Dept. for Speech, Music and Hearing

Quarterly Progress and Status Report

Computer program for VT-resonance frequency calculations

Liljencrants, J. and Fant, G.

journal: STL-QPSR
volume: 16
number: 4
year: 1975
pages: 015-020

B. COMPUTER PROGRAM FOR VT-RESONANCE FREQUENCY CALCULATIONS

J. Liljencrants and G. Fant

Abstract

Two programs for deriving vocal-tract resonance frequencies from area functions have been developed. The purpose has been to obtain maximum exactness in implementing a multiple tube representation of the loss-less vocal tract by treating the length of any part tube as a continuous parameter and avoiding quantizations of the wave propagation quantization. One program is based on the volume velocity transfer from the glottis to the lips, the other program is based on the input impedance as seen from the glottis. The two methods gave results that differed by less than 1 Hz in any resonance frequency. The computer program plots a quasi-spectrum envelope with a print-out of resonance frequencies.

Introduction

The need for these programs grew out of a project on deriving theory for vocal-tract length perturbations, a report of which is included in this issue of the STL-QPSR, Fant (1975). The programs also serve the purpose of a general tool for VT-calculations. The circuit theory basis for the two programs is well established, Fant (1960), whilst the computer programs and print-out technique are more specific.

Theory

A representation of the vocal tract in terms of a cascade of transmission lines, one for each tube of the acoustic system is schematized in Fig. I-B-1. The three-element exact representation of one such tube, disregarding losses is:

$$a_n = Z_n j\,tg\left(\frac{1}{2} \cdot \frac{\omega \ell_n}{c}\right) \qquad (1)$$

$$b_n = \frac{Z_n}{j \sin(\omega \ell_n / c)} \qquad (2)$$

$$a_n + b_n = \frac{Z_n}{j\,tg(\omega \ell_n / c)} \qquad (3)$$

$$Z_n = \rho c / A_n \qquad (4)$$

where ℓ_n is the length and A_n the cross-sectional area of tube nr n. The standard value of c = 35.300 cm/s is used for the velocity of sound at 35^{o}C.

The radiation inductance is

$$\left.\begin{aligned} b_o &= j\omega \ell_o / A_o \\ \ell_o &= 0.8\sqrt{A_o/\pi} \end{aligned}\right\} \qquad (5)$$

Corrections associated with radiation from a narrow tube into a joining, much wider, tube are generally not included since actual VT configurations seldom display such discontinuities. These effects may, however, be taken into account by adding internal end corrections, Fant (1960).

The notations for elements within the system determinant $\Delta_q = i_q/i_o$ are simplified by the substitutions $y = \omega/c$, $k_n = A_{n+1}/A_n$, and

$$d_{o1} = (b_o + a_1 + b_1)/b_1 \qquad (6)$$

$$d_{n,n+1} = \frac{a_n+b_n+a_{n+1}+b_{n+1}}{b_{n+1}} = \cos(y\ell_{n+1}) + k_n \cos(y\ell_n)\frac{\sin(y\ell_{n+1})}{\sin(y\ell_n)} \quad (7)$$

$$b_{n,n+1} = \frac{b_n}{b_{n+1}} = k_1 \frac{\sin(y\ell_{n+1})}{\sin(y\ell_n)} \quad (8)$$

As an example, the four-tube system determinant may be written

$$\Delta_4 = \begin{vmatrix} d_{01} & -b_{12} & 0 & 0 \\ -1 & d_{12} & -b_{23} & 0 \\ 0 & -1 & d_{23} & -b_{34} \\ 0 & 0 & -1 & d_{34} \end{vmatrix} \quad (9)$$

In general, starting from the lip end, the calculation involves the following sequence of operations

$$\left.\begin{aligned} \Delta_1 &= d_{01} \\ \Delta_2 &= d_{12} \cdot \Delta_1 - b_{12} \\ \Delta_3 &= d_{23} \cdot \Delta_2 - b_{23}\Delta_1 \\ \Delta_4 &= d_{34} \cdot \Delta_3 - b_{34}\Delta_2 \end{aligned}\right\} \quad (10)$$

etc. up to the glottal end

$$\Delta_q = d_{q-1,q}\Delta_{q-1} - b_{q-1,q}\Delta_{q-2} \quad (11)$$

This is the general solution. The resonance frequencies are found by setting $\Delta_q = 0$. If all tubes are of the same length ℓ but not small compared to the wavelength

$$\left.\begin{aligned} d_{01} &= \cos(\ell y) - \ell_o y \sin(\ell y) \\ d_{n,n+1} &= (1+k_n)\cos(\ell y) \\ b_{n,n+1} &= k_n \end{aligned}\right\} \quad (12)$$

With a large number of tubes of an equal quantal length $\Delta\ell$, one may approximate

$$\left.\begin{aligned} \cos(\Delta \ell y) &= \left[1 - \frac{(\Delta \ell y)^2}{2} \right] \\ \ell_o \cdot y \sin(\Delta \ell y) &= \frac{\ell_o}{\Delta \ell} (\Delta \ell y)^2 \end{aligned}\right\} \quad (13)$$

to be substituted in Eq. (12)

The second approach is to calculate the input impedance of the vocal tract at the glottis starting by the radiation impedance

$$Z_{io} = b_o \quad (14)$$

We proceed to the impedance Z_{i1} at the intersection between tubes 1 and 2 looking towards the lips

$$Z_{i1} = Z_1 j tg(\frac{\omega \ell_1}{c} + artg \frac{b_o}{Z_o}) \quad (15)$$

and similarly

$$Z_{i2} = Z_2 j tg[\frac{\omega \ell_2}{c} + artgk_1 tg(\frac{\omega \ell_1}{c} + artg \frac{b_o}{Z_o}) \quad (16)$$

$$Z_{iq} = Z_q j tg \left\{ \frac{\omega \ell_q}{c} + \ldots \; artgk_2 tg[\frac{\omega \ell_2}{c} + artgk_1 tg(\frac{\omega \ell_1}{c} + artg \frac{b_o}{Z_o})] \right\} \quad (17)$$

The recurrent operation is a calculation of phase shift

$$\varphi_{im} = \frac{\omega \ell_m}{c} + artgk_{m-1} tg \varphi_{i,\, m-1} \quad (18)$$

The VT resonance frequencies are found when

$$\left.\begin{aligned} &Z_{iq} = \infty \\ \text{or} \quad &\varphi_{iq} = (2n-1)\frac{\pi}{2} \end{aligned}\right\} \quad (18)$$

This is conceptually simpler than the transfer determinant calculation and the complete computer program is somewhat but not crucially faster. However, the phase function is not so easy to handle correctly as the transfer function.

Program

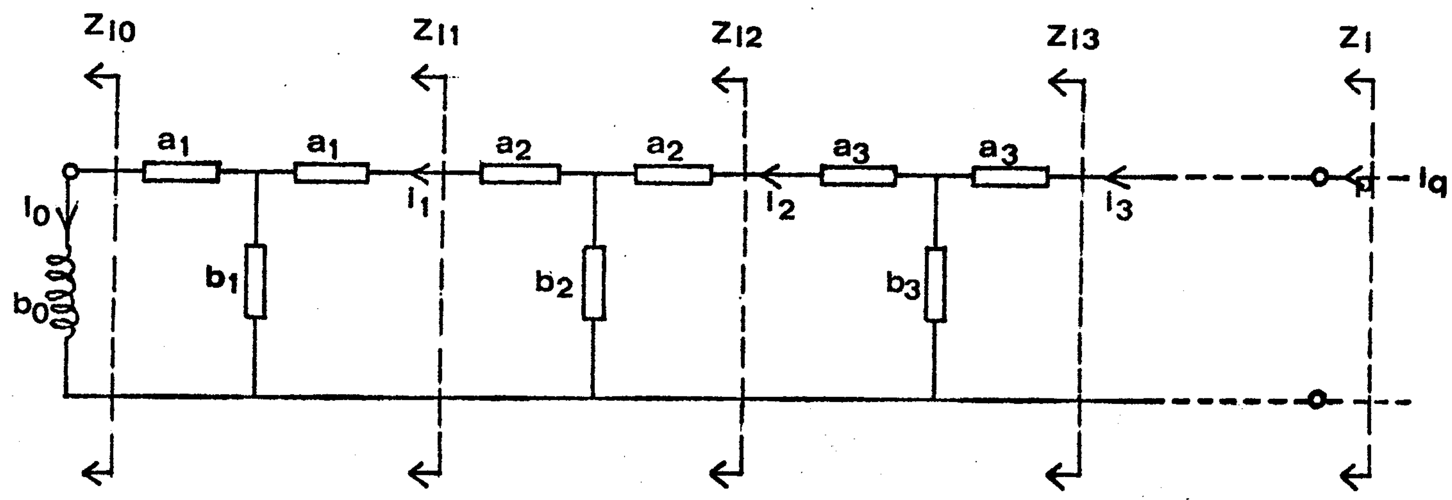
The program starts with routines for manual input of the number of tubular sections, M, and the tables A(N) of section areas, and L(N) of section lengths. The shape of the model can then be tabulated and plotted as diameter vs axial coordinate. The operator selects which computational method to be used, and if the alternative with sections of equal lengths was selected an auxiliary area table is derived from A and L.

In the computations the program will branch to different subroutines according to the method selected. The determination of resonance frequencies, result-plotting etc are common parts of the program. To determine the resonance frequencies the main program scans the frequency from zero and upwards in increments. For each frequency the computation subroutine is called, and it ends up with a number Y. This is designed to reach a maximum of one at the resonance frequencies, and to vary continuously between zero and one for other frequencies. The task of the main program is thus to find the peaks in Y.

In the impedance method phase angles are found where the angle goes through $\pi(2n-1/2)$. A complication will rise, however, because the angle is determined by repeated uses of the arctg function. This function always gives a result in the range $\pm\pi/2$ and does not keep track of which branch of the function that should be used. Because of this, a plot of the phase angle vs frequency will display numerous discontinuities due to branch changes in the arctg somewhere in the computational chain. Apart from being visually disturbing this makes it difficult to locate the points of interest. One simple remedy is to compute the end result Y as the squared sine of the phase angle. This will then oscillate continuously when the frequency is swept and the formant frequencies are located at the peaks. In the same time a plot of Y vs frequency will show a superficial similarity to the transfer function corresponding to the model.

In the transfer method the criterion of a formant frequency match is that the system determinant is zero. To maintain compatibility to the impedance method the end value is recomputed as $Y = \cos^2(\text{arctg}(Y))$.

Now to determine the resonance frequencies the main program scans the frequency from zero and up. For each frequency Y is computed with the pertinent subroutine and the difference in Y from the previous value is formed. This difference in turn is compared to the difference obtained last time, and when there is a change of sign from positive to negative a formant frequency has been coarsely located.



$$b_o = j\omega\rho\, l_o / A_o \qquad a_n = Z_n\, j\, tg\, (\omega l_n / 2c)$$

$$l_o = 0{,}8\sqrt{A_o/\pi} \qquad b_n = Z_n / j \sin\, (\omega l_n / c)$$

$$Z_n = \rho c / A_n \qquad a_n + b_n = Z_n / j\, tg\, (\omega l_n / c)$$

Fig. I-B-1. T-network analog of loss-less multisection vocal tract.

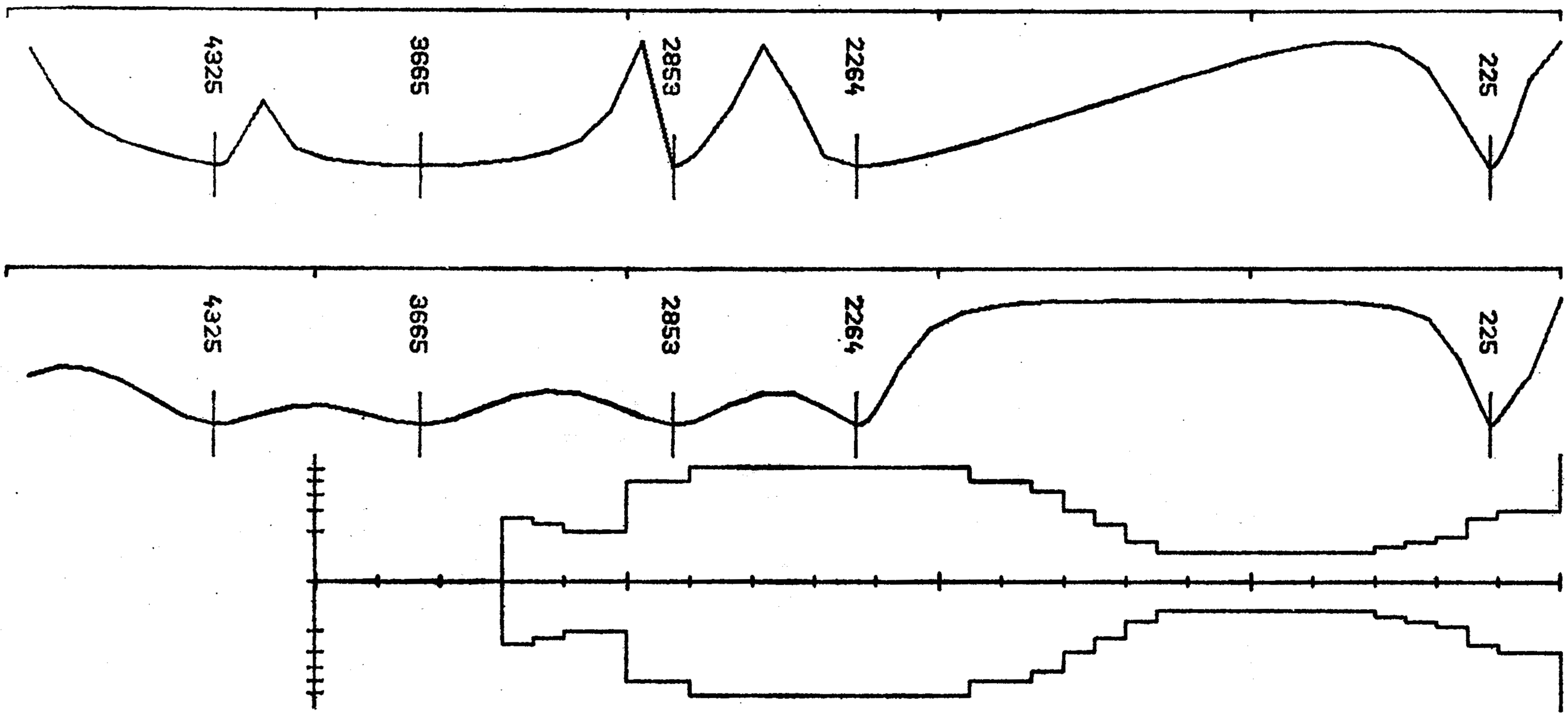


Fig. I-B-2. Computer print-outs for the two methods, vowel [i]. From top the diameter function and the Y functions for the transfer determinant and impedance phase methods.

The frequency step should be taken as large as possible in order to minimize computing time. On the other hand, with increasing step there is a risk that the peaks in Y are missed altogether if they are narrower than the step. Also, in some instances it was found that these peaks are heavily asymmetric so that interpolation would give little improvement in accuracy as compared to just selecting the frequency with the largest value of Y. To resolve these difficulties the peak finding algorithm initially uses a comparatively large frequency step, 125 Hz. When a peak has been grossly located the frequency is reset to the value just below the peak and a new search is done with a 5 times smaller step. When the peak now is detected the scan is repeated still another time, now with a 1 Hz step which eventually gives the final result. Then the step is reset to the initial 125 Hz and the scan proceeds to locate the next peak in Y.

During execution Y is plotted vs frequency. To make a clean plot without backspacing along the frequency axis the plotting is made to lag two frequency steps behind the one being computed and the backspaced points are suppressed. Print-outs for the two methods are shown in Fig. I-B-2. There is some perceptual relevance of the Y vs frequency graph of the transfer method since regions containing two or more resonances that build up an area of prominence show Y-values close to 1 even between peaks.

Appendix

A. Subroutine to compute Y from F, L(N), and A(N) with the transfer determinant method.

B. Subroutine to compute Y from the frequency F, the lengths L(N), and the areas A(N), the impedance phase shift method. C is the speed of sound.

References

FANT, G. (1960): Acoustic Theory of Speech Production, Mouton, 's-Gravenhage (2nd edition 1970).

FANT, G. (1975): "Vocal-tract area and length perturbations", STL-QPSR 4/1975, pp. 1-14.

APPENDIX A, TRANSFER DETERMINANT

```
      WC=6.2832*F/C
      N=1
      V=WC*L(1)
      SV=SIN(V)
      CV=COS(V)
      D1=1
      Y=CV-WC*L(0)*SV
312   N=N+1
      IF M-N,319,313,313              BRANCH TO EXIT
313   SV1=SV
      CV1=CV
      V1=V
      V=WC*L(N)
      D2=D1
      D1=Y
      SV=SIN(V)
      CV=COS(V)
      KP=A(N)/A(N-1)
      BP=KP*SV/SV1
      DP=CV+BP*CV1
      Y=DP*D1-BP*D2
      GOTO 312
319   Y=(COS(ATG(Y)))**2
```

APPENDIX B, IMPEDANCE PHASE SHIFT

```
      WC=6.2832*F/C
      N=0
      VNP=ATG(WC*L(0))
122   N=N+1
      IF M-N,124,123,123              BRANCH TO EXIT
123   VNP=WC*L(N)+ATG(A(N)*SIN(VNP)/(A(N-1)*COS(VNP)))
      GOTO 122
124   Y=(SIN(VNP))**2
```