences of playing organ in continental concert halls, usually of the well-established rectangular shape, to the varying shapes of churches, also referring to the Finlandia House in Helsingfors (where acoustics change radically from section to section or even from seat to seat). Also for the church choirs there are great problems in finding the ideal spot - if there is one - for their performances. The congregation of churches very often get widely different impressions according to where they are sitting on that special occasion.

He strongly advised setting up standards for the construction of new churches where the requirements of organ and choir were respected.

Siegfried Naumann stressed the importance of joining the experience of two different "experience groups", one for the musician's acoustical feeling and the other for the acoustician's field. As an example of the dangers involved in insufficient contacts, he pointed to the evolution of "orchestral" sound: nowadays, acousticians judge the values of the big romantic symphony orchestra, whereas there are several earlier sound experiences that escape us completely, the renaissance music the baroque ensemble, the Mannheim school, etc.

He also pointed out that sound ideas like the renaissance "cori spezzati" are renewed by contemporary composers, who do not always write music for the "philharmonic society"-type of concerts but with quite different parameters. Halls where "movement", ambulating ensembles, instrumental theatre, improvisation can be performed raise new and varying acoustical problems that must be solved in collaboration between musicians and acousticians.

It seemed to him a complicated method to judge acoustics on the basis of orchestral sound, open to so much correction and subjective judgment; he would prefer "neutral" sound effects.

To this Ulf Rosenberg replied that impulse sound can be used and analysed, but the problem is that technique is one matter and music another, and one of the reasons for the meeting was to try and find possibilities to join the two fields. Later on in the following discussion, an acoustician maintained that they normally use nothing but "neutral" sounds, since music gives values that are difficult to measure.

It was asked whether microphones and amplifiers were used to improve the acoustics of a hall. To this Max Mathews said: "Well, I think that is an excellent suggestion to use microphones and loudspeakers to change the characteristics of concert halls. Certainly, there have been a number of halls already that have been treated in this way, and basically, the audience nowadays is too big to be reached by the power that comes out of today's instruments. A violin of this size is an insufficient instrument to come out well in the Philharmonic Hall of New York, for example. I was going to comment that: the Philharmonic Hall is planned to be changed so that the inside is to be taken out and a new structure put in; I think that an alternative to this scheme might rather have been a particular sound reinforcement system - which would also be much cheaper ..."

The acoustician mentioned above also stressed the fact that there <u>are</u> already hundreds of engineers working in the field of music acoustics and that you only have to ask them for advice. But this costs money, and authorities like city councils and others hesitate to spend the few thousand dollars extra that their advice would cost.

The same is true about churches, where several experts are at hand. Much was also said about acoustics in new schools, although not strictly belonging to music acoustics, and there too it was pointed out that the School authorities (Skolöverstyrelsen) do have information and standards worked out for different school rooms.

Since much was said about the two groups - acousticians and musicians - not always speaking to each other and not always using the same language or understanding its implications, Siegfried Naumann indicated one cathegory of experts for whom no regular education is yet arranged in Sweden, i.e. the "Tonmeister", who should basically be able to serve as an interpreter between the two groups, normally being well equipped both in techniques and in music. To this Krister Malm added that the Organisation occupied in forming new education schedules (OMUS) already had a reference group asked to investigate links between science and music - this seemed to be a task for them. Malm also had asked if the conference was not too much centered on one sort of music culture, that for instance already 80, maybe 90 % of the music played in western countries - live music, that is - is transmitted by loudspeakers. Electric instruments dominate our music today. The development of music seems to be far ahead of the problems discussed here ....

Ulf Rosenberg finally asked if we did not often play the right kind of music in the wrong place - which seems well to demonstrate one basic problem, i.e. that a hall is expected to serve for almost any kind of music, which is demanding very much of any hall.

## Sound examples

Side A

Mathews: Analysis and synthesis of timbres

Track I

1. J.C. Risset's trumpet tone synthesis (25")
2. D. Morill's brass and percussion sound synthesis, c.f. Fig. 4 (1'25")

Track II

3. Ever-ascending tone pitch paradox, c.f. Fig. 6 (1'30")
4. Ever-descending glissando pitch paradox, c.f. Fig. 6 (1'10")
5. Ever-increasing tempo rhythm paradox (50")

Track III

6. Scale played with three different adjustments of the Q values of an electronic violin resonances, c.f. Fig. 9, I, II and IV (1'10")
7. Duet with electronic violin and acoustic violin (40")

Side B

Mathews: Analysis and synthesis of timbres (continued)

Track I

8. Violin with variable resonance, c.f. Fig. 10. From M. Urbaniak: "Fusion", Columbia Record KC32852 (45")
9. Violin with brass timbre, c.f. Fig. 11 (25")
10. Super-bass violin, excerpt from Mathews's "Elephants Can Safely Graze" (1'40")

Track II

11. Excerpt from J. Olive's "Mar-ri-ia-a". Text: see Appendix (2'55")

Bengtsson-Gabrielsson: Rhythm research in Uppsala

Track III

1. Excerpt from "Kettenbrücke Waltz" by Johann Strauss senior, performed by the Boskowsky ensemble, c.f. Figs. 4, 5 and 6 Philips SGL5757 (10")
2. Excerpt from "Bridal March" by Gelotte, performed on a key-fiddle by E. Sahlström, c.f. Fig. 7 (17")

3. Excerpt from "Polska" from Dalecarlia (Sweden), performed on a "spelpipa" by E. Åhs, c.f. Fig. 8 (25")
4. Three examples of monophonic rhythm stimuli, performed on a bongo drum or a side drum and used in the experiments concerning rhythm experience (30")

Side C

Bengtsson-Gabrielsson: Rhythm research in Uppsala (continued)

Track I

5. Examples of polophonic rhythm stimuli produced by an electronic "rhythm box" (rock'nroll, bossanova, mixture of habanera and beguine) and used in the experiments concerning rhythm experience (35")
6. The beginning of four pieces of dance music (mambo, swing, Swedish waltz, duet-polska) used in the experiments concerning rhythm experience (1')

Sundberg: Singing and timbre

Track II

1. Soprano singing and synthesis
   a) a vowel sung at four pitches
   b) synthesis of a) with pitch dependent formant frequencies as used by the singer
   c) synthesis of a) with constant formant frequencies as used by the singer in the lowest pitch (55")
2. Effect of the "singing formant", c.f. Fig. 7
   a) noise with the same average spectrum as an orchestra
   b) singing without and with "singing formant", 3 times
   c) a) and b) simultaneously (1'05")
3. Bass, baritone and a tenor synthesis of the vowels in hard, heed and who'd
   a) bass
   b) baritone
   c) tenor (45")
4. Synthesis of a sung vowel
   a) the singer's production of a vowel
   b) synthesis of a) (10")

Track III

5. Effect of increasing the number of harmonic partials of equal acoustic amplitudes
   a) partials 1-2 (3 times)
   b) partials 1-4 (3 times)
   c) partials 1-6 (3 times)
   d) partials 1-8 (3 times) (1'30")
6. Alto and tenor synthesis alternating 3 times (25")
7. Smooth and rough in organ timbre
   A piece of music is played twice on an organ, first on a Gedackt 8' stop, which generates odd-numbered partials only, and then on a Principal 8' stop, which generates a complete series of partials (30")
8. Dissonance and consonance in dyads
   a) pure major third (100 Hz and 125 Hz), major seventh (440 Hz and 831 Hz), and pure octave (440 Hz and 880 Hz) with sinusoids
   b) same as a) but with complex tones (1')

Side D

<u>Jansson: On the acoustics of musical instruments</u>

Track I

1. Excerpts from recordings of four violins in a reverberation chamber, c.f. Fig. 11
   a) violin 1 and violin 2
   b) violin 1 and violin 3
   c) violin 1 and violin 4 (35")
2. Excerpts from recordings of a Violino Grande played by B. Eichenholtz in a reverberation chamber, c.f. Fig. 12
   a) the scale
   b) from J.S. Bach: "Sarabande" from Suite nr 6 for Violoncello solo, D major, BWV 1012 (30")

<u>Jansson: On sound radiation of musical instruments</u>

Track II

1. Excerpts from recordings of a violin in four different directions in an unechoic chamber, c.f. Fig. 8
   a) the microphone to the right of the player and in front of the player

b) the microphone to the right of the player and to the left of the player

c) the microphone to the right of the player and behind the player (45")

Rosenberg: The room, the instrument, the microphone

Track III

Symphony orchestra in Stockholm Concert Hall

1. Orchestra arranged according to Fig. 1. Microphones on the stage (50")
2. Orchestra arranged according to Fig. 2. Microphones on the stage (50")
3. Orchestra arranged according to Fig. 1. Microphones in the stall (50")
4. Orchestra arranged according to Fig. 2. Microphones in the stall (50")

Sjögren: Music and the grammophone record

Track IV

Music performed by a pop-group

a) 50 - 15 000 Hz

b) 10 000 - 15 000 Hz

c) 50 - 15 000 Hz

d) 5 000 - 15 000 Hz

e) 50 - 15 000 Hz (1')

Appendix

## Excerpts from Mar-ri-ia-a

### A Miniature Opera for Soprani, Computer and Chamber Ensemble

Libretto: Joan S. Olive     Music: Joseph P. Olive

Machine: Wonderful (repeatedly).

Maria: Wonderful, your second word,
my wonderful talking machine.

Machine: I am a talking machine.

Maria: Wow. Incredible.
Now, the button.

Machine: "Good evening, how are you?
Sociable program number one now in progress."

* * *

Machine: Maria, Maria my darling sweet heart,
The most amorous lovely and gentle girl,
My dear, you're so cute, you're so terribly nice,
I love you whenever you're near.
My pretty and my most charming love,
When you come near, my motor is hot,
My discs whiz and my voltage soars,
I need you so very much.
I love you, I adore you and cherish you so,
That I won't let you leave me alone.

Chorus: So that he won't let her leave him alone,
Loves and won't let her leave him alone.

Maria: Machine, my machine my shiny bright love,
The most talkative sturdy and handsome one,
Machine you're so smart and so verbally clear,
Your talk brings my gentle heart near.

I love my metalic creation,

Machine: I love my dearest creator,

Maria: I want to hold you so tightly,

Machine: Come close my metal is yearning.

Maria: Your frame is so hard you feel so cold,
Machine: Your skin is so soft you feel so good,
Chorus: He loves her... she loves him...

Maria: Your knobs are so smooth to my gentle touch,
Machine: Your curves are so smooth to my gentle touch,

M & M: Your are mine, I am yours to love and adore,

M & M: I am mad with my love for you.

M & M: Love, love, I love you, love, love, I love you
Chorus: Dear love... Sweet love...

Maria: Love, machine.
Machine: Love, Maria.
Chorus: Love, love.