A QUASI-ONE-DIMENSIONAL DIGITAL SIMULATION FOR THE TIME-VARYING VOCAL TRACT

by

Michael Rodney Portnoff

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREES OF

BACHELOR OF SCIENCE,

MASTER OF SCIENCE,

and

ELECTRICAL ENGINEER

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1973

Signature of Author_____
Department of Electrical Engineering, May 11, 1973

Certified by_____
Thesis Supervisor (Academic)

Certified by_____
Thesis Supervisor (VIA Cooperating Company)

Accepted by_____
Chairman, Departmental Committee on Graduate Students

A QUASI-ONE-DIMENSIONAL DIGITAL SIMULATION FOR THE TIME-VARYING VOCAL TRACT

by

Michael Rodney Portnoff

Submitted to the Department of Electrical Engineering on May 11, 1973 in partial fulfillment of the requirements for the Degrees of Bachelor of Science, Master of Science, and Electrical Engineer.

ABSTRACT

A digital simulation which models certain aspects of the human speech-production process is developed from fundamental principles. Specifically, the transmission of sound in the vocal tract is modeled as quasi-one-dimensional acoustic wave propagation in a nonuniform time-varying tube with yielding walls. The vocal tract is terminated by a radiation load based on a simplified model of a piston in a baffle.

The formulation of the model is developed in two stages. The first stage is the formulation of a mathematical description of the physics of sound propagation in the vocal tract. This description takes the form of a system of partial-differential equations and boundary conditions derived from fundamental physical principles. The simplifying assumptions invoked in the derivation, together with the conditions for which they are valid, are clearly stated. The second stage of the modeling process is the actual implementation of the model; that is, the formulation of finite-difference equations derived from the differential equations and the solution of these equations on a digital computer. In this latter stage, techniques and insights from digital signal processing and numerical analysis prove valuable and provide novel approaches to some of the modeling problems.

THESIS SUPERVISOR (Academic): Kenneth N. Stevens
TITLE: Professor of Electrical Engineering

THESIS SUPERVISOR (VIA Cooperating Company): Ronald W. Schafer
TITLE: Member of the Technical Staff, Bell Laboratories

ACKNOWLEDGMENT

I wish to express my deep appreciation to Dr. Ronald Schafer for his guidance, encouragement, and remarkable patience throughout the course of this study. His insights and understanding of the problems considered here, suggested in our numerous discussions, are reflected throughout this thesis.

Also, I wish to express my gratitude to Professor Kenneth Stevens, who served as advisor for this thesis, for his time interest, and critical suggestions.

I am indebted to Dr. James Flanagan, head of the Acoustics Research Department at Bell Laboratories, for his interest, support, and critical reading of the manuscript; and to the other members of his department for their friendly help and encouragement. Particularly, I would like to thank Dr. Mohan Sondhi for our many discussions and for his often embarassing questions; and Miss Carol McGonegal for her willing help in my constant battle with the DDP-516 computer.

TABLE OF CONTENTS

CHAPTER I

Introduction

The production of speech and speech-like sounds had been attempted as early as nine-hundred years ago,[1] but it was not until the late Eighteenth Century that the first analog of the speech production mechanism was demonstrated. In 1779, the Imperial Academy of St. Petersburg awarded its annual prize to C. G. Kratzenstein for demonstrating the physiological differences among five vowels. Kratzenstein constructed a model of the speech production mechanism consisting of a system of acoustic resonators excited by vibrating reeds.[2] In 1791, von Kempelen[3] demonstrated a more sophisticated speaking machine capable of producing connected utterances. In later years, many notable scientists including Wheatstone,[4] Helmholtz[5] and Alexander Graham Bell[6] became interested in the problem of artificially producing speech.

Although the first speech synthesizers were, for the most part, objects of curiosity and playthings, they did provide some insight into the mechanisms of speech production and perception. Today, interest in synthetic speech is not only motivated by a desire for a better understanding of the natural speech processes, but also by more practical considerations viz., the need for efficient means of storing and transmitting the information contained in speech signals. Moreover, the successful medical diagnosis

and treatment of speech disorders can be facilitated by a clear understanding of the critical parameters of the speech production mechanism. It is, then, applications such as these that provide the present motivation for modeling the human articulatory system.

The principal features of the human vocal apparatus are shown in the sagittal-plane X ray of Fig. 1.1. The vocal tract is a nonuniform acoustic tube which, in the adult male, is approximately 17 cm in length. It is terminated at one end by the lips, and at the other end by the glottis. In addition, the vocal tract may be shunted by the nasal tract through the trapdoor action of the velum.

Figure 1.2 shows a schematized diagram of the vocal apparatus which will be the basis of the present simulation. Also included in the figure is the subglottal system consisting of the lungs, bronchi, and trachea. The lungs are represented as the air reservoir at the left of the diagram. The pressure of the air in the lungs, P_s, is increased by application of muscular force from the thoracic muscles thus producing a flow of air with volume velocity U_g through the glottis. The vocal cords are modeled as a system of coupled mechanical oscillators driven by the subglottal pressure and the local Bernoulli pressure resulting from the air flow through the glottal orifice. The vocal tract and nasal tract are modeled as nonuniform acoustic ducts with yielding walls. An acoustic wave propagates through this system and generates a radiated wave at the lips and nostrils. The speech signal

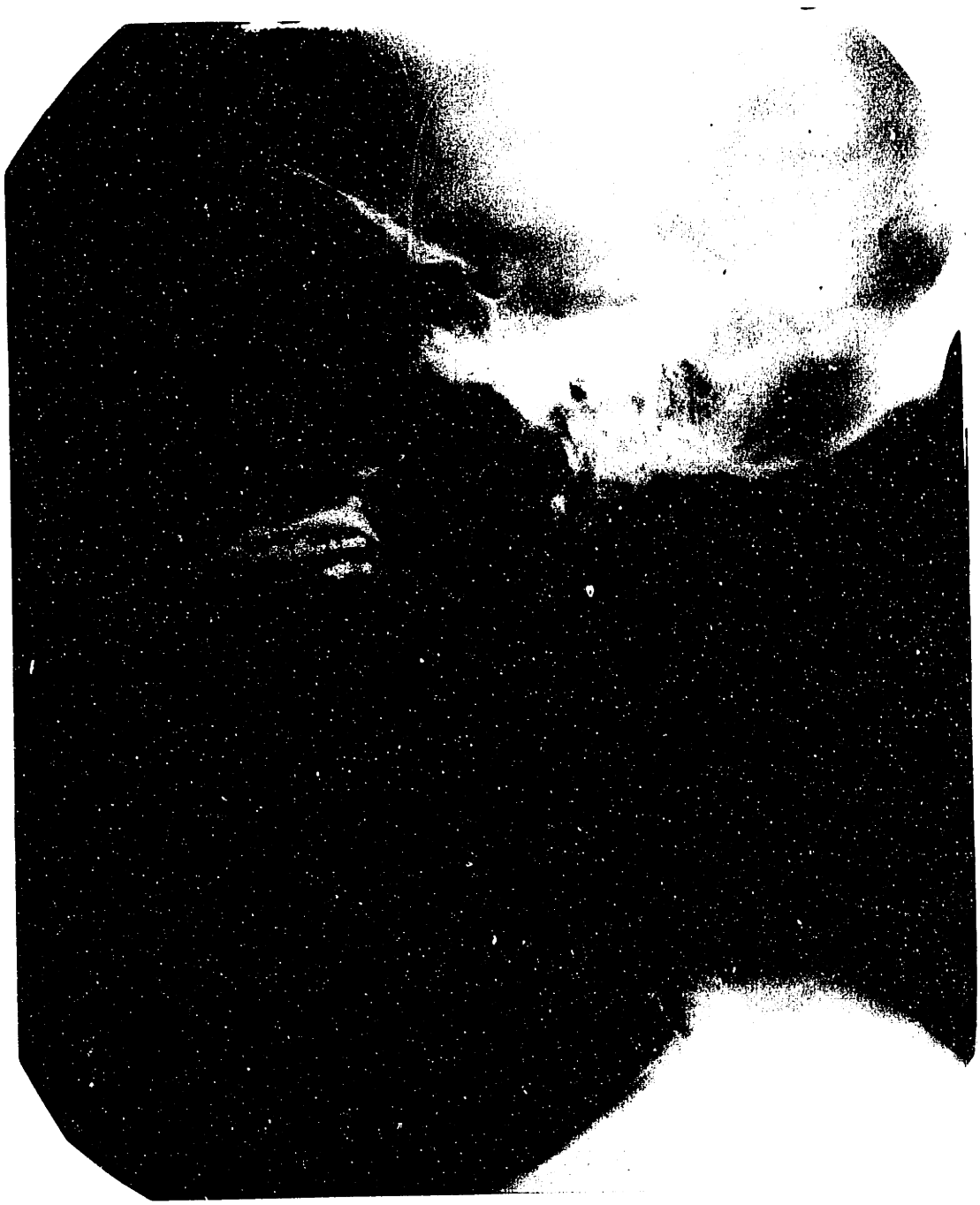


FIGURE 1.1

SAGITTAL-PLANE X RAY

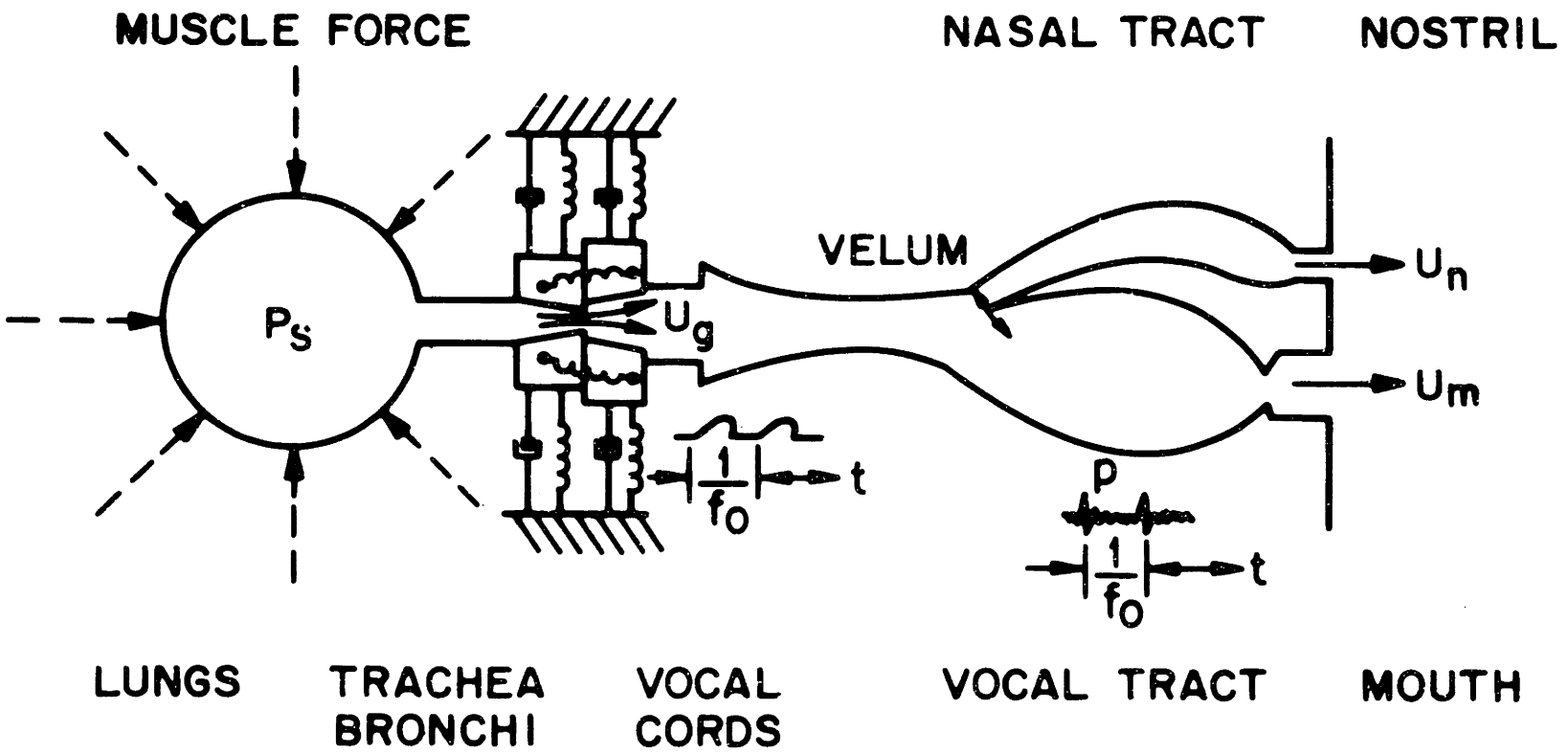


FIGURE 1.2

SCHEMATIC DIAGRAM FOR MODEL OF HUMAN VOCAL APPARATUS

(Adapted from Flanagan, et. al.[21])

is then taken as a linear combination of the pressure waves radiated at the lips and nostrils.

Recently, considerable effort has been applied to modeling the vocal cords.[7-11] Therefore, this paper will be concerned primarily with modeling the vocal tract and radiation load at the lips and nostrils. The formulation of the model will be developed in two stages. The first stage of the modeling process consists of formulating a mathematical description of the physics of sound propagation in the vocal tract. This description is in the form of a system of partial-differential equations and boundary conditions derived from such fundamental physical laws as the conservation of mass and the conservation of momentum. The approximations made at this level are based primarily on physical considerations and previous experience modeling acoustic systems. The second stage of the modeling process is the actual implementation of the model; that is, the formulation of finite-difference equations derived from the partial-differential equations and the solution of these equations on a digital computer. At this level, approximations are based primarily on mathematical considerations and draw on previous experience from numerical analysis and digital signal processing. In fact, some of these considerations, such as numerical stability and round-off noise, have no "physical" basis at all.

It is this latter aspect of the modeling process that has often received too little attention in previous vocal-tract simulations. As an example, consider the following

analysis which demonstrates the unsuitability of a common approach to the problem of simulating the acoustic transmission properties of the vocal tract. According to this approach, the vocal tract is approximated as a piecewise uniform acoustic tube, each section of which is approximated as a cascade of "T" networks containing analogous circuit elements.* The resulting lumped-parameter network is then simulated by replacing the time derivatives by backward (finite) differences. For simplicity, suppose the vocal tract is modeled as a lossless uniform acoustic tube with rigid walls. The analogous network based on cascaded "T" sections is shown in Fig. 1.3. The analogous inductance per unit length is given by L, the analogous capacitance per unit length by C, and the current in the k^{th} loop by $U_k(t)$. The loop equation for the k^{th} current loop gives

$$\hat{U}_{k-1}(s) - 2\hat{U}_k(s) + \hat{U}_{k+1}(s) = LC\Delta x^2 s^2 \hat{U}_k(s)$$

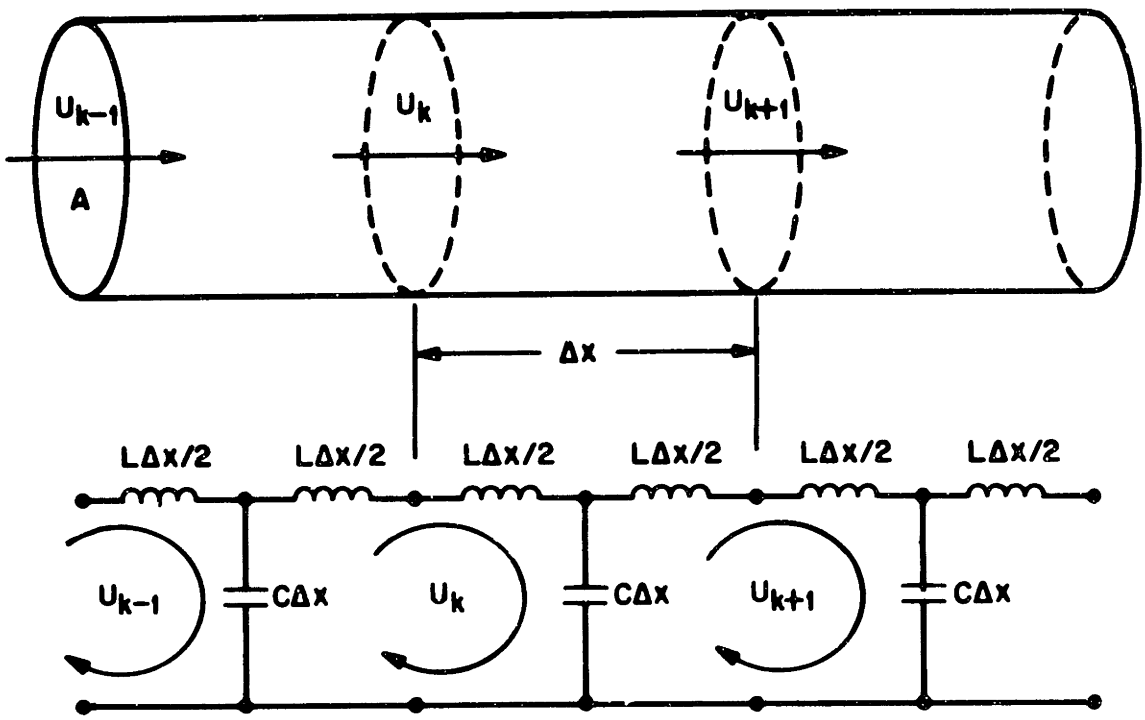
where $\hat{U}_k(s)$ is the complex amplitude (Laplace Transform) of $U_k(t)$. Therefore,

$$U_{k-1}(t) - 2U_k(t) + U_{k+1}(t) = LC\Delta x^2 \frac{d^2U_k(t)}{dt^2} \qquad (1.1)$$

Replacing the time derivative with a backward difference gives

---

* With voltage and current being analogous to acoustic pressure and volume velocity, respectively.



ANALOGOUS CIRCUIT ELEMENTS:

$$L = \rho/A$$
$$C = A/\rho c^2$$

kTH LOOP EQUATION:

$$\hat{U}_{k-1}(s) - 2\hat{U}_k(s) + \hat{U}_{k+1}(s) = LC\Delta x^2 s^2 \hat{U}_k(s)$$

OR

$$U_{k-1}(t) - 2U_k(t) + U_{k+1}(t) = LC\Delta x^2 \frac{d^2U_k(t)}{dt^2}$$

FIGURE 1.3

T-SECTION LUMPED PARAMETER ANALOG FOR LOSSLESS UNIFORM ACOUSTIC TUBE

$$U_{k-1}^{n} - 2U_{k}^{n} + U_{k+1}^{n} = \frac{1}{\eta^2}\left(U_{k}^{n}-2U_{k}^{n-1}+U_{k}^{n-2}\right) \tag{1.2}$$

where

$$U_{k}^{n} = U_{k}(t=n\Delta t)$$

and

$$\frac{1}{\eta^2} = LC\left(\frac{\Delta x}{\Delta t}\right)^2 = \frac{1}{c^2}\left(\frac{\Delta x}{\Delta t}\right)^2$$

Equation (1.2) is the equation which is solved, in an appropriate domain, to obtain the simulation.

We assume a solution of the form

$$U_{k}^{n} = Ae^{j\beta k\Delta x+sn\Delta t}. \tag{1.3}$$

This ansatz is chosen so that any allowed spatial distribution at a particular time $t = n\Delta t$ can be synthesized as a Fourier sum (or integral) over the modes specified by β. Substituting the ansatz in Eq. (1.2) it is clear that it satisfies this equation if and only if β and s are related by the characteristic equation

$$e^{-2s\Delta t} - 2e^{-s\Delta t} + 1 + 4\eta^2 \sin^2\left(\frac{\beta\Delta x}{2}\right) = 0 \tag{1.4}$$

Thus,

$$e^{-s\Delta t} = 1 \pm j2\eta \sin\left(\frac{\beta\Delta x}{2}\right). \tag{1.5}$$

Letting $s = \sigma + j\omega$ and equating the real and imaginary parts of Eq. (1.5) gives

$$\sigma = -\frac{1}{2\Delta t} \ln\left[1+4\eta^2 \sin^2\left(\frac{\beta\Delta x}{2}\right)\right]$$

$$\omega = \pm\frac{1}{\Delta t} \tan^{-1}\left[2\eta \sin\left(\frac{\beta\Delta x}{2}\right)\right] \tag{1.6}$$

Therefore, a particular mode with wave number β decays as $\exp[-\frac{n}{2}\ln(1+4\eta^2 \sin^2\beta\Delta x/2)]$. Since $\mathrm{Re}(s) \leq 0$, Eq. (1.2) is stepwise stable[12] for all values of the sampling ratio (which is proportional to η). In addition, the simulation can be shown to be pointwise stable.[13]

It is well known however, that the volume velocity of an acoustic wave propagating in a lossless uniform hard-walled duct obeys the wave equation

$$\frac{\partial^2 U}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} . \tag{1.7}$$

Consequently, the finite-difference Eq. (1.2) is a simulation of Eq. (1.7) obtained by replacing spatial derivatives by divided central differences and time derivatives by backward differences. If the ansatz

$$U(x,t) = A' e^{j\beta' x + s' t} \tag{1.8}$$

is substituted into the partial-differential Eq. (1.7) we find

$$\omega' = \pm\beta'/c$$

and

$$\sigma' = 0.$$

Thus, the modes of Eq. (1.7) oscillate periodically in time (the physically correct behavior), whereas the modes of the discrete-variable simulation decay exponentially. This phenomenon is called stepwise overstability and the error that it introduces can be quite large.[12] Because of this improper damping, it is expected that a vocal-tract simulation based on "T" sections and backward differences will result in formants with excessive bandwidths, and is therefore undesirable in spite of its stability, and in spite of the fact that the approximate solution will converge to the true solution as the sampling rates are increased. The point of this example is simply that one must use considerable care when modeling a continuous system as a discrete system and it is not sufficient to rely solely on intuitive arguments.

In addition to the philosophical reasons for drawing a distinction between the two stages of the modeling process, there is also a practical reason. Once a particular simulation has been implemented, it is not immediately obvious how to evaluate it, and should it be found inadequate, how to determine at what level the inadequacy originates. By separating the modeling process into two steps, it is possible to investigate the characteristics of each step independently. The adequacy of the partial-differential equations to describe the physical system may be investigated by using very accurate, albeit inefficient, numerical techniques, or by transforming the equations to the frequency domain and investigating the

behavior of the eigenmodes of the system. The properties of a particular finite-difference formulation, on the other hand, may be investigated from a numerical analysis or digital-signal-processing viewpoint, independently from the specific physical model. In fact, from such an analysis it is possible to predict the effects of a particular simulation on the physical model. An illustration of this kind of analysis is the previous discussion of an acoustic tube modeled as a cascade of "T" networks, though a more general approach is suggested in Appendix A.

This thesis presents the formulation of a vocal-tract simulation which models certain aspects of the speech-production process. Specifically, the transmission of sound in the vocal tract is modeled as quasi-one-dimensional* acoustic wave propagation in a nonuniform time-varying tube with yielding walls. The vocal tract is excited by, and interacts with, a self-oscillating two-mass model of the vocal cords.[11] The vocal tract is terminated by a radiation load based on a simplified model of a piston in a baffle.

In addition to the obvious objective of ultimately realizing a carefully formulated and accurate digital simulation of the vocal tract, this thesis has several other important

---

* The term quasi-one-dimensional is used to emphasize that acoustic wave propagation in a duct with yielding walls cannot be strictly one dimensional, since the wall motion introduces a pressure gradient in the transverse direction. However, when the wavelength of this transverse mode is large compared with the diameter of the duct, the transverse pressure gradient may be treated as a second-order effect.

objectives. The formulation of the mathematical model for acoustic transmission in the vocal tract begins with fundamental physical laws and is developed in such a manner that each of the underlying assumptions and approximations introduced is clearly stated. Moreover, in developing from this model a discrete-variable simulation, suitable for implementation on a digital computer, techniques and insights drawn from digital signal processing and numerical analysis prove valuable and provide novel approaches to some of the modeling problems. It is felt that some of the approaches suggested by these techniques might be applied profitably to other simulation problems as well.

In Chapter II, a mathematical description of acoustic wave propagation in the vocal tract is formulated. This description is a set of partial-differential equations derived from fundamental physical principles. The simplifying assumptions invoked in the derivation, together with the conditions for which they are valid, are clearly stated. In addition, an equation of motion describing the wall vibration is derived, assuming small oscillations about equilibrium.

In Chapter III, a mathematical description of the effects of radiation at the lips and nostrils is given. The model is based on a simplification of the well known radiation impedance for a piston in baffle.[14] The model leads to an algebraic equation in the frequency domain suitable as a boundary condition for the frequency-domain solution of the vocal-tract equations (eigenvalue problem) and also leads to

an ordinary differential equation suitable as a boundary condition for a time-domain solution.

In Chapter IV, the partial-differential equations derived in Chapter II are transformed to ordinary differential equations in the (temporal) frequency domain. A two-point boundary-value problem is formulated assuming an ideal current source at the glottis. The problem is solved numerically and the glottis-to-mouth transfer ratio Um/Ug is computed. Furthermore, this formulation is of interest for it allows the investigation of the behavior of the eigenmodes of the vocal tract.

In Chapter V, the partial-differential equations derived in Chapter II are written as partial-difference equations in discrete variables. The formulation chosen is an implicit formulation equivalent to integrating the equations by the trapezoid rule in both the time and space directions. The equation of motion for the vibrating duct wall is transformed to a difference equation using the technique of impulse invariance.[15]

The first part of Chapter VI discusses several simulations of the radiation load. The radiation impedance is calculated and plotted for each of the simulations and it is apparent that the simulation corresponding to integration by the trapezoid rule is superior. The remainder of the chapter gives a brief description of the two-mass model of the vocal cords recently described by Ishizaka and Flanagan.[11]

In Chapter VII, a complete simulation is discussed. The system of equations for sound transmission in the vocal tract is written compactly in matrix form and a procedure is given for solving these equations simultaneously with the equations for the glottal system, given by Ishizaka and Flanagan.

In Chapter VIII, some experimental results are presented, the significant results of the present study are summarized, and some areas for further study are suggested.

In addition to the main body of the thesis there are two appendices which discuss topics of global significance. Appendix A presents an interpretation of numerical integration and differentiation rules as transformations between frequency spaces. This point of view is profitable because it often allows one to predict the implications of a particular simulation in advance and provides an intuitive picture of distortions that might be introduced by the simulation. Appendix B presents an efficient algorithm for solving the complete system of algebraic equations for the simulation when both boundary conditions are explicitly given.*

---

* When one of the boundary conditions is not explicitly known (as with the two-mass model or at a junction of three ducts) a modification of the procedure is given in Chapter VII which permits the solution of several systems simultaneously.

CHAPTER II

The Equations of Motion for Acoustic Wave Propagation in the Vocal Tract

The equations of motion for acoustic wave propagation in the vocal tract will now be derived from the fundamental physical principles of conservation of mass and conservation of momentum. The vocal tract is modeled as a nonuniform time-varying duct with yielding walls and the wave motion is considered as essentially plane-wave propagation with perturbations introduced by the yielding walls.

The quasi-one-dimensional model is justifiable for several reasons. It is a universally used model and is valid so long as the wavelength of sound is large compared with the transverse dimensions of the vocal tract.[16-21] The validity of this assumption does, however, become borderline at the highest audible frequencies. A one-dimensional model is desirable in order to keep the equations simple enough to be used in a simulation and the question remains to be answered: just how accurately can observed phenomena be accounted for with such a model? By far the best justification of the one-dimensional model is the recent work by Lesser and Lewis.[22] They have shown, using modern perturbation theory in the form of matched asymptotic expansions, that the

first-order correction term to the Webster Horn equation* at an abrupt change in the duct cross section is very small. This means that the one-dimensional model is valid over a wider range of duct cross sections than would be expected from the classical theory.

The internal energy losses in the fluid, due to viscosity and thermal conduction, will also be neglected. The losses in the bulk of the medium are negligible because of the small coefficient of viscosity and thermal conductivity of the air in the vocal tract. The losses in the boundary layer at the duct wall can be shown to vary inversely as the three-halves power of duct's cross-sectional area and directly with the square root of the sound frequency.[14,19] In the case of laminar (nonturbulent) flow, for the cross-sectional areas and frequencies of interest, the effects of the boundary-layer losses are small when compared with the losses introduced by the yielding walls and radiation impedance. If these losses are included, the equations of motion are modified by the addition of a frequency dependent resistive term.[14,19]

The yielding wall of the duct is considered as locally reacting. That is, adjacent surface elements are

---

* The Webster Horn equation:

$$\frac{1}{A(x)} \frac{\partial}{\partial x} \left[A(x) \frac{\partial p}{\partial x} \right] = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}$$

describes an acoustic pressure wave $p(x,t)$ propagating in a loss free horn with rigid walls and cross-sectional area $A(x)$.[23]

not strongly coupled and hence there is no wave propagation in the wall itself. The wall motion not only introduces loss due to the real part of the (mechanical) wall impedance, but results in dispersive propagation of the acoustic wave and allows for pressure build-up and vocal-cord vibration with the lips closed.

A sufficient set of quantities for describing acoustic wave propagation in a fluid such as air is the density ρ, the pressure P, the temperature T, the entropy S, and the local or particle velocity $\vec{V}$ of the fluid.[14] These quantities are interdependent and related by equations of state. Since several approximations are necessary to model the vocal tract as an essentially one-dimensional system, the equations will be derived in their general form, as a system of nonlinear partial-differential equations in three spatial dimensions, and then reduced to a one-dimensional system in such a manner that each approximation and its rationale is clearly stated. In this way, should the present model prove inadequate, or a more detailed model be desired, it is only necessary to return to the derivation and relax the appropriate assumptions. The general nonlinear set of state equations in three dimensions will be derived first, then linearized about the equilibrium values of the state variables, and finally applied to a non-uniform time-varying acoustic tube.

Using Newton's Second Law ($\vec{F}=m\vec{a}$), we can write the equation for conservation of fluid momentum. We must be careful, however, to write this equation in the rest frame

of the fluid, and then properly transform it to the fixed or "lab" frame. Consider a differential mass element of the fluid $dm = \rho d\tau$ ($d\tau$ represents a differential volume element). This element may deform with time, but its mass remains constant. If the mass element is acted upon by a force $\vec{F} = -\vec{\nabla}P d\tau$ and its velocity is $\vec{V}$, then its momentum is $\rho\vec{V}d\tau$ and Newton's Second Law gives

$$-\vec{\nabla}P d\tau = \frac{D}{Dt}(\rho\vec{V}d\tau) \tag{2.1}$$

where $\frac{D}{Dt}$ denotes the "convective-derivative" and indicates that the quantity being differentiated is measured in the "proper" or rest frame of the fluid.[24] Since the mass dm, by definition, remains constant with time,

$$\frac{D}{Dt}(\rho d\tau) = 0$$

and Eq. (2.1) becomes

$$-\vec{\nabla}P d\tau = \rho \frac{D\vec{V}}{Dt} d\tau$$

hence

$$-\vec{\nabla}P = \rho \frac{D\vec{V}}{Dt} \tag{2.2}$$

Finally, since the operator $\frac{D}{Dt}$ transforms to the lab frame as $\frac{\partial}{\partial t} + \vec{V}\cdot\vec{\nabla}$, the equation for conservation of fluid momentum is

$$-\vec{\nabla}P = \rho \frac{\partial\vec{V}}{\partial t} + \rho\vec{V}\cdot\vec{\nabla}\vec{V} \tag{2.3}$$

The second equation of motion is derived from the conservation of total fluid mass and will be written in integral form. Consider a deformable macroscopic region of fluid τ bounded by the closed surface $\vec{\partial\tau}$. Denote a volume element in this region by $d\tau$ and a surface element of the boundary by $d\vec{\sigma} = \hat{n}d\sigma$ defined by the outward pointing unit normal vector $\hat{n}$. The total mass contained in τ is given by

$$\begin{matrix}\text{total mass} \\ \text{in volume } \tau\end{matrix} = \iiint_{\tau} \rho d\tau$$

and the net rate of flow of mass into τ by

$$\begin{matrix}\text{rate of mass} \\ \text{flow into } \tau\end{matrix} = -\oiint_{\vec{\partial\tau}} \rho\vec{V}\cdot d\vec{\sigma}$$

If there are no sources or sinks in τ, then the net rate of flow of mass into τ must be equal to the time rate of increase of mass contained in τ, thus

$$-\oiint_{\vec{\partial\tau}} \rho\vec{V}\cdot d\vec{\sigma} = \frac{d}{dt}\iiint_{\tau} \rho d\tau \qquad (2.4)$$

This is the equation of continuity, in integral form, for the fluid mass.

Equations (2.3) and (2.4) are nonlinear equations in the quantities P, $\vec{V}$ and ρ. We now seek and third relation among these quantities and linearize the resulting set of equations about the equilibrium values of the variables. Assume that with no passage of sound the fluid is in

equilibrium and denote by P_o, $\vec{V}_o$, and ρ_o the equilibrium values of the pressure, velocity, and density of the fluid. Now, the propagation of sound through the fluid causes small perturbations in these quantities (denoted by p, $\vec{v}$, ρ') so that we may write

$$P = P_o + p$$

$$\vec{V} = \vec{V}_o + \vec{v}$$

and

$$\rho = \rho_o + \rho'$$

For sound frequencies below about 10^9 Hz, the compression of a gas due to the passage of a sound wave is approximately <u>adiabatic</u> i.e., the entropy content of the gas remains nearly constant during the compression.[14] Thus we may write

$$\rho' = \rho_o \kappa_s p$$

and

$$\rho = \rho_o + \rho_o \kappa_s p \qquad (2.5)$$

where κ_s is the adiabatic compressibility of the fluid.

We now substitute $P = P_o + p$, $\vec{V} = \vec{V}_o + \vec{v}$, and $\rho = \rho_o(1+\kappa_s p)$ into Eqs. (2.3) and (2.4) neglect all products of the perturbations, and observe that the equilibrium values must satisfy the appropriate equations (2.3 and 2.4) at

equilibrium. Equation (2.3) becomes

$$-\vec{\nabla}p = \rho_o \frac{\partial \vec{v}}{\partial t} + \rho_o \vec{V}_o \cdot \vec{\nabla}\vec{v} \tag{2.6}$$

and Eq. (2.4) becomes

$$-\oiint_{\partial\tau} \rho_o(\vec{v}+\kappa_s p\vec{V}_o)\cdot d\vec{\sigma} = \frac{d}{dt}\iiint_{\tau} \rho_o(1+\kappa_s p)d\tau$$

or since ρ_o is a constant

$$-\oiint_{\partial\tau} (\vec{v}+\kappa_s p\vec{V}_o)\cdot d\vec{\sigma} = \frac{d}{dt}\iiint_{\tau} (1+\kappa_s p)d\tau \tag{2.7}$$

Equations (2.6) and (2.7) will now be applied to a nonuniform time-varying duct with yielding walls. Consider the incremental length Δx of duct shown in Fig. 2.1. The surface corresponding to the duct wall, between parallel planes at x and x + Δx, is denoted by Σ and the cross-sectional area at x is denoted by A(x). Although A may be time varying, only the spatial dependency is shown explicitly here. Applying the conservation of mass equation (2.7) gives

$$-\left\{\iint_{A(x+\Delta x)} [v_x+\kappa_s pV_{ox}]dA - \iint_{A(x)} [v_x+\kappa_s pV_{ox}]dA \right.$$

$$\left. +\iint_{\Sigma} \rho\vec{v}\cdot\hat{n}_\Sigma d\Sigma\right\} = \frac{d}{dt}\int_x^{x+\Delta x} dx \iint_{A(x)} (1+\kappa_s p)dA \tag{2.8}$$

Since no mass passes through the wall of the duct, the last integral (over the surface Σ) on the left-hand side of Eq. (2.8) vanishes. Dividing both sides of the equation

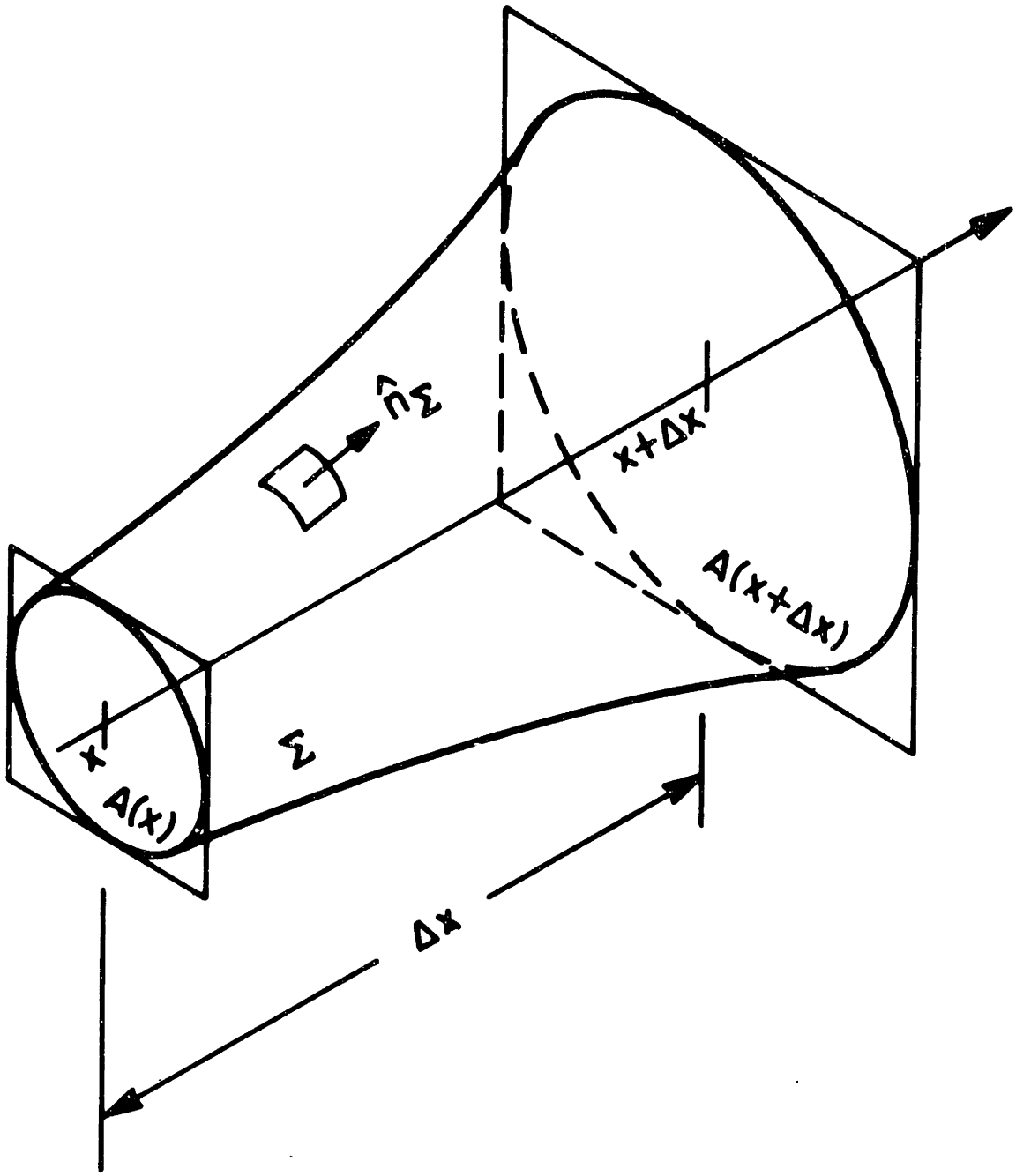


FIGURE 2.1
INCREMENTAL LENGTH Δx OF DUCT

by Δx and taking the limit as $\Delta x \to 0$ gives

$$- \lim_{\Delta x \to 0} \frac{1}{\Delta x} \left\{ \iint_{A(x+\Delta x)} [v_x + \kappa_s p V_{ox}] dA - \iint_{A(x)} [v_x + \kappa_s p V_{ox}] dA \right\}$$

$$= \lim_{\Delta x \to 0} \frac{1}{\Delta x} \frac{d}{dt} \int_x^{x+\Delta x} dx \iint_{A(x)} (1 + \kappa_s p) dA \tag{2.9}$$

Applying the definition of the derivative to the left-hand side and the Fundamental Theorem of Integral Calculus to the right-hand side of Eq. (2.9) gives*

$$- \frac{\partial}{\partial x} \iint_{A(x)} [v_x + \kappa_s p V_{ox}] dA = \frac{\partial}{\partial t} \iint_{A(x)} (1 + \kappa_s p) dA \tag{2.10}$$

Defining the volume velocity at x as

$$U(x,t) = \iint_{A(x)} v_x dA$$

and approximating

$$\iint_{A(x)} p dA \approx pA$$

gives

$$- \frac{\partial U}{\partial x} - \kappa_s V_{ox} \frac{\partial (pA)}{\partial x} = \frac{\partial A}{\partial t} + \kappa_s \frac{\partial (pA)}{\partial t} \tag{2.11}$$

---

* Note that both derivative operators are partial-derivatives because the functions on which they operate are functions of both x and t.

The conservation of momentum equation is applied simply as follows. We neglect, to first order, the effects of the radial and circumferential components of the pressure gradient on the fluid itself. Thus, the laminar (x) component of Eq. (2.6) gives

$$-\frac{\partial p}{\partial x} = \rho_o \frac{\partial}{\partial t}\left(\frac{U}{A}\right) + \rho_o V_{ox} \frac{\partial}{\partial x}\left(\frac{U}{A}\right) \tag{2.12}$$

with the approximation

$$v_x \approx \frac{U}{A}$$

Equations (2.11) and (2.12) both contain the quantity A which is <u>not</u> known a priori. Although the cross-sectional area may be specified at equilibrium (no sound propagation), the change in fluid pressure at the duct wall accompanying an acoustic wave will cause a small change in the cross-sectional area if the wall admittance is nonzero. Treating this small change as a linear perturbation, we write

$$A = A_o + \delta A$$

where A_o is the equilibrium value of the cross-sectional area (thus A_o is the parameter of the model which we are free to specify). Since $\delta A << A_o$, we write

$$\begin{aligned}\frac{U}{A} &= \frac{U}{A_o + \delta A} \\ &= \frac{U}{A_o}\left[1 - \frac{\delta A}{A_o} + \left(\frac{\delta A}{A_o}\right)^2 - \cdots\right] \\ &\approx \frac{U}{A_o}\end{aligned}$$

and

$$pA = p(A_o+\delta A)$$

$$\approx pA_o$$

so that Eqs. (2.11) and (2.12) become, respectively

$$-\frac{\partial U}{\partial x} - \kappa_s V_{ox} \frac{\partial (pA_o)}{\partial x} = \frac{\partial (A_o+\delta A)}{\partial t} + \kappa_s \frac{\partial (pA_o)}{\partial t} \qquad (2.13)$$

and

$$-\frac{\partial p}{\partial x} = \rho_o \frac{\partial}{\partial t}\left(\frac{U}{A_o}\right) + \rho_o V_{ox} \frac{\partial}{\partial x}\left(\frac{U}{A_o}\right) \qquad (2.14)$$

We now proceed to determine the perturbation δA in terms of the fluid pressure at the duct wall. Referring to Fig. 2.2, denote by $\xi(x,t)$ the normal displacement of the duct wall from its equilibrium position. If $S_o(x,t)$ denotes the duct perimeter at equilibrium ($\xi=0$) then the approximate change in area, δA, due to the displacement ξ is

$$\delta A = S_o \xi \qquad (2.15)$$

The wall admittance is modeled as a spring force constant, $K_w(x)$, per unit surface area, a damping constant, $b_w(x)$, per unit surface area, and a mass, $M_w(x)$, per unit surface area. Referring to Fig. 2.3, the equation of motion for small oscillations about equilibrium for a differential element of wall area, $d\Sigma$, is

$$pd\Sigma - (K_w d\Sigma)\xi - (b_w d\Sigma)\dot{\xi} = (m_w d\Sigma)\ddot{\xi}$$

$$A(x,t) = A_0(x,t) + \delta A(x,t)$$
$$= A_0(x,t) + S_0(x,t)\xi(x,t)$$

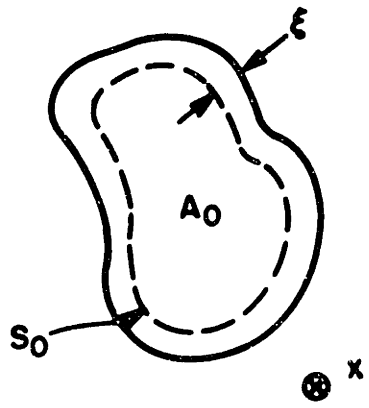


FIGURE 2.2
CROSS SECTION OF TIME-VARYING DUCT

$$p = M_w\ddot{\xi} + b_w\dot{\xi} + K_w\xi$$

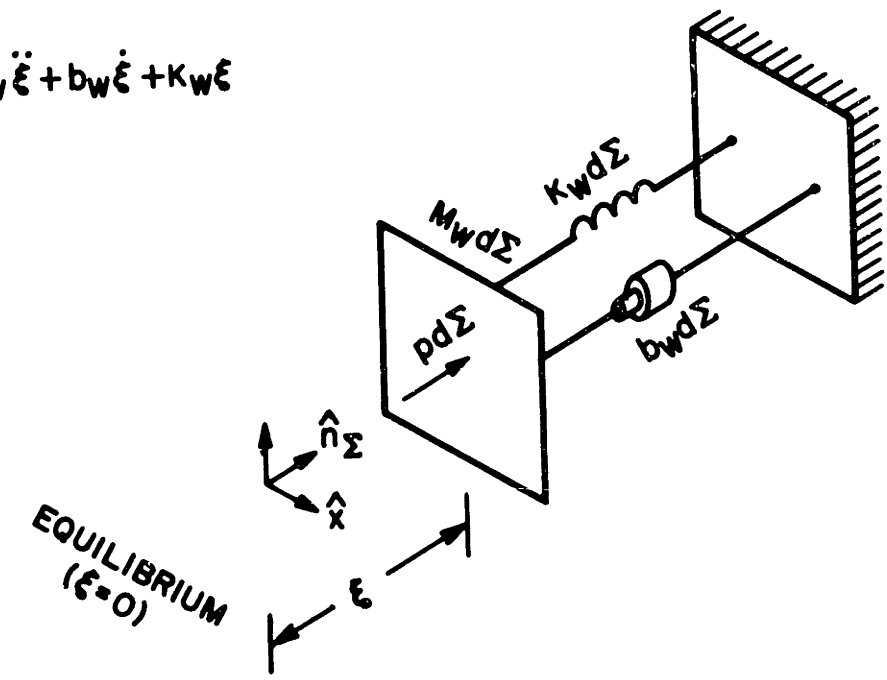


FIGURE 2.3
MECHANICAL MODEL FOR DIFFERENTIAL SURFACE ELEMENT $d\Sigma$ OF DUCT WALL

or

$$p = M_w\ddot{\xi} + b_w\dot{\xi} + k_w\xi \tag{2.16}$$

The solution to Eq. (2.16) may be expressed as the convolution integral

$$\xi(x,t) = \int_{-\infty}^{t} p(x,\alpha)h(x,t-\alpha)d\alpha \tag{2.17}$$

where $h(x,t)$ is the impulse response of Eq. (2.16) at the position x along the duct.

It can now be argued that the convective terms of the form $V_{ox}\frac{\partial}{\partial x}$ in Eqs. (2.13) and (2.14), may be neglected, to a good approximation, when the D.C. flow velocity, V_{ox}, is small compared with the velocity of sound, c. We know that in the case of a duct with rigid walls and piecewise constant cross-sectional area, when there is no D.C. flow, the system of Eqs. (2.13) and (2.14) is satisfied by a superposition of traveling waves in each section of the form

$$p(x,t) = p_+(x-ct) + p_-(x+ct)$$

and

$$U(x,t) = U_+(x-ct) + U_-(x+ct).$$

Observing that for any function of the form $f(x\pm ct)$

$$\frac{\partial f(x\pm ct)}{\partial t} = \pm cf'(x\pm ct)$$

whereas

$$V_{ox} \frac{\partial f(x \pm ct)}{\partial x} = V_{ox} f'(x \pm ct)$$

it can be seen that the error introduced by neglecting the convective terms is small when $V_{ox} << c$. It is expected that the introduction of yielding walls and nonuniform cross-sectional area should not affect the solution so as to make this assumption unreasonable.

The final set of equations for sound propagation in the vocal tract is then the following:

$$- \frac{\partial p}{\partial x} = \rho_o \frac{\partial}{\partial t} \left(\frac{U}{A_o} \right) \qquad \text{(conservation of momentum)} \qquad (2.18)$$

$$- \frac{\partial U}{\partial x} = \kappa_s \frac{\partial (pA_o)}{\partial t} + \frac{\partial A}{\partial t} \qquad \text{(conservation of mass)} \qquad (2.19)$$

where

$$A(x,t) = A_o(x,t) + S_o(x,t)\xi(x,t) \qquad (2.20)$$

and $\xi(x,t)$ satisfies the differential equation

$$p(x,t) = M_w(x)\ddot{\xi}(x,t) + b_w\dot{\xi}(x,t) + K_w(x)\xi(x,t) \qquad (2.21)$$

The parameters which msut be specified are:

$A_o(x,t)$ the cross-sectional area of the vocal tract (at equilibrium)

$S_o(x,t)$ the perimeter of the vocal tract (at equilibrium)

$M_w(x)$ the surface density or mass/unit area of the vocal-tract wall

$b_w(x)$ the damping/unit area of the vocal-tract wall

$K_w(x)$ the stiffness or spring force constant/unit area of the vocal-tract wall

ρ_o the air density[17] $= 1.14\times10^{-3}$ gm/cm^3 (moist air at body temperature 37°C)

κ_s the adiabatic compressibility of air which will be shown to be given by

$$\kappa_s = \frac{1}{\rho_o c^2}$$

where $c = 3.5\times10^4$ cm/sec, is the speed of sound in moist air at body temperature, 37°C.

It will now be shown that if the duct wall is rigid and the cross-sectional area is time invariant, then the system of Eqs. (2.18) and (2.19) reduces to the familiar Webster Horn Equation.[23] Since $\frac{\partial A}{\partial t} = 0$ we can write

$$-\frac{\partial p}{\partial x} = \frac{\rho_o}{A(x)}\frac{\partial U}{\partial t} \tag{2.22}$$

and

$$-\frac{\partial U}{\partial x} = \kappa_s A(x)\frac{\partial p}{\partial t} \tag{2.23}$$

Multiplying Eq. (2.22) by $A(x)$ and differentiating with respect to x gives

$$-\frac{\partial}{\partial x}\left[A(x)\frac{\partial p}{\partial x}\right] = \rho_o \frac{\partial^2 U}{\partial x \partial t} \tag{2.24}$$

Differentiating Eq. (2.23) with respect to t gives

$$-\frac{\partial^2 U}{\partial t \partial x} = \kappa_s A(x) \frac{\partial^2 p}{\partial t^2} \tag{2.25}$$

Finally, eliminating the mixed partial derivative, $\frac{\partial^2 U}{\partial x \partial t}$, in Eqs. (2.24) and (2.25) gives

$$\frac{1}{A(x)} \frac{\partial}{\partial x}\left[A(x) \frac{\partial p}{\partial x}\right] = \kappa_s \rho_o \frac{\partial^2 p}{\partial t^2} \tag{2.26}$$

which is recognized as the Webster Horn equation with

$$\frac{1}{c^2} = \kappa_s \rho_o \tag{2.27}$$

where c is the velocity of sound in the horn.

In this chapter we have derived a set of linear partial-differential equations describing quasi-one-dimensional acoustic wave propagation in a nonuniform time-varying duct with yielding walls. These equations will be used to simulate sound propagation in the vocal tract under the assumptions that

1. the transverse modes of propagation are small (valid so long as the wavelength is larger than the duct cross section);
2. the acoustic wave may be treated as a linear perturbation of the acoustic medium (valid as long as amplitude of the pressure wave is small compared with the ambient pressure);
3. the D.C. velocity is small compared with the speed of sound;

4. the losses due to the viscosity and thermal conductivity in the medium bulk and boundary layer are small, especially when compared with loss due to the glottal impedance, radiation impedance, and finite wall impedance (valid for laminar flow, sound frequencies less than about 10^9 Hz, and cross-sectional areas greater than about 1 cm^2);
5. the vocal-tract cross section doesn't change "too rapidly" in the x direction (recent work by Lesser and Lewis indicates this approximation is better then expected from classical theory);
6. the elements of the vocal-tract wall are locally reacting i.e., adjacent elements are not coupled;
7. the vibration of the vocal-tract wall may be treated as "small oscillations" about equilibrium.

CHAPTER III

The Radiation Load at the Mouth and Nostrils

The partial-differential equations derived in Chapter II describe the vocal tract as an acoustic transmission line. It is a well known result from the study of boundary-value problems that the eigenmodes of such a system are determined by both a differential (or integral) equation describing the system and the specification of a suitable set of boundary conditions. In this chapter we derive the boundary condition corresponding to the radiation load at the mouth or nostrils.

The mechanism whereby acoustic energy is radiated by the articulatory system and diffracted about the head is indeed complicated. However, we must remember that the eignemodes of the vocal tract are determined by the impedance "seen" by the vocal tract looking toward the outside world. Thus, a reasonable model for the radiation load at the mouth (or nostrils), consistent with the one-dimensional assumption of Chapter II, is the radiation impedance for a piston set in the side of a rigid sphere. The expression for this impedance is derived by Morse and Ingard[14] as an infinite series of various spherical harmonics, and cannot be written in closed form. In a word, even this simplified model is algebraically intractable.

As a second approximation, we assume the diameter of the piston is small compared to the diameter of the sphere.

In this case, we choose for the radiation load the radiation impedance for a piston set in an infinite plane baffle which is expressible in closed form.[14] A comparison between the radiation impedances for a piston in a spherical baffle and a piston in a plane baffle shows that the change from a spherical to a plane baffle makes little change in the average impedance load on the piston, although the radiation directionality patterns differ considerably.[14] The radiation impedance for a piston set in an infinite plane baffle, expressed as a normalized acoustic impedance $z_n = Z_A/Z_o = \left(\frac{p}{U}\right)/\left(\frac{\rho_o c}{A}\right)$, is given by

$$Z_n = \theta_o + jX_o$$

$$\theta_o = 1 - \frac{1}{\mu} J_1(2\mu)$$

$$X_o = \int_o^{\pi/2} \sin(2\mu\cos\alpha)\sin^2\alpha d\alpha = \frac{1}{\mu} H_1(2\mu) \qquad (3.1)$$

where $\mu = ka = \frac{\omega a}{c}$, $k = \frac{2\pi}{\lambda} = \frac{\omega}{c}$ is the wavenumber, a is the radius of the piston, A is the area of the piston, $J_1(2\mu)$ is the first order Bessel function[24] of the first kind given by

$$J_1(2\mu) = \sum_{m=0}^{\infty} \frac{(-1)^m \mu^{2m+1}}{m!(m+1)!} \qquad (3.2)$$

and $H_1(2\mu)$ is the first order Struve function[25] given by*

* $(2m+1)!! = (2m+1)(2m-1) \ldots 5\cdot3\cdot1$

$$H_1(2\mu) = \frac{2}{\pi} \sum_{m=0}^{\infty} \frac{(-1)^m (2\mu)^{2m+2}}{(2m+1)!!(2m+3)!!} \tag{3.3}$$

We can thus write

$$\theta_o = \sum_{m=0}^{\infty} \frac{(-1)^m \mu^{2m+2}}{m!(m+1)!}$$

$$= \frac{\mu^2}{2!} - \frac{\mu^4}{2!3!} + \frac{\mu^6}{3!4!} - \cdots \tag{3.4}$$

and

$$\chi_o = \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{(-1)^m (2\mu)^{2m+1}}{(2m+1)!!(2m+3)!!}$$

$$= \frac{4}{\pi} \left[\frac{2\mu}{3} - \frac{(2\mu)^3}{3^2 \cdot 5} + \frac{(2\mu)^5}{3^2 \cdot 5^2 \cdot 7} - \cdots \right] \tag{3.5}$$

Now consider an approximation to z_n given by

$$\tilde{z}_n = \frac{j\alpha\mu}{1+j\beta\mu} \tag{3.6}$$

where α and β are constants to be determined. If both numerator and denominator of Eq. (3.6) are multiplied by $1-j\beta\mu$ and the result expanded as a Taylor series in the variable $(\beta\mu)^2$ we have

$$\tilde{z}_n = \frac{j\alpha\mu(1-j\beta\mu)}{1+(\beta\mu)^2}$$

$$= (\alpha\beta\mu^2 + j\alpha\mu)[1-(\beta\mu)^2+(\beta\mu)^4-(\beta\mu)^6+\cdots]. \tag{3.7}$$

Writing

$$\tilde{z}_n = \tilde{\theta}_o + j\tilde{\chi}_o \tag{3.8}$$

gives

$$\tilde{\theta}_o = \alpha\beta[\mu^2 - \beta^2\mu^4 + \beta^4\mu^6 - \ldots] \tag{3.9}$$

and

$$\tilde{\chi}_o = \alpha[\mu - \beta^2\mu^3 + \beta^4\mu^5 - \ldots] \tag{3.10}$$

If we set $\alpha = 8/3\pi$ and $\beta = 3\pi/16$, then the lowest order terms in the expansions for $\tilde{\theta}_o$ and $\tilde{\chi}_o$ correspond with those of θ_o and χ_o. Since $\mu = ka = 2\pi a/\lambda$, the higher order terms in the expansions (3.4, 3.5, 3.9, and 3.10) become decreasingly small under our assumption that $a << \lambda$. Moreover, the higher order terms of $\tilde{\theta}_o$ and $\tilde{\chi}_o$ have the same exponential dependence as those of θ_o and χ_o.

As a simplified model for the radiation load we use the analogous impedance derived from $\tilde{z}_n$, that is,

$$Z_A = \frac{p}{U} = \left(\frac{\rho_o c}{A}\right)\tilde{z}_n = \left(\frac{\rho_o c}{A}\right)\frac{j\,\frac{8}{3\pi}\,\mu}{1+j\,\frac{3\pi}{16}\,\mu} \tag{3.11}$$

Substituting $\mu = \frac{\omega a}{c}$, gives

$$Z_A = \frac{p}{U} = \left(\frac{\rho_o c}{A}\right)\frac{j\omega\left(\frac{8a}{3\pi c}\right)}{1+j\omega\left(\frac{3\pi a}{16c}\right)} . \tag{3.12}$$

Observing that the impedance of a resistance and inductance in parallel is

$$Z = \frac{j\omega L}{1+j\omega \frac{L}{R}} , \tag{3.13}$$

the normalized acoustic impedance, $\tilde{z}_n$, associated with Eq. (3.12) may be interpreted as an analogous inductance of $8a/3\pi c$ in parallel with an analogous resistance of $128/9\pi^2$.

In order to be used as boundary condition for the wave equations derived in Chapter II, Eq. (3.12) must be transformed to the time domain. Using the usual analogy of acoustic pressure to electrical voltage and acoustic volume velocity to electric current, Eq. (3.12) may be written in the time domain as the integral equation

$$U(t) = \frac{A}{\rho_o c}\left[\frac{1}{L_{rad}}\int_{-\infty}^{t} p(\alpha)d\alpha + \frac{p(t)}{R_{rad}}\right] \tag{3.14}$$

where

$$R_{rad} = \frac{128}{9\pi^2}$$

$$L_{rad} = \frac{8a}{3\pi c}$$

Equation (3.14) is then the desired boundary condition for the radiation load at the mouth or nostrils.

CHAPTER IV

Frequency-Domain Solution of the Vocal-Tract Transmission Equations

If the vocal-tract configuration is constrained to be static, then the steady-state solution to the vocal-tract transmission equations (2.18 - 2.21) can be obtained by transforming the equations to the frequency domain. These equations, when expressed in the frequency domain, become a system of first-order ordinary differential equations; and together with the appropriate boundary conditions define a two-point boundary-value problem.[26,27]

The frequency-domain solution of the vocal-tract transmission equations is of interest for several reasons. Since the solution to the equations in the frequency domain is the steady-state acoustic pressure and volume-velocity distribution for a complex exponential excitation, we can observe these distributions for arbitrary frequencies. We can also compute such quantities as the driving-point impedance of the vocal tract and the glottis-to-mouth transfer ratio. Not only are such data of interest themselves, but they can also be compared with spectral data obtained from real speech and thus provide an additional basis for evaluating the vocal-tract model. Finally, since the viscous and thermal boundary-layer losses are expressed simply in the validity of neglecting these losses may be checked.

The vocal-tract transmission equations are transformed to the frequency domain by assuming a steady-state solution of the form

$$p(x,t) = \hat{p}(x,\omega)e^{j\omega t}$$

$$U(x,t) = \hat{U}(x,\omega)e^{j\omega t}$$

$$\xi(x,t) = \hat{\xi}(x,\omega)e^{j\omega t} \tag{4.1}$$

and substituting this into Eqs. (2.18) through (2.21). The specified cross-sectional area A_o and perimeter S_o are assumed to be time invariant, although the duct wall may vibrate. This gives

$$-\frac{d\hat{p}}{dx} = j\omega\rho_o \frac{\hat{U}}{A_o} \tag{4.2}$$

$$-\frac{d\hat{U}}{dx} = j\omega\kappa_s A_o \hat{p} + j\omega S_o \hat{\xi} \tag{4.3}$$

$$\hat{p} = -\omega^2 M_w \hat{\xi} + j\omega b_w \hat{\xi} + K_w \hat{\xi} \tag{4.4}$$

Solving Eq. (4.4) for $\hat{\xi}$ gives

$$\hat{\xi} = \frac{\hat{p}}{(K_w - \omega^2 M_w) + j\omega b_w}$$

$$= \left[\frac{(K_w - \omega^2 M_w) - j\omega b_w}{(K_w - \omega^2 M_w)^2 + \omega^2 b_w^2}\right] \hat{p}. \tag{4.5}$$

Now $\hat{\xi}$ may be eliminated from the equations by substituting Eq. (4.5) for $\hat{\xi}$ in Eq. (4.3) so that

$$-\frac{d\hat{U}}{dx} = j\omega\hat{p}\left[\kappa_s A_o + \frac{(K_w-\omega^2 M_w) - j\omega b_w}{(K_w-\omega^2 M_w)^2 + \omega^2 b_w^2} S_o\right] \tag{4.6}$$

Finally, we have

$$\frac{d\hat{p}}{dx} + Z\hat{U} = 0 \tag{4.7}$$

$$\frac{d\hat{U}}{dx} + Y\hat{p} = 0 \tag{4.8}$$

where

$$Z = j\omega\rho_o/A_o \tag{4.9}$$

and

$$Y = j\omega\left[\kappa_s A_o + \frac{(K_w-\omega^2 M_w) - j\omega b_w}{(K_w-\omega^2 M_w)^2 + \omega^2 b_w^2} S_o\right] \tag{4.10}$$

(Note that Z and Y are functions of both ω and x.)

In order to assess the effects of the viscous and thermal boundary-layer losses which were neglected in the time-domain formulation, Eqs. (4.7) and (4.8) may be modified to include these losses. The frequency-domain expressions for these losses are well known and derived by Flanagan.[19] The effects of viscous drag at the duct wall can be accounted for by the addition of an analogous resistive term in Eqs. (4.7) and the effects of thermal conduction by the addition of a analogous conductive term in Eq. (4.8). The modified set of equations is:

$$\frac{d\hat{p}}{dx} + Z\hat{U} + R_a\hat{U} = 0 \tag{4.11}$$

$$\frac{d\hat{U}}{dx} + Y\hat{p} + G_a\hat{p} = 0 \tag{4.12}$$

where

$$R_a = \frac{S}{A^2}\sqrt{\frac{\omega\rho_o\mu}{2}} \tag{4.13}$$

and

$$G_a = S\,\frac{\eta-1}{\rho_o c^2}\sqrt{\frac{\lambda\omega}{2C_p\rho_o}} \tag{4.14}$$

μ is the coefficient of viscosity of the fluid, λ is the coefficient of heat conduction of the fluid. η is the adiabatic constant of the fluid and C_p is the heat capacity of the fluid at constant pressure.

The frequency-domain expression for the radiation load was derived in Chapter III; it is

$$Z_{rad} = \frac{\hat{p}}{\hat{U}} = \frac{\rho_o c}{A}\,\frac{j\omega L_{rad}}{1+j\omega L_{rad}/R_{rad}} \tag{4.15}$$

where

$$R_{rad} = 128/9\pi^2$$

$$L_{rad} = 8a/3\pi c$$

$$A = \pi a^2$$

Equation (4.15) implies the homogeneous condition

$$\hat{p} - Z_{rad}\hat{U} = 0. \tag{4.16}$$

For the purpose of studying the frequency-domain characteristics of the vocal tract, the boundary condition of the glottis may be idealized as a Norton equivalent network.[28] Thus we assume a condition of the form

$$Y_g\hat{p} + \hat{U} = \hat{U}_g \tag{4.17}$$

where Y_g is the Norton-equivalent admittance of the glottis and $\hat{U}_g$ is the equivalent volume-velocity source. If the glottal impedance is assumed much greater than the input impedance of the vocal tract then Eq. (4.17) is simply $\hat{U} = \hat{U}_g$.

The differential equations (4.7) and (4.8) are now transformed to a set of finite-difference equations which may be solved efficiently using the algorithm presented in Appendix B. Suppose the length of the vocal tract is ℓ. Let the glottis be located at $x = -\ell$ and the mouth at $x = 0$. We seek the acoustic pressure and volume velocity at N equally spaced points ($k = -N+1, \ldots, -1, 0$) on the interval $[-\ell, 0]$. The differential equations are transformed to finite-difference equations according to the rule:

$$\left[\frac{df}{dx} + g\right]_{x=(k-\frac{1}{2})\Delta x} = 0 \quad \rightarrow \quad \frac{1}{\Delta x}(f_k - f_{k-1}) + \tfrac{1}{2}(g_k + g_{k-1}) = 0 \tag{4.18}$$

where

$$\Delta x = \ell/(N-1)$$

$$f_k = f(k\Delta x)$$

$$g_k = g(k\Delta x)$$

It can be shown that this transformation is equivalent to integration by the trapezoid rule or to a bilinear transformation in the (spatial) frequency domain (see Appendix A).

The transformation (4.18) is now applied to Eqs. (4.11-4.12) at the (N-1) midpoints between each of the N sample points to obtain the system of (2N-2) equations:

$$\frac{1}{\Delta x}(\hat{p}_k - \hat{p}_{k-1}) + \tfrac{1}{2}(Z_k\hat{U}_k + Z_{k-1}\hat{U}_{k-1}) = 0$$

$$\frac{1}{\Delta x}(\hat{U}_k - \hat{U}_{k-1}) + \tfrac{1}{2}(Y_k\hat{p}_k + Y_{k-1}\hat{p}_{k-1}) = 0$$

$$K = -N+2,\ldots,-1,0$$

or equivalently,

$$\hat{p}_k + \frac{\Delta x}{2} Z_k\hat{U}_k - \hat{p}_{k-1} + \frac{\Delta x}{2} Z_{k-1}\hat{U}_{k-1} = 0$$

$$\frac{\Delta x}{2} Y_k\hat{p}_k + \hat{U}_k + \frac{\Delta x}{2} Y_{k-1}\hat{p}_{k-1} - \hat{U}_{k-1} = 0$$

$$k = -N+2,\ldots,-1,0 \qquad (4.19)$$

where Z_k is taken to include R_a and Y_k is taken to include G_a, i.e.,

$$Z_k = Z(k\Delta x) + R_a(k\Delta x)$$

$$Y_k = Y(k\Delta x) + G_a(k\Delta x).$$

We have 2N-2 equations in 2N variables. The two additional equations necessary for a unique solution are supplied by the boundary conditions at the mouth (4.16) and glottis (4.17). The complete system of 2N equations determining the acoustic pressure and volume velocity distributions is the following:

$$\begin{bmatrix}
1 & -Z_{rad} & & & & & & & & \\
1 & Z'_o & -1 & Z'_{-1} & & & & & & \\
Y'_o & 1 & Y'_{-1} & -1 & & & & & & \\
 & & 1 & Z'_{-1} & -1 & Z'_{-2} & & & & \\
 & & Y'_{-1} & 1 & Y'_{-2} & -1 & & & & \\
 & & & & \ddots & \ddots & & & & \\
 & & & & 1 & Z'_{-N+2} & -1 & Z'_{-N+1} & & \\
 & & & & Y'_{-N+2} & 1 & Y'_{-N+1} & -1 & & \\
 & & & & & & Y_g & 1 &
\end{bmatrix}
\begin{bmatrix}
\hat{p}_o \\ \hat{U}_o \\ \hat{p}_{-1} \\ \hat{U}_{-1} \\ p_{-2} \\ \vdots \\ \hat{U}_{-N+2} \\ \hat{p}_{-N+1} \\ \hat{U}_{-N+1}
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \hat{U}_g
\end{bmatrix}
\tag{4.20}$$

where $Z'_k = \frac{\Delta x}{2} Z_k$ and $Y'_k = \frac{\Delta x}{2} Y_k$. Although the dimensionality of the system (4.20) must be large in order to obtain a very accurate solution, the system can be solved efficiently with the algorithm of Appendix B since most of the elements of the matrix are zero.

These equations were solved on the G.E. 635 computer at Bell Laboratories, Murray Hill, N. J. Figures 4.1 through 4.16 show plots of the magnitude of the glottal-to-mouth transfer ration $\hat{U}_m/\hat{U}_g$ for various cross-sectional area configurations and various types of losses. In all cases the pressure and volume velocity are computed at 96 sample points along the vocal tract, and the glottal impedance is approximated as infinite. In those cases with yielding walls, the wall parameters are those given by Flanagan:[19]

$$M_w = 0.4 \text{ gm/cm}^2$$

$$b_w = 6500 \text{ dyne-sec/cm}^3$$

$$K_w = 0$$

For lack of more precise data, at this time, these parameters are assumed constant along the length of the vocal tract. Figures 4.1 through 4.3 show $|\hat{U}_m/\hat{U}_g|$ for a 17.5 cm long uniform tube with a 5 cm^2 cross-sectional area and "short circuit" (p=0) at the mouth. Figures 4.4 through 4.6 show $|\hat{U}_m/\hat{U}_g|$ for the same tube, but terminated with the radiation load. Figures 4.7-4.16 show $|\hat{U}_m/\hat{U}_g|$ for vocal-tract configurations corresponding to the Russian vowels /a/, /e/, /i/, /o/, and /u/ (with the radiation load). The cross-sectional area data were obtained from Fant.[20] It should be noted that his data are quantized to the nearest multiple of $\sqrt[3]{2}$ and given at only 35 to 40 points along the vocal tract. The values of the area functions at intermediate sample points were obtained by linear interpolation.

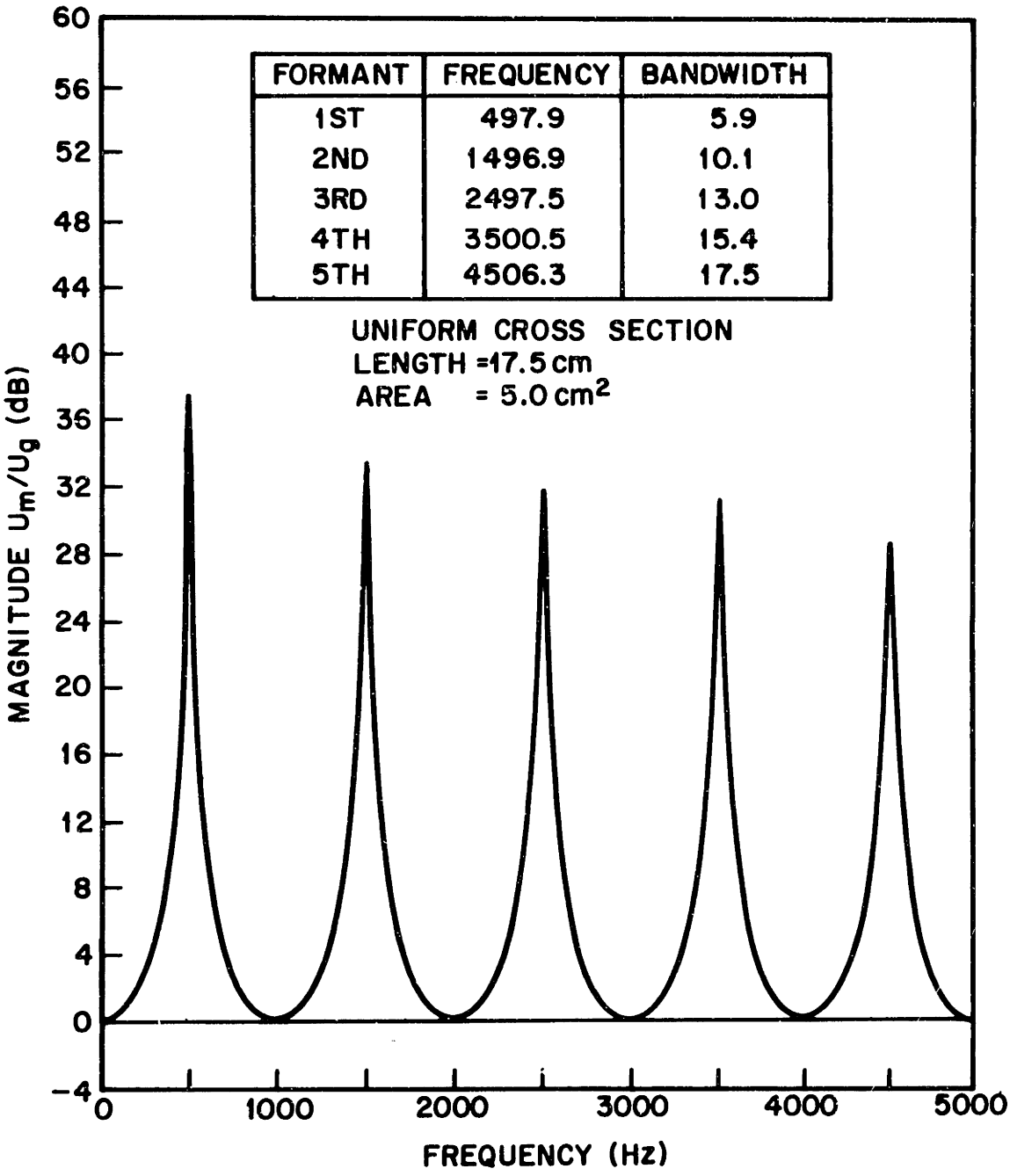


FIGURE 4.1

MAGNITUDE OF TRANSFER RATIO U_m/U_g FOR UNIFORM ACOUSTIC TUBE WITH RIGID WALL, VISCOUS AND THERMAL LOSSES, AND TERMINATED IN A SHORT CIRCUIT (p_m=0)

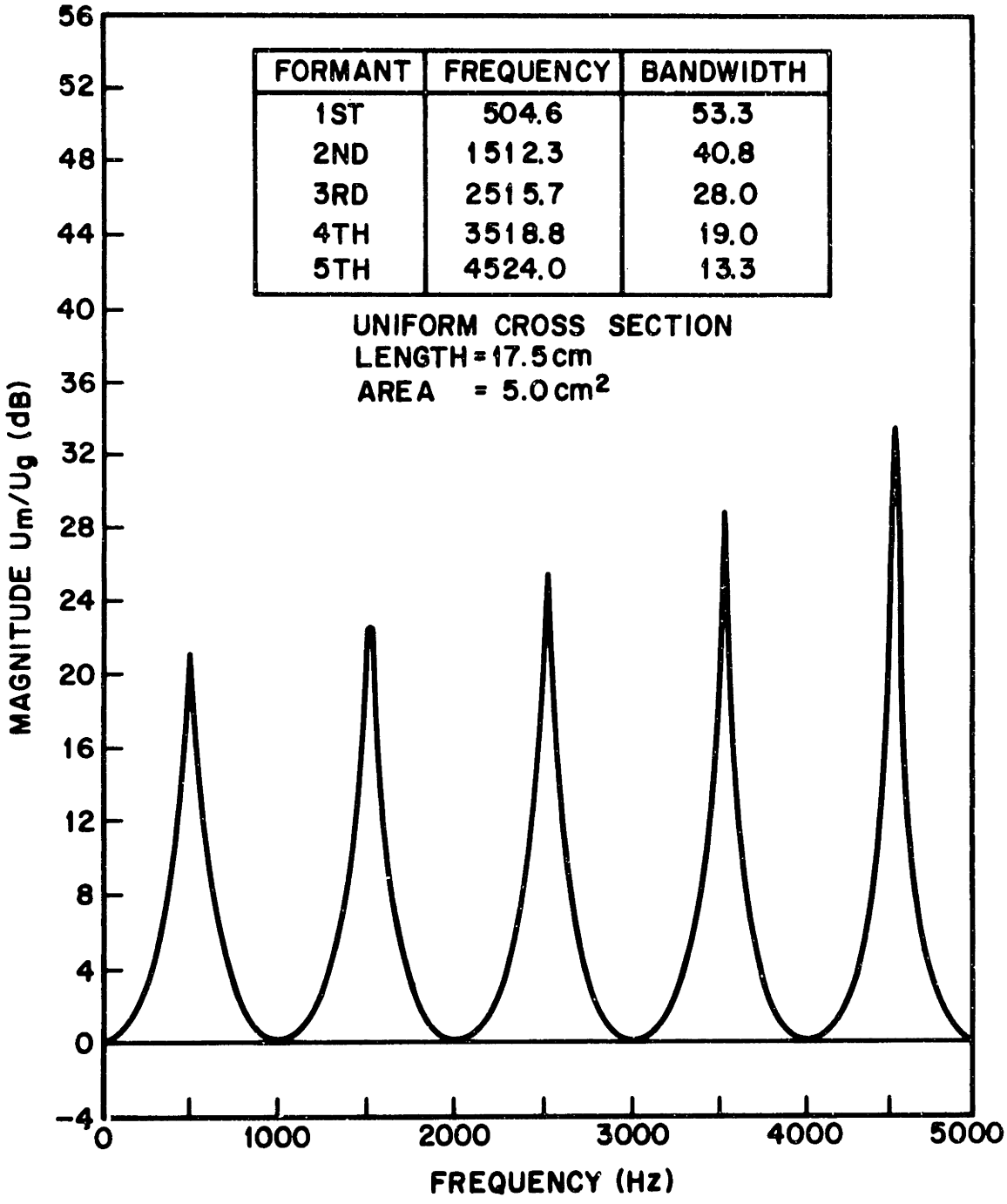


FIGURE 4.2

MAGNITUDE OF TRANSFER RATIO U_m/U_g FOR UNIFORM ACOUSTIC TUBE WITH YIELDING WALL, NO OTHER LOSSES, AND TERMINATED IN A SHORT CIRCUIT ($p_m = 0$)

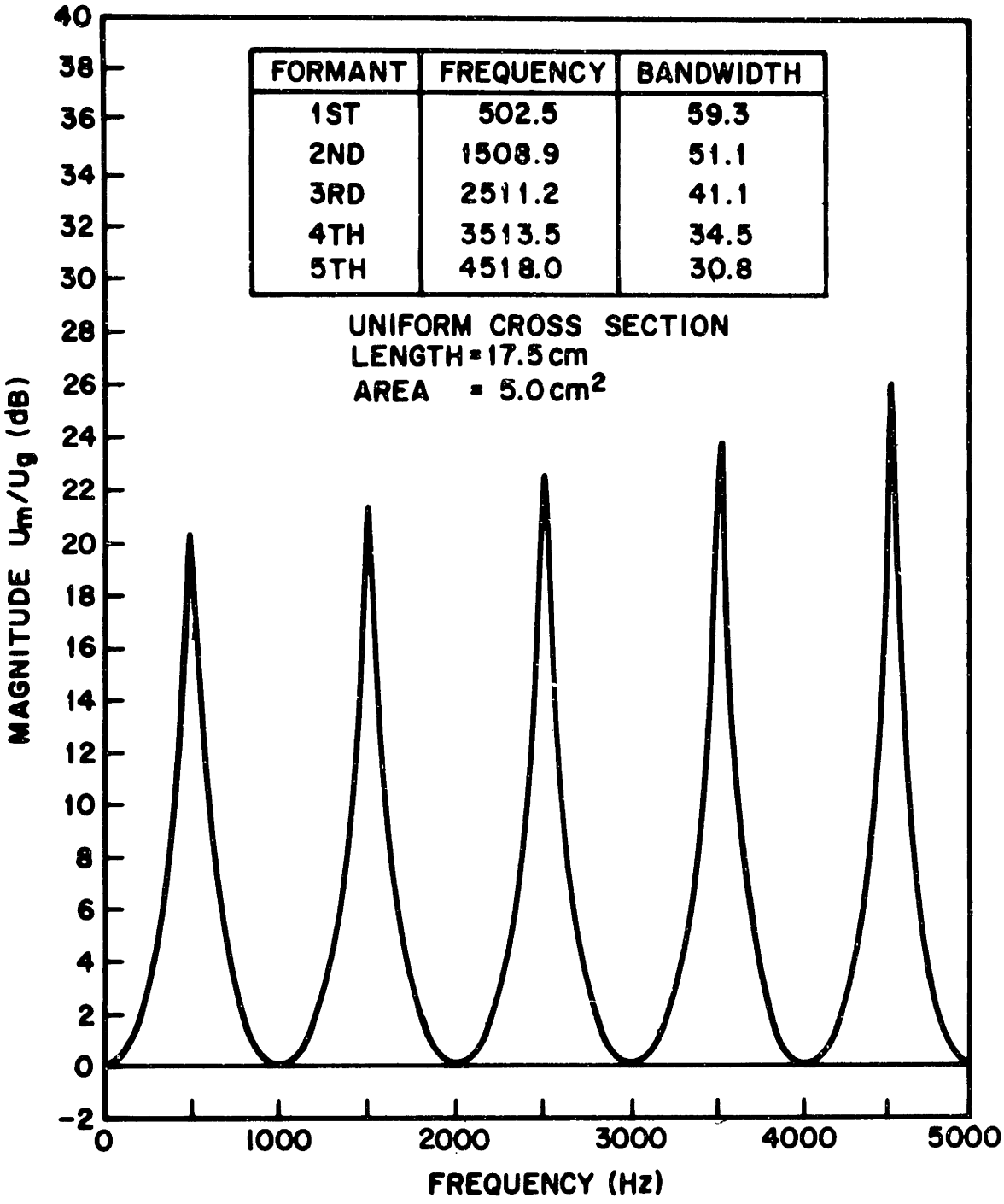


FIGURE 4.3

MAGNITUDE OF TRANSFER RATIO U_m/U_g FOR UNIFORM ACOUSTIC TUBE WITH YIELDING WALL, VISCOUS AND THERMAL LOSSES, AND TERMINATED IN A SHORT CIRCUIT ($p_m = 0$)

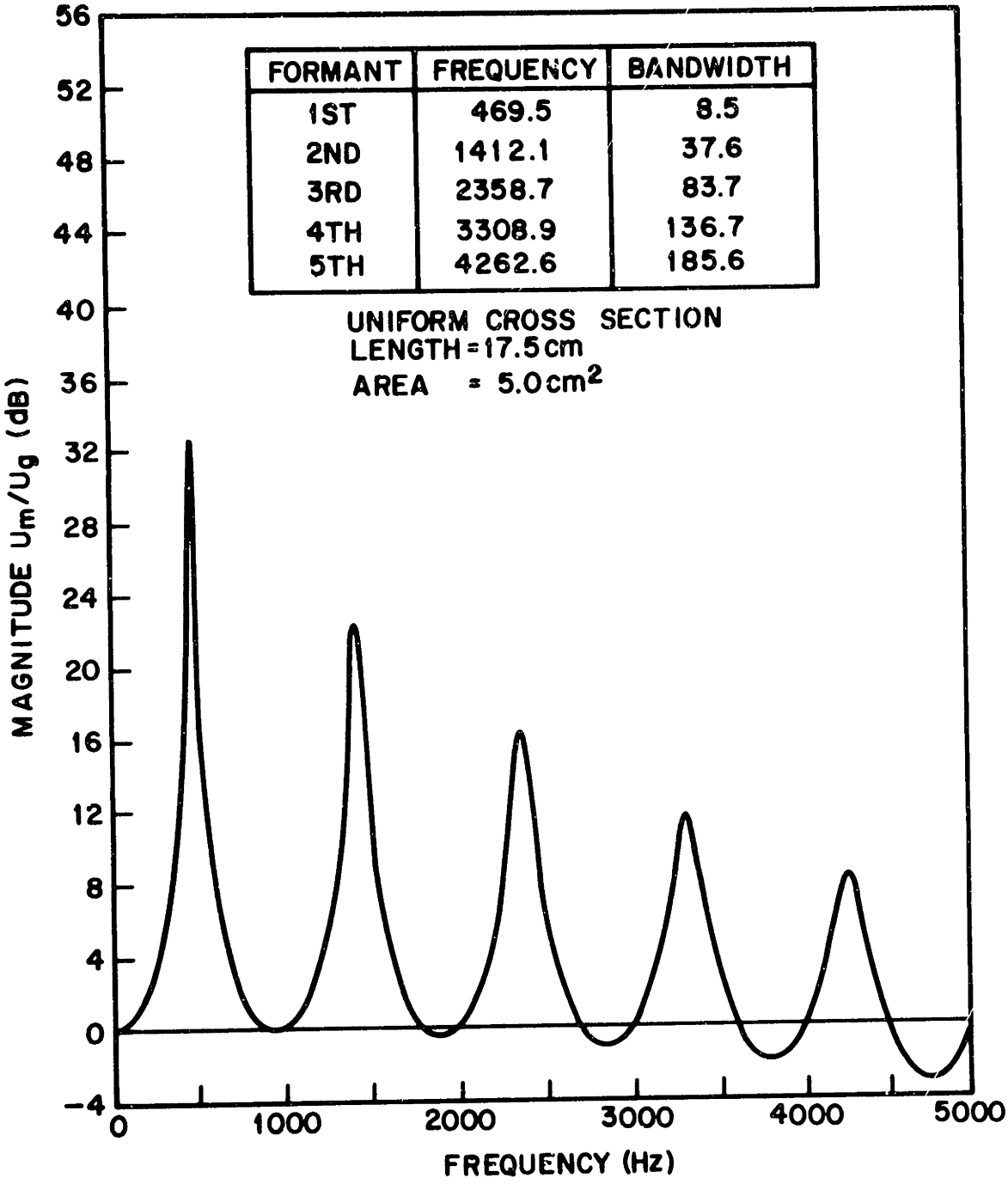


FIGURE 4.4

MAGNITUDE OF TRANSFER RATIO U_m/U_g FOR UNIFORM ACOUSTIC TUBE WITH RIGID WALL, VISCOUS AND THERMAL LOSSES, AND TERMINATED WITH RADIATION LOAD

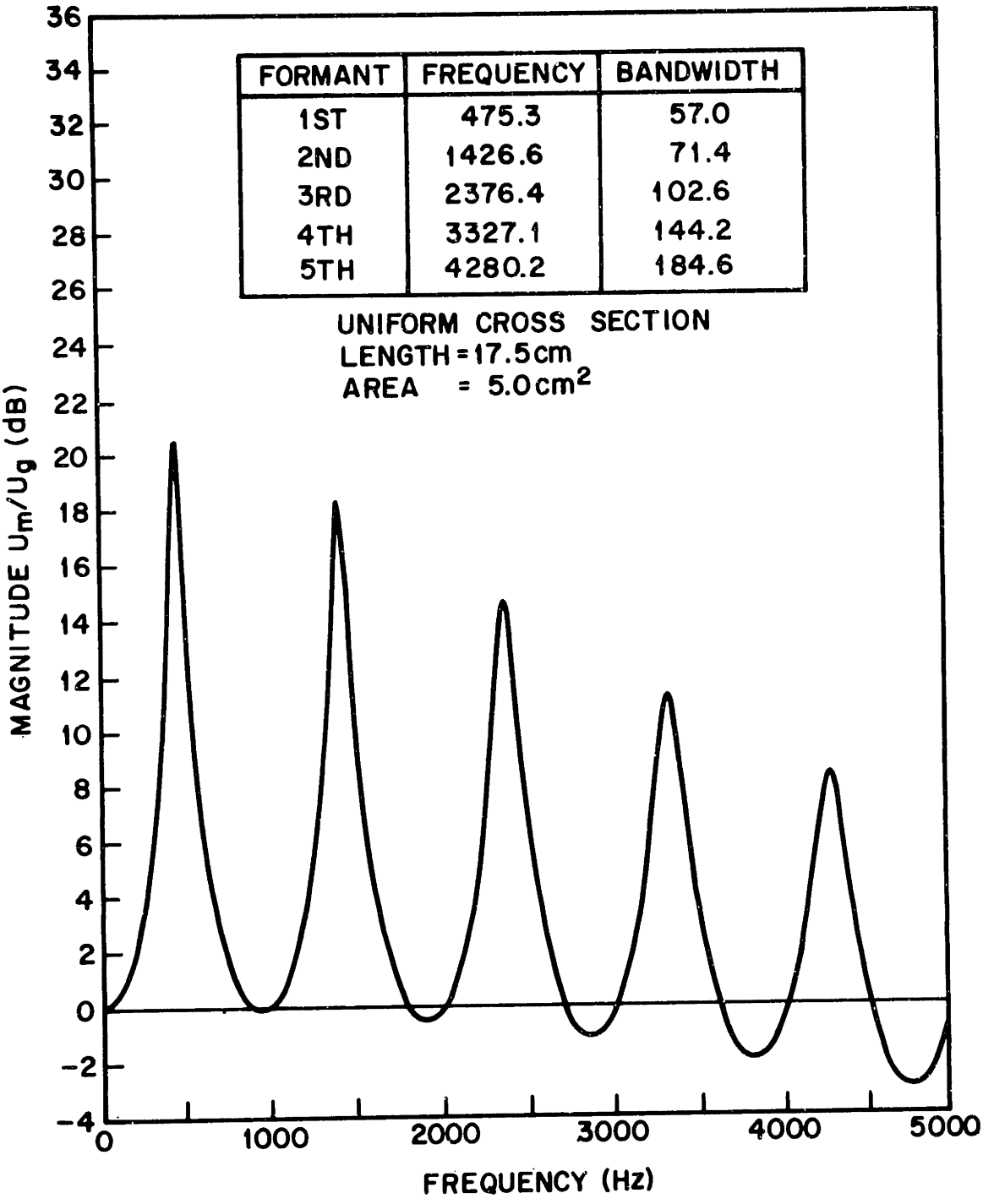


FIGURE 4.5

MAGNITUDE OF TRANSFER RATIO Um/Ug FOR UNIFORM ACOUSTIC TUBE WITH YIELDING WALL, NO OTHER LOSSES, AND TERMINATED WITH RADIATION LOAD

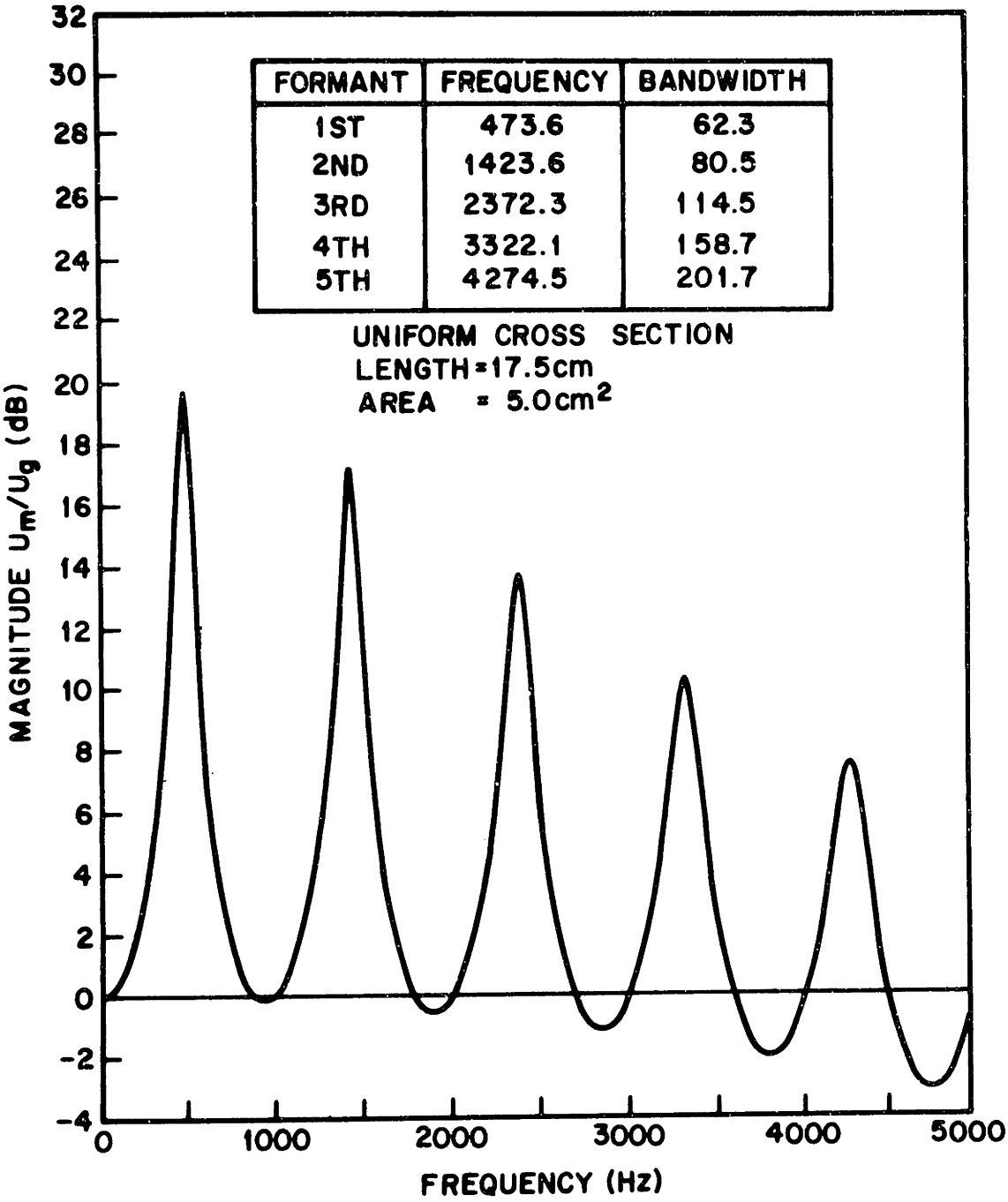


FIGURE 4.6

MAGNITUDE OF TRANSFER RATIO U_m/U_g FOR UNIFORM ACOUSTIC TUBE WITH YIELDING WALL, VISCOUS AND THERMAL LOSSES, AND TERMINATED WITH RADIATION LOAD

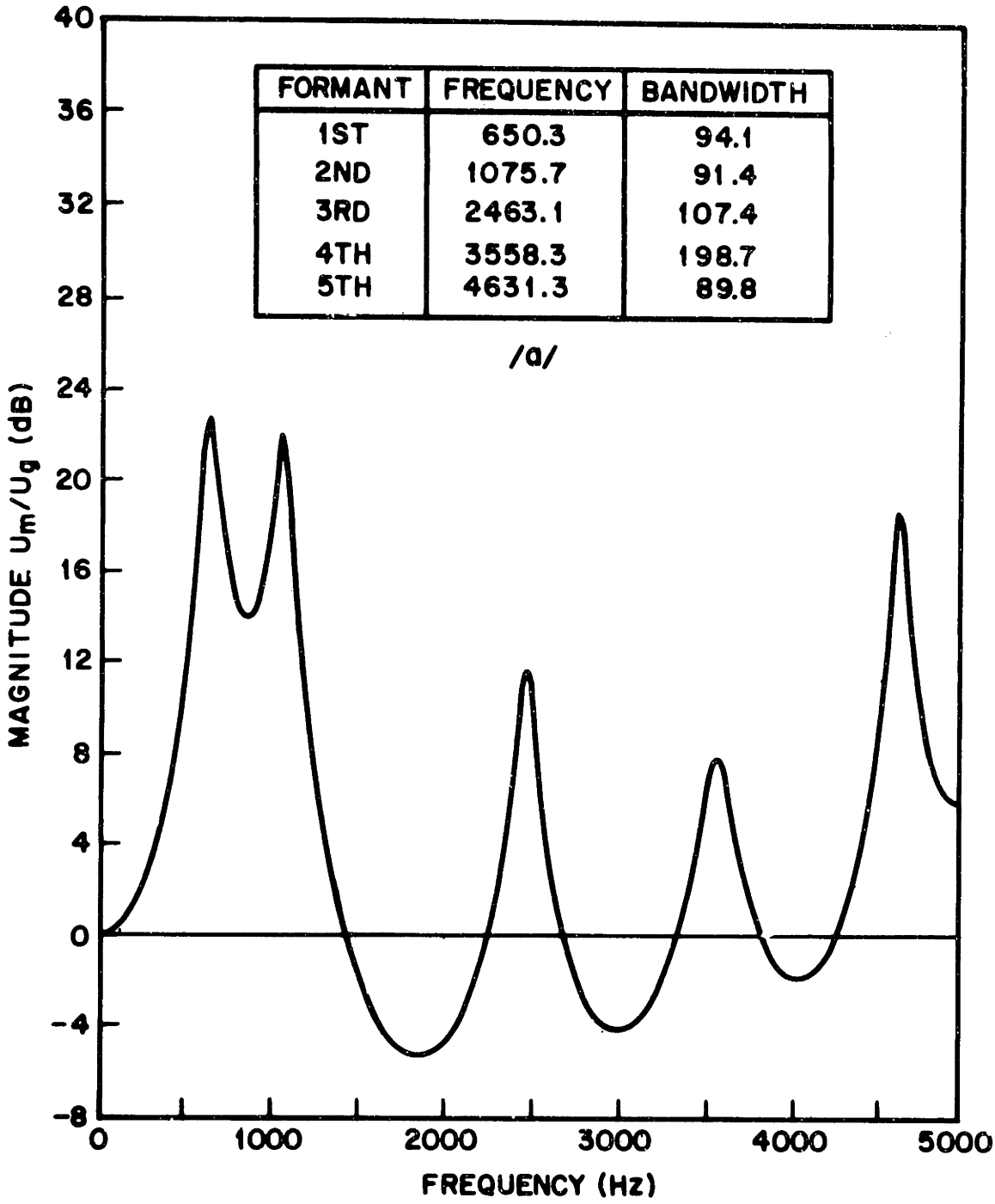


FIGURE 4.7

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /a/ INCLUDING YIELDING WALL AND VISCOUS AND THERMAL LOSSES

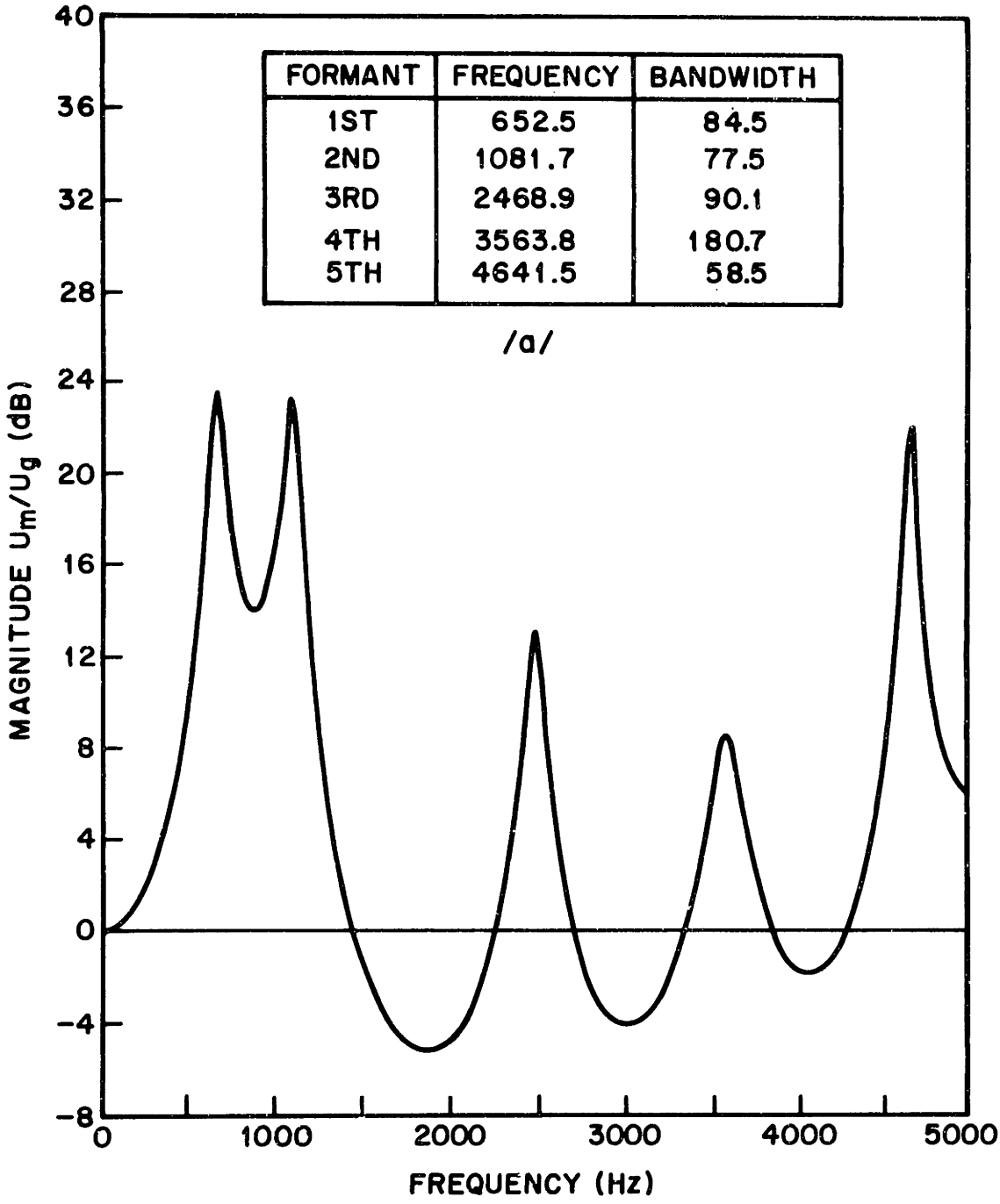


FIGURE 4.8

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /a/ INCLUDING YIELDING WALL, BUT NEGLECTING VISCOUS AND THERMAL LOSSES

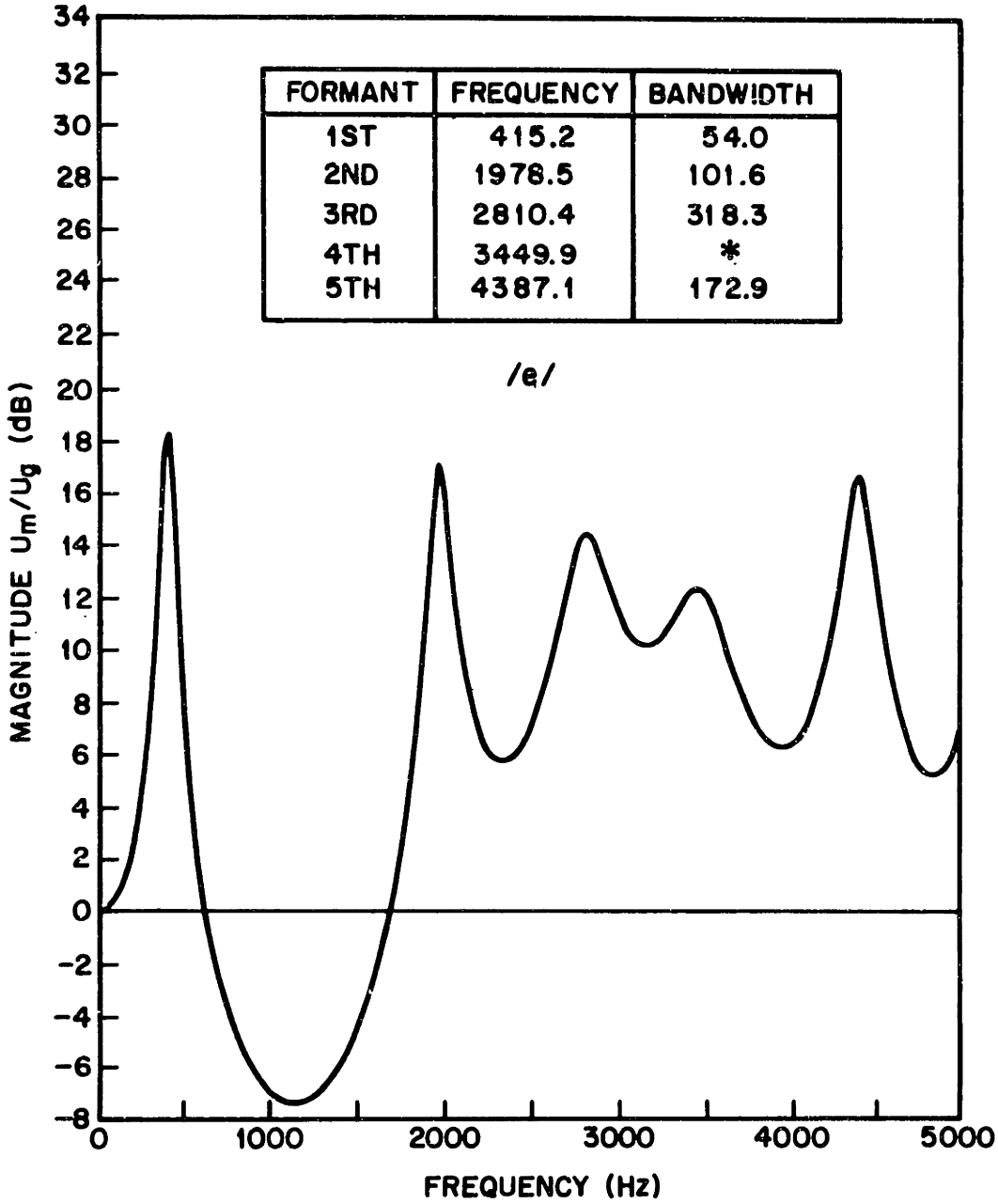


FIGURE 4.9

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /e/ INCLUDING YIELDING WALL AND VISCOUS AND THERMAL LOSSES

*BANDWIDTH COULD NOT BE DETERMINED BY LOCATING HALF-POWER POINTS.

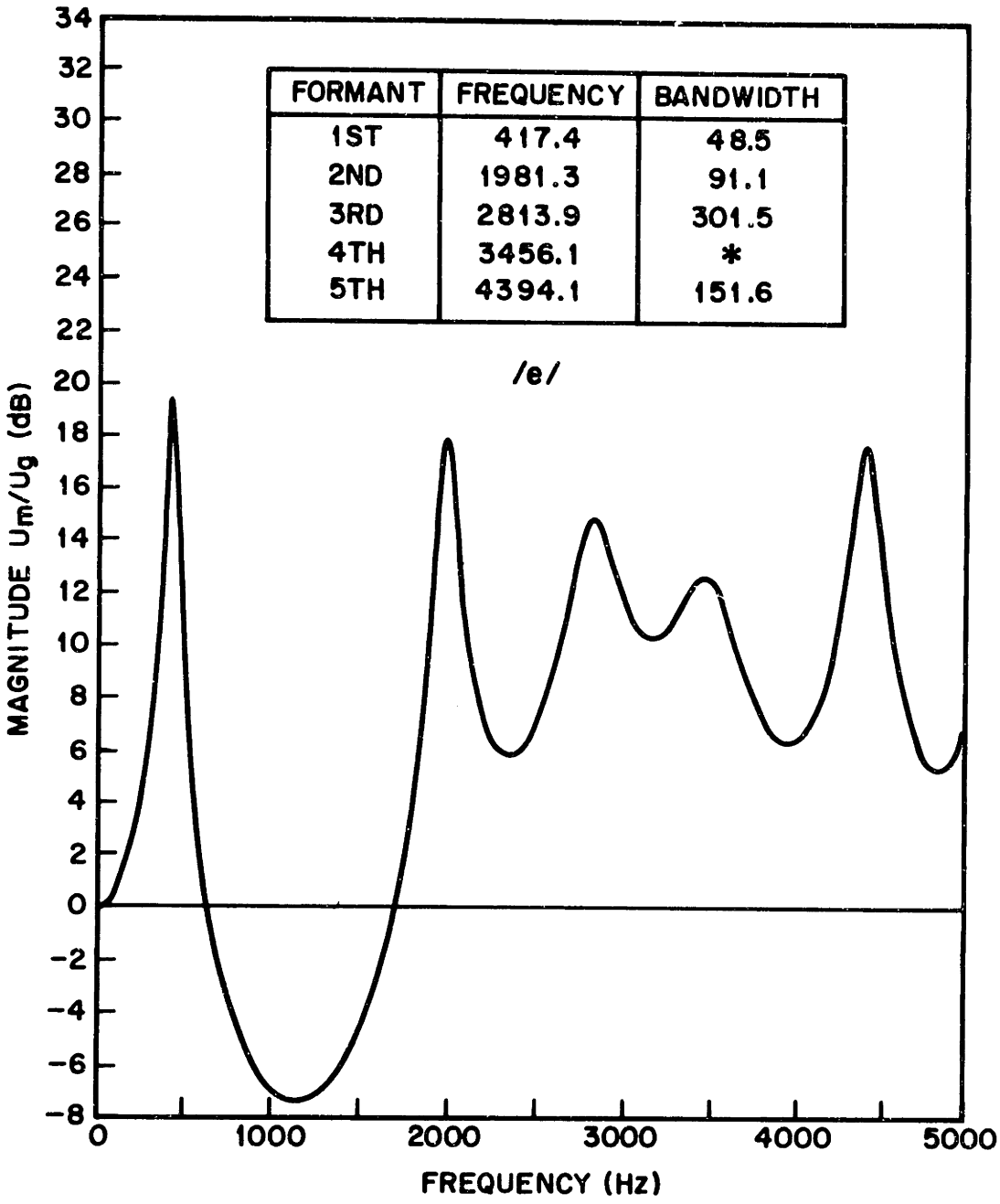


FIGURE 4.10

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /e/ INCLUDING YIELDING WALL, BUT NEGLECTING VISCOUS AND THERMAL LOSSES

*BANDWIDTH COULD NOT BE DETERMINED BY LOCATING HALF-POWER POINTS.

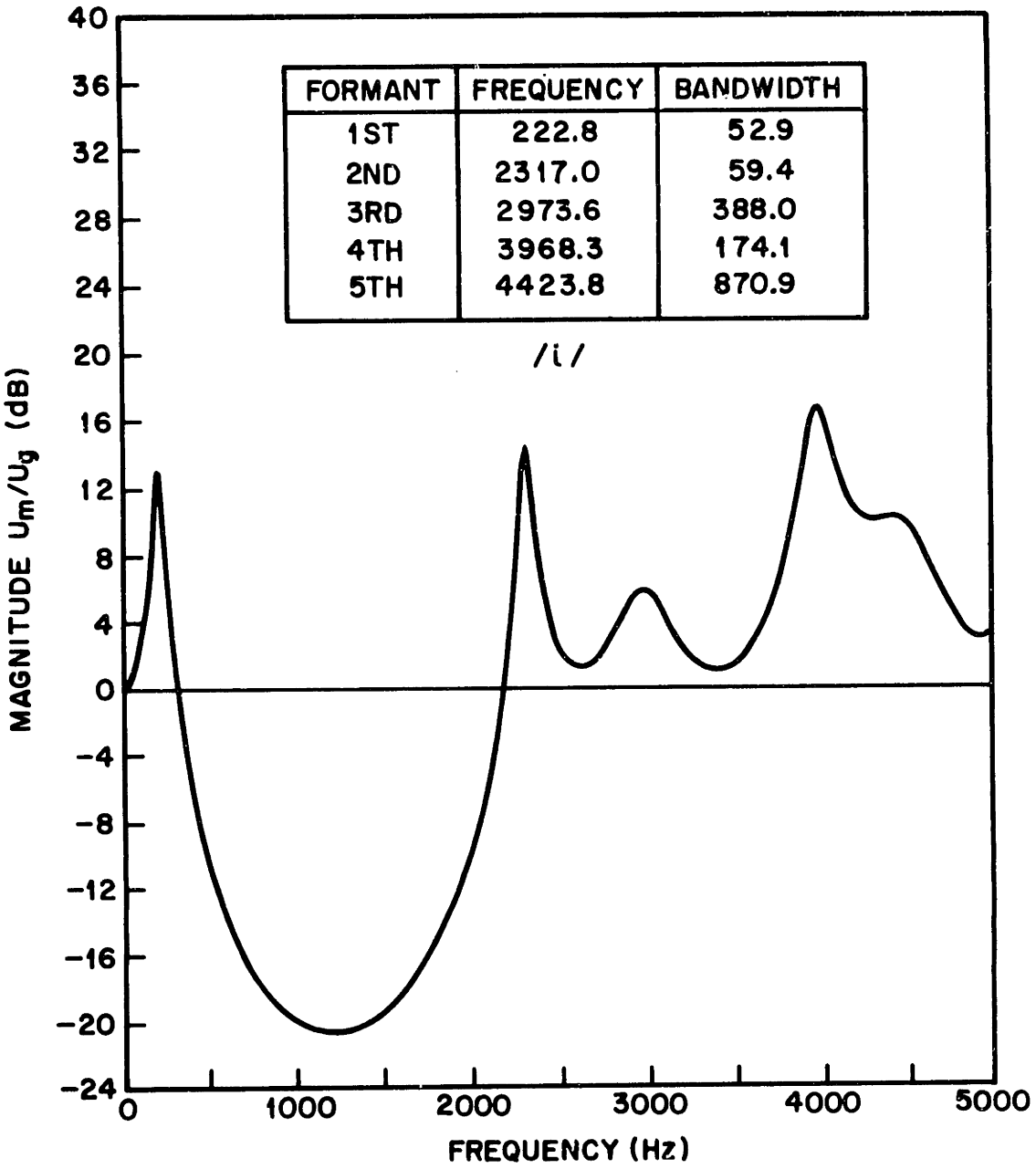


FIGURE 4.11

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /i/ INCLUDING YIELDING WALL AND VISCOUS AND THERMAL LOSSES

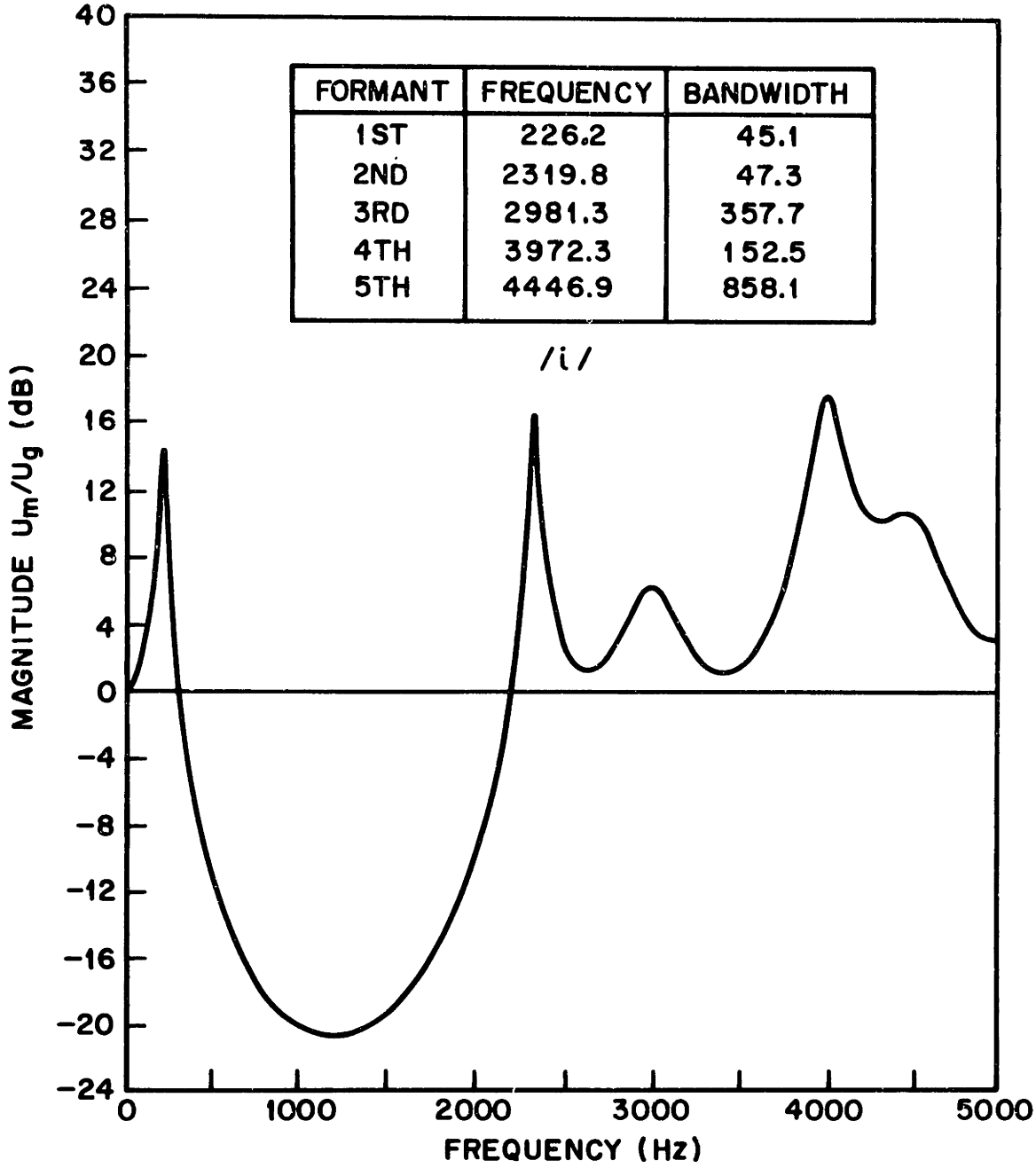


FIGURE 4.12

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /i/ INCLUDING YIELDING WALL, BUT NEGLECTING VISCOUS AND THERMAL LOSSES

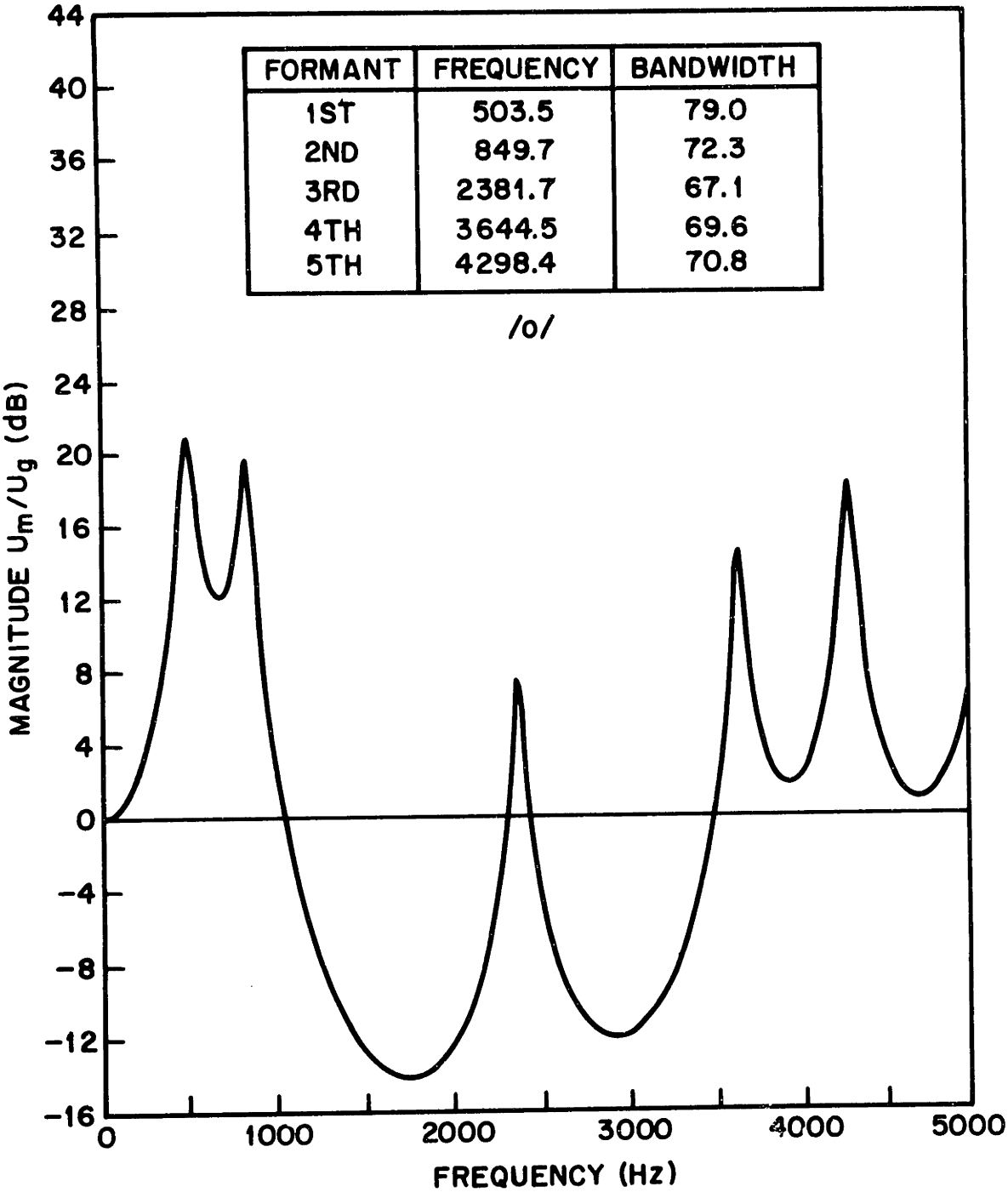


FIGURE 4.13

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /o/ INCLUDING YIELDING WALL AND VISCOUS AND THERMAL LOSSES

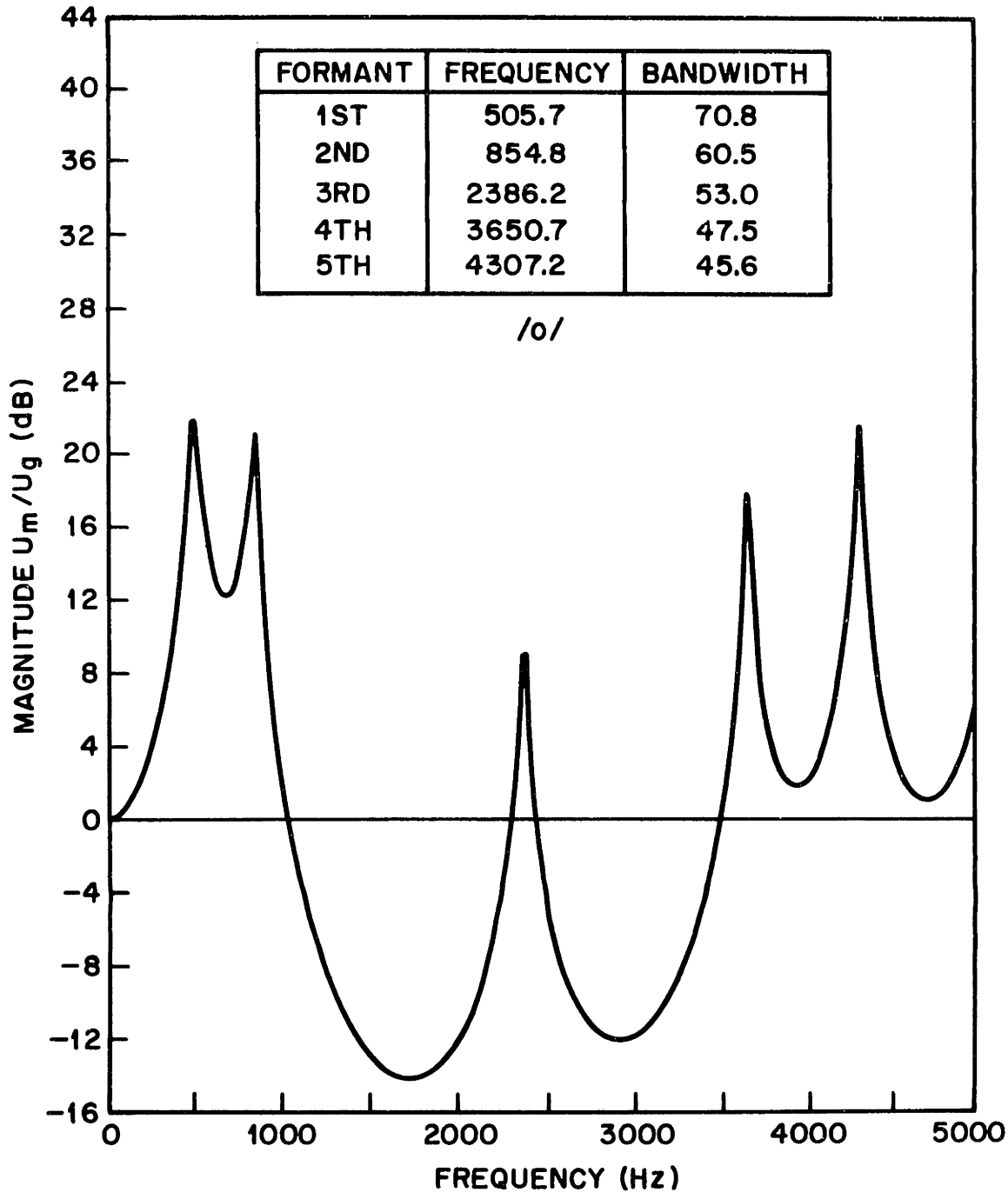


FIGURE 4.14

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /o/ INCLUDING YIELDING WALL BUT NEGLECTING VISCOUS AND THERMAL LOSSES

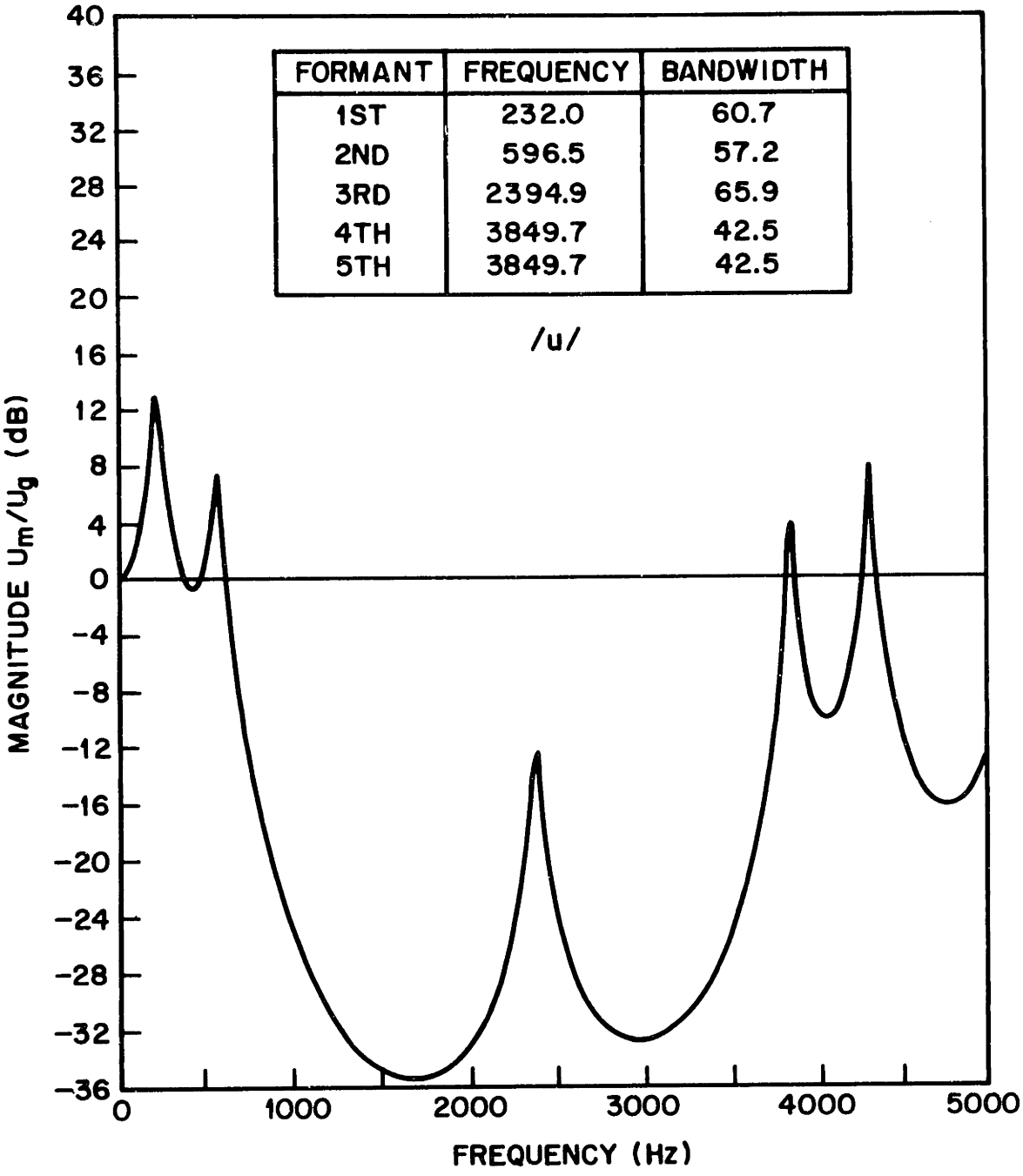


FIGURE 4.15

MAGNITUDE OF TRANSFER RATIO Um/Ug OF VOCAL TRACT MODEL FOR /u/ INCLUDING YIELDING WALL AND VISCOUS AND THERMAL LOSSES

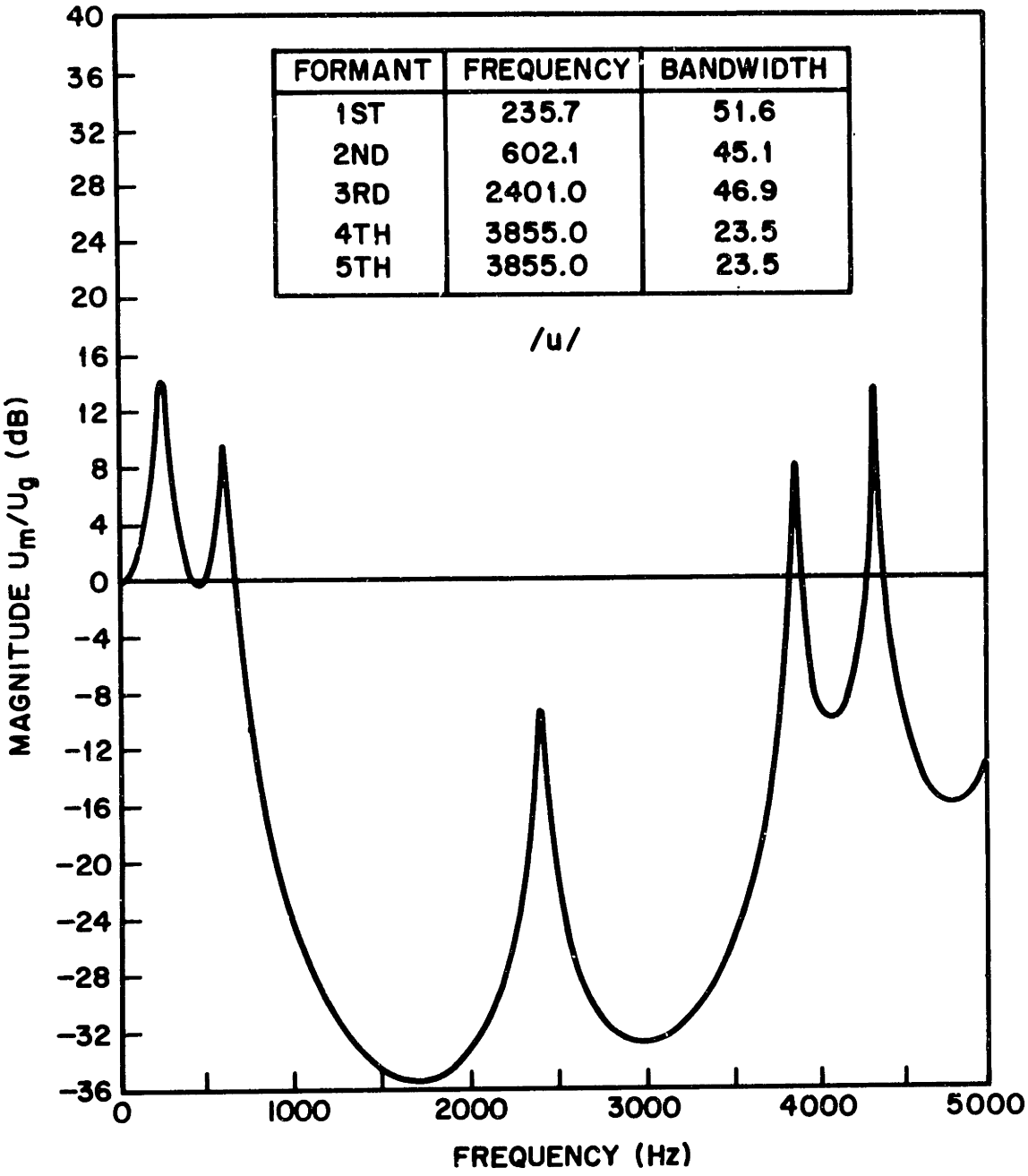


FIGURE 4.16

MAGNITUDE OF TRANSFER RATIO U_m/U_g OF VOCAL TRACT MODEL FOR /u/ INCLUDING YIELDING WALL, BUT NEGLECTING VISCOUS AND THERMAL LOSSES

From these plots of the transfer ratio $|U_m/U_g|$ for vocal-tract geometries corresponding to typical vowel sounds, some important properties of the model can be observed. In particular, the formant frequencies and bandwidths behave, with respect to the types of losses included in the model, as predicted by Flanagan's analysis.[19] The effect of the viscous and thermal boundary-layer losses is to slightly decrease the formant frequencies and increase the formant bandwidths, while the effect of the yielding duct wall is to increase both the formant frequencies (slightly) and bandwidths. The viscous and thermal effects are more apparent at high frequencies because the energy loss due to fluid viscosity and thermal conductivity in the boundary layer is proportional to the square root of the frequency; whereas the effect of the nonrigid duct wall is more apparent at low frequecies because the vibrating wall is a lossy, low-frequency mechanical oscillator whose response is essentially zero for frequencies above a few hundred Hertz. In fact, for the values of the wall parameters given by Flanagan, it can be seen that for low frequencies, the effects due to the yielding duct wall dominate the effects due to the boundary-layer losses. These effects can be seen clearly in Figs. 4.1-4.3, which show $|U_m/U_g|$ for a uniform acoustic tube terminated in a short circuit.

The effect of terminating the duct with the radiation load derived in Chapter III is to decrease the formant

frequencies and increase the formant bandwidths. This effect is greatest at high frequencies because the radiation load is modeled by a resistance and inductance in parallel, and thus behaves as a short circuit for DC and a resistance for high frequencies. From Figs. 4.4-4.6 it can be seen that at high frequencies, the effects of the radiation load on the ratio $|U_m/U_g|$ dominate the effects of the boundary-layer losses.

The important result of the frequency-domain analysis is that the data obtained from this analysis, shown in Figs. 4.7-4.16, compare favorably with characteristics of real speech.[20] Moreover, the error introduced by neglecting the boundary-layer losses, in the cases considered, is small enough to be considered acceptable. This error is on the order of a percent for the formant frequencies and ten percent for the formant bandwidths, which compares favorably with the difference limens of 3-5% for formant frequencies and 20-40% for formant bandwidths.[19] It is therefore reasonable to accept Eqs. (2.18-2.21) as providing a satisfactory description of the acoustical properties of the vocal tract.

CHAPTER V

Discrete-Variable Formulation of the Vocal-Tract Transmission Equations

The time-domain Eqs. (2.18-2.21) governing acoustic wave propagation in the vocal tract are a system of linear (hyperbolic) partial-differential equations. These equations are considerably more difficult to deal with than the ordinary-differential equations in the frequency-domain formulation of Chapter IV and raise issues that are outside the realm of the usual (one-dimensional) sampled-data theory.

The problem of simulating the one-dimensional model of the vocal tract in the time-domain has a two-dimensional character: there is one time dimension and one space dimension. If we use the same convention as in Chapter IV, i.e., locate the glottis at $x = -\ell$ and the mouth at $x = 0$, then the problem is to solve the partial-differential equations for $p(x,t)$ and $U(x,t)$ on the semi-infinite strip of the x - t plane, shown in Fig. 5.1. This requires that boundary conditions be specified along the boundaries of the strip. The condition on the line $x = -\ell$ is supplied by the vocal-cord model, the condition on the line $x = 0$ is supplied by the radiation load, and the condition on the line segment $t = 0$, $-\ell \leq x \leq 0$ is supplied by assuming initial rest.

In order to realize a digital simulation of the partial-differential equations, it is necessary to formulate

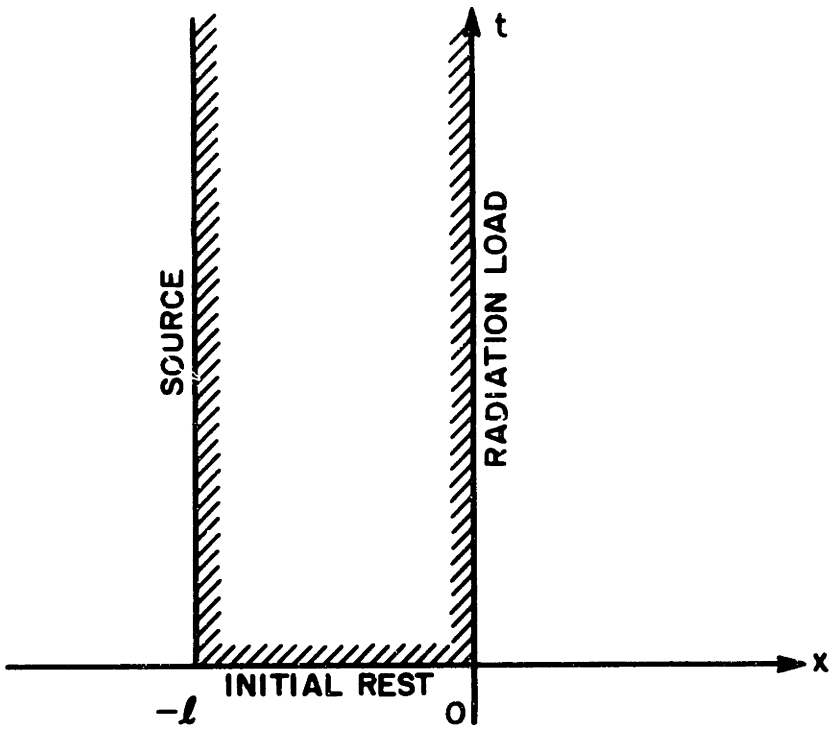


FIGURE 5.1
SEMI-INFINITE STRIP IN x-t PLANE ON WHICH PARTIAL-DIFFERENTIAL EQUATIONS ARE TO BE SOLVED

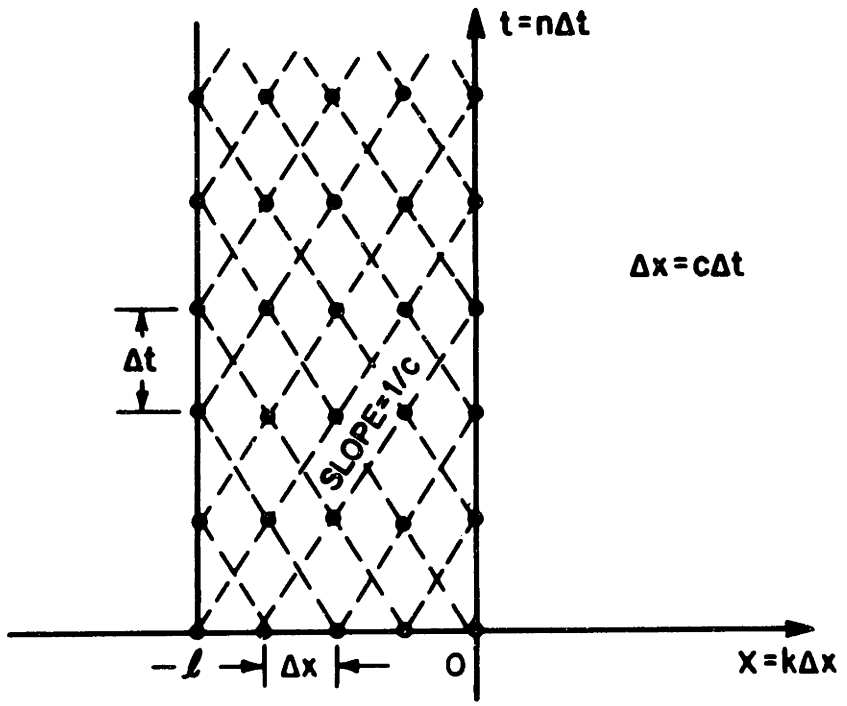


FIGURE 5.2
RECTANGULAR NET CONSTRUCTED ON CHARACTERISTICS OF LOSS-FREE WAVE EQUATIONS

the problem in terms of discrete variables and finite differences. We construct a "net" of discrete points on the semi-infinite strip in the x-t plane and seek the values of the acoustic pressure and volume velocity at all points on the net. The arrangement of the points on the net is an important issue and one without parallel in one-dimensional signal processing theory.

It is a well known result from the theory of partial-differential equations[12,13,24] that discontinuities or irregularities in the solution of equations of the hyperbolic type tend to propagate along certain curves in the domain called "characteristics". It is thus desirable, although not necessary, to construct the net so that the net points lie on the characteristics. The construction of such a "characteristic net" is generally not an easy task since it may be impossible to determine the characteristic curves in advance of the determination of the solution itself.[12,13] However, we know certain properties of the vocal-tract transmission equations which allow us to construct, a priori, a good approximation to the characteristic net since these equations are closely related to the simple wave equations of the form

$$\frac{\partial f}{\partial x} + \frac{1}{c}\frac{\partial g}{\partial t} = 0$$

$$\frac{\partial g}{\partial x} + \frac{1}{c}\frac{\partial f}{\partial t} = 0. \qquad (5.1)$$

The characteristics of the simple wave equations (5.1) are straight lines in the x-t plane with slope $\pm 1/c$. We expect, therefore, that a net whose points lie on these lines will be a good approximation to the actual characteristic net for the more complicated equations. It is clear that such a net can be constructed as a rectangular mesh with uniform spacing in both the x and t directions as shown in Fig. 5.2. If Δx is the net spacing in the x direction and Δt is the net spacing in the t direction then a uniformly spaced characteristic net requires that

$$\Delta t = \Delta x/c. \tag{5.2}$$

Relations between the spatial and temporal sampling rates will appear again and arise from various considerations. In this case, although a characteristic net is desirable (in fact, nearly optimum), Eq. (5.2) is not in general a requirement for a working simulation.

The vocal-tract transmission equations will now be transformed to a set of partial-difference equations relating the values of acoustic pressure and volume velocity at only those points on the net. Formulations of the partial-difference equations may be classified as either explicit or implicit formulations. Explicit formulations are those for which it is possible to determine the solution to the difference equations at a particular net point explicitly from those net points at which the solution has already been computed. Implicit formulations are those formulations which require the solution of a system of simultaneous

equations in order to obtain the solution of the difference equation at a set of net points. For example, in an implicit formulation, the solution at all the net points on a particular characteristic or at a particular time step might be determined simultaneously as the solution to a set of algebraic equations. The explicit formulations appear to have the advantage of speed and simplicity. Indeed, most previous vocal-tract simulations have used explicit formulations. Explicit formulations of wave equations all suffer from the restriction that the sampling intervals must satisfy the constraint*

$$\Delta x \geq c\Delta t. \tag{5.3}$$

If this condition is violated, the simulation will be unstable, in spite of the fact that the solution to the partial-differential equations may itself be stable.[29] Moreover, many explicit formulations are unstable for any choice of sampling intervals.[12] Explicit formulations of the two-point boundary-value problem also present certain difficulties in simulating the boundary conditions. Implicit formulations, on the other hand, can be found which are stable for any choice of sampling intervals and which allow formulation of the boundary conditions of the two-point boundary-value problem in a natural way.

---

* Although this requirement is derived by analysis, it can be interpreted intuitively as a statement of causality.

An implicit formulation is chosen here to simulate the vocal-tract equations. This formulation has been used successfully by researchers[30-36] to simulate similar types of equations and can be shown to be stable for any choice of sampling intervals.[35] The formulation is called a centered-difference scheme, for reasons that will become apparent shortly, and is a generalization of the trapezoid rule to two dimensions obtained by using divided-central differences and averages. The formulation of the finite-difference equations proceeds according to the following rule:

$$\frac{\partial}{\partial x} \rightarrow \frac{1}{\Delta x} \mu_t \delta_x$$
$$\frac{\partial}{\partial t} \rightarrow \frac{1}{\Delta t} \mu_x \delta_t \tag{5.4}$$

where δ_ν is the central difference operator and μ_ν is the averaging operator in the ν direction.[12] Thus, we make the substitutions:

$$\left.\frac{\partial f}{\partial x}\right|_{\substack{t=(n+\frac{1}{2})\Delta t \\ x=(i+\frac{1}{2})\Delta x}} = \frac{1}{2\Delta x}\left[f_{i+1}^{n+1} - f_i^{n+1} + f_{i+1}^{n} - f_i^{n}\right]$$
$$\left.\frac{\partial f}{\partial t}\right|_{\substack{t=(n+\frac{1}{2})\Delta t \\ x=(i+\frac{1}{2})\Delta x}} = \frac{1}{2\Delta t}\left[f_{i+1}^{n+1} - f_{i+1}^{n} + f_i^{n+1} - f_i^{n}\right]. \tag{5.5}$$

Equations (5.5) may be interpreted as simulating the derivative of a function f, evaluated at the center of each rectangle comprising the net, in terms of the values of the function on the vertices of the rectangle. Specifically,

the value of each derivative at the center of a rectangle is obtained by differencing the function on the two sides of the rectangle parallel to the direction of differentiation and averaging in the other direction. Because of the averaging, this formulation tends to suppress (high frequency) roundoff noise as does the trapezoid rule in one dimension.[37]

In addition to having favorable noise characteristics, this formulation has several other desirable attributes. It has second order accuracy and is stable for any ratio of sampling intervals $\Delta x/\Delta t$. When applied to a system of linear constant-coefficient partial-differential (such as Eq. 5.1), the formulation may be interpreted as a bilinear transformation of the analog frequency space to the digital frequency space producing a tangential warping of the frequency axes (ω and β).* Moreover, in the special case of a characteristic net, $\Delta x = c\Delta t$, the simulation of Eqs. (5.1) is exact. That is, any solution to the finite-difference equations is also a solution to the partial-differential equations. This property of the formulation may be interpreted as leaving the dispersion relation, ω vs. β, invariant under the frequency-warping transformation. In this case, the only error in the simulation is the error which may be introduced by the simulation of the boundary conditions.

Another important attribute of the formulation is that certain important characteristics of the physical model

* See Appendix A.

are preserved in the simulation. Specifically, the partial-differential equations which we wish to simulate are conservation laws: Eq. (2.18) requires that fluid momentum be conserved and Eq. (2.19) requires that fluid mass be conserved. In general, a set of finite-difference equations derived from these partial-differential equations need not be conservation laws themselves. However, the formulation presented here preserves this property of the model and the finite-difference equations are, in fact, conservation laws.

Applying the transformations (5.4) to the partial-differential equations (2.18) and 2.19) gives, after rearranging terms, the system of (2N-2) linear equations in 2N unknowns:

$$(ZU)_{i+1}^{n+1} + p_{i+1}^{n+1} + (ZU)_i^{n+1} - p_i^{n+1} = (ZU)_{i+1}^{n} - p_{i+1}^{n} + (ZU)_i^{n} + p_i^{n} \tag{5.6}$$

and

$$U_{i+1}^{n+1} + (Yp)_{i+1}^{n+1} - U_i^{n+1} + (Yp)_i^{n+1} = -U_{i+1}^{n} + (Yp)_{i+1}^{n} + U_i^{n} + (Yp)_i^{n} - \frac{\Delta x}{\Delta t}\left[A_{i+1}^{n+1} - A_{i+1}^{n} + A_i^{n+1} - A_i^{n}\right] \tag{5.7}$$

for

$$i = -N+1, \ldots, -2, -1$$

$$n = 0, 1, 2, \ldots$$

where

$$Z_i^n = \rho_o \frac{\Delta x}{\Delta t} \left[\frac{1}{A_o(x,t)} \right]_{x=i\Delta x}^{t=n\Delta t}$$

and

$$Y_i^n = \kappa_s \frac{\Delta x}{\Delta t} \left[A_o(x,t) \right]_{x=i\Delta x}^{t=n\Delta t} .$$

The remaining two equations necessary to determine a unique solution come, of course, from the boundary conditions at the mouth and glottis, and will be discussed in the next chapter.

At this point, we consider the problem of determining the value of the perturbed cross-sectional area, A_i^n appearing in Eq. (5.7). In Chapter II it was assumed that $A(x,t)$ could be approximated as

$$A(x,t) = A_o(x,t) + S_o(x,t)\xi(x,t) \tag{2.17}$$

where $\xi(x,t)$ satisfies

$$p(x,t) = M_w(x)\ddot{\xi}(x,t) + b_w(x)\dot{\xi}(x,t) + K_w(x)\xi(x,t). \tag{2.18}$$

At any position x along the vocal tract, Eq. (2.18) is an ordinary linear constant-coefficient differential equation. Thus it is appropriate to define the (bilateral) Laplace transforms

$$\hat{p}(x,s) = \int_{-\infty}^{\infty} p(x,t)e^{-st}dt$$
$$\hat{\xi}(x,s) = \int_{-\infty}^{\infty} \xi(x,t)e^{-st}dt. \tag{5.8}$$

Laplace transforming both sides of Eq. (2.18) with respect to t gives

$$\hat{p}(x,s) = M_w(x)s^2\hat{\xi}(x,s) + b_w(x)s\hat{\xi}(x,s) + K_w(x)\hat{\xi}(x,s)$$

or,

$$\frac{\hat{\xi}(x,s)}{\hat{p}(x,s)} = \frac{1}{M_w(x)s^2 + b_w(x)s + K_w(x)} \tag{5.9}$$

Equation (5.9) may be interpreted as the system function of a linear time-invariant system whose input is the acoustic pressure $p(x,t)$ and whose output is the normal wall displacement $\xi(x,t)$.

A discrete-time system analogous to Eq. (5.9) can be synthesized using the technique of impulse invariance.[15] According to this technique, a rational transfer function is expanded in partial fractions which transform as

$$\frac{R_i}{s-s_i} \rightarrow \frac{R_i\Delta t}{1-e^{s_i\Delta t}z^{-1}} \tag{5.10}$$

and

$$\frac{R_i}{(s-s_i)^2} \rightarrow \frac{R_i(\Delta t)^2 e^{s_i \Delta t} z^{-1}}{1-2e^{s_i \Delta t} z^{-1} + e^{2s_i \Delta t} z^{-2}} \tag{5.11}$$

We consider four distinct cases and apply impulse invariance to each:

Case I:

$$M_w = 0 \quad \text{(first-order system)}$$

$$\frac{\hat{\xi}(x,s)}{\hat{p}(x,s)} = \frac{1/b_w}{(s-\alpha_o)} \rightarrow \frac{\hat{\xi}(x,z)}{\hat{p}(x,z)} = \frac{\Delta t/b_w}{1-e^{\alpha_o \Delta t} z^{-1}}$$

$$\alpha_o = -K_w/b_w$$

$$\therefore \quad \xi(x,n\Delta t) = [\Delta t/b_w]\, p(x,n\Delta t) + e^{\alpha_o \Delta t}\, \xi(x,n\Delta t - \Delta t)$$

Case II:

$$b_w^2 - 4M_w K_w < 0, \quad M_w \neq 0 \quad \text{(underdamped)}$$

$$\frac{\hat{\xi}(x,s)}{\hat{p}(x,s)} = \frac{1/M_w}{(s-\sigma_o-j\omega_o)(s-\sigma_o+j\omega_o)} = \frac{\Delta t}{2j\omega_o M_w}\left[\frac{1}{s-\sigma_o-j\omega_o} - \frac{1}{s-\sigma_o+j\omega_o}\right]$$

$$\rightarrow \frac{\hat{\xi}(x,z)}{\hat{p}(x,z)} = \frac{\Delta t}{\omega_o M_w} \frac{e^{\sigma_o \Delta t} \sin(\omega_o \Delta t) z^{-1}}{1-2e^{\sigma_o \Delta t} \cos(\omega_o \Delta t) z^{-1} + e^{2\sigma_o \Delta t} z^{-2}}$$

$$\sigma_o = -b_w/2M_w \qquad \omega_o = \sqrt{K_w/M_w - (b_w/2M_w)^2}$$

$$\therefore \quad \xi(x,n\Delta t) = \left[2e^{\sigma_o \Delta t}\cos\omega_o\Delta t\right]\xi(x,n\Delta t-\Delta t) - e^{2\sigma_o\Delta t}\xi(x,n\Delta t-2\Delta t)$$

$$+ \left[\frac{\Delta t}{\omega_o M_w}\, e^{\sigma_o\Delta t}\sin\omega_o\Delta t\right] p(x,n\Delta t-\Delta t)$$

Case III:

$$b_w^2 - 4M_wK_w > 0, \quad M_w \neq 0 \quad \text{(overdamped)}$$

$$\frac{\hat{\xi}(x,s)}{\hat{p}(x,s)} = \frac{1/M_w}{(s-\sigma_o-\alpha_o)(s-\sigma_o+\alpha_o)} = \frac{\Delta t}{2\alpha_o M_w}\left[\frac{1}{s-\sigma_o-\alpha_o} - \frac{1}{s-\sigma_o+\alpha_o}\right]$$

$$\rightarrow \frac{\hat{\xi}(x,z)}{\hat{p}(x,z)} = \frac{\Delta t}{\alpha_o M_w}\,\frac{e^{\sigma_o\Delta t}\sinh(\alpha_o\Delta t)z^{-1}}{1-2^{\sigma_o\Delta t}\cosh(\alpha_o\Delta t)z^{-1} + e^{2\sigma_o\Delta t}z^{-2}}$$

$$\sigma_o = b_w/2M_w \qquad \alpha_o = \sqrt{(b_w/2M_w)^2 - K_w/M_w}$$

$$\therefore \quad \xi(x,n\Delta t) = \left[2e^{\sigma_o\Delta t}\cosh\alpha_o\Delta t\right]\xi(x,n\Delta t) - e^{2\sigma_o\Delta t}\xi(x,n\Delta t-2\Delta t)$$

$$+ \left[\frac{\Delta t}{\alpha_o M_w}\, e^{\sigma_o\Delta t}\sinh\alpha_o\Delta t\right] p(x,n\Delta t-\Delta t)$$

Case IV:

$$b_w^2 - 4M_wK_w = 0, \quad M_w \neq 0 \quad \text{(critically damped)}$$

$$\frac{\hat{\xi}(x,s)}{\hat{p}(x,s)} = \frac{1/M_w}{(s-\sigma_o)^2}$$

$$\rightarrow \frac{\hat{\xi}(x,z)}{\hat{p}(x,\hat{z})} = \frac{\Delta t^2}{M_w} \frac{e^{\sigma_o \Delta t} z^{-1}}{1-2e^{\sigma_o \Delta t} z^{-1} + e^{2\sigma_o \Delta t} z^{-2}}$$

$$\sigma_o = -b_w/2M_w$$

$$\therefore \xi(x,n\Delta t) = 2e^{\sigma_o \Delta t} \xi(x,n\Delta t-\Delta t) - e^{2\sigma_o \Delta t} \xi(x,n\Delta t-2\Delta t)$$

$$+ [\Delta t^2/M_w]\, e^{\sigma_o \Delta t} p(n\Delta t-\Delta t)$$

Case I, for which $M_w = 0$, is not realistic for the vocal-tract wall, and will not be considered further. For cases II, III, and IV, ξ and p satisfy the second order difference equation

$$\xi(x,n\Delta t) = c_o(x)p(x,n\Delta t-\Delta t) + c_1(x)\xi(x,n\Delta t-\Delta t)$$

$$+ c_2(x)\xi(n\Delta t-2\Delta t) \qquad (5.12)$$

at any position x. Therefore, at each point on the net

$$A_i^n = A_{oi}^n + S_{oi}^n \xi_i^n \qquad (5.13)$$

and

$$\xi_i^n = c_{oi} p_i^{n-1} + c_{1i} \xi_i^{n-1} + c_{2i} \xi_i^{n-2} \qquad (5.14)$$

where

$$A_{oi}^n = A_o(x,t)\Big|_{\substack{t=n\Delta t \\ x=i\Delta x}}$$

$$c_{\nu i} = c_\nu(x)\Big|_{x=i\Delta x} \quad \nu = 0,1,2$$

$$c_o = \begin{cases} \dfrac{\Delta t}{\omega_o M_w} e^{\sigma_o \Delta t} \sin\omega_o \Delta t & \text{if} \quad b_w^2 - 4M_w K_w < 0 \\ \dfrac{\Delta t}{\alpha_o M_w} e^{\sigma_o \Delta t} \sinh\alpha_o \Delta t & \text{if} \quad b_w^2 - 4M_w K_w > 0 \\ \dfrac{\Delta t^2}{M_w} e^{\sigma_o \Delta t} & \text{if} \quad b_w^2 - 4M_w K_w = 0 \end{cases}$$

$$c_1 = \begin{cases} 2e^{\sigma_o \Delta t} \cos\omega_o \Delta t & \text{if} \quad b_w^2 - 4M_w K_w < 0 \\ 2e^{\sigma_o \Delta t} \cosh\omega_o \Delta t & \text{if} \quad b_w^2 - 4M_w K_w > 0 \\ 2e^{\sigma_o \Delta t} & \text{if} \quad b_w^2 - 4M_w K_w = 0 \end{cases}$$

$$c_2 = -e^{2\sigma_o \Delta t}$$

and

$$\sigma_o = -b_w/2M_w$$

$$\omega_o = \sqrt{K_w/M_w - (b_w/2M_w)^2}$$

$$\alpha_o = \sqrt{(b_w/2M_w)^2 - K_w/M_w} \qquad (5.15)$$

The impulse-invariant simulation of the yielding duct wall was chosen for several reasons. First, the impulse-invariance technique has the property that the frequency axis is not distorted (warped) by the transformation. Also, since the yielding wall of the vocal tract is a highly damped mechanical system, its frequency response falls off essentially to zero as the Nyquist frequency appropriate for an audio signal is approached. Therefore, the effects of aliasing are negligible and the frequency response of the digital simulation is identical with that of the mechanical system when sampling frequencies appropriate for audio systems are used. Another reason for using an impulse-invariant simulation is of great practical significance, for it immeasurably simplifies the entire simulation. According to Eq. (5.14), ξ_i^n does not depend on p_i^n, therefore, it is not necessary to solve Eqs. (5.13) and (5.14) simultaneously with the partial-difference equations (5.6) and (5.7). Instead, the value of the wall displacement ξ_i^{n+1}, and hence the perturbed cross-sectional area A_i^{n+1} appearing in Eq. (5.7), can be determined

from only a knowledge of the past values of the acoustic pressure and wall displacement.

Equations (5.6, 5.7, 5.13, and 5.14) are the finite-difference analogs of the vocal-tract transmission Eqs. (2.18-2.21) and can be solved together with the boundary conditions at the mouth and glottis to obtain the values of the acoustic pressure and volume velocity at all points on the net. The finite-difference analogs of these boundary conditions will be discussed in Chapter VI and the solution to the complete system in Chapter VII.

CHAPTER VI

Discrete-Time Formulation of the Boundary Conditions

6.1 The Radiation Load

In Chapter III an expression for the radiation impedance at the mouth or nostrils was derived. Based on this expression, an integral equation relating the acoustic pressure and volume velocity at the mouth or nostrils was inferred. This equation is

$$U(t) = \frac{A}{\rho_o c}\left[\frac{1}{L_{rad}}\int_{-\infty}^{t} p(\alpha)d\alpha + \frac{p(t)}{R_{rad}}\right] \qquad (3.14)$$

where

$$R_{rad} = 128/9\pi^2$$

$$L_{rad} = 8a/3\pi c$$

$$A = \pi a^2.$$

In the first portion of this chapter, several digital simulations of Eq. (3.14), based on various integration rules, will be considered. It is shown in Appendix A that numerical integration of a linear constant-coefficient differential (or integral) equation corresponds to a mapping of the s plane to the z plane. We will thus investigate the effects on the radiation impedance (3.12) resulting from

various integration rules applied to Eq. (3.14) when the cross-sectional area is held fixed.

Consider first the simulation of Eq. (3.14) using an ideal integrator. The (simulated) radiation impedance is therefore simply that given by Eq. (3.12). The normalized acoustic impedance given by

$$z_n(j\omega) = Z_A(j\omega)/Z_o; \quad Z_o = \frac{\rho_o c}{A}$$

$$= \frac{j\omega\, L_{rad}}{1+j\omega\, L_{rad}/R_{rad}}$$

$$= \frac{j\omega \left(\frac{8a}{3\pi c}\right)}{1+j\omega \left(\frac{3\pi a}{16c}\right)} \qquad (6.1)$$

depends only on the ratio of the piston circumference ($2\pi a$) to the wavelength ($\lambda=2\pi c/\omega$), and not on the cross-sectional area itself. Because z_n depends only on this single parameter, $2\pi a/\lambda = \omega a/c$, we will compare the normalized acoustic impedances, z_n, for the different simulations rather than their analogous impedances Z_A. Figure 6.1 shows plots of the real and imaginary parts of the normalized acoustic radiation impedances corresponding to digital simulations of Eq. (3.14) according to various numerical integration rules when the cross-sectional area is held fixed. For purpose of comparison, a sampling rate of $1/\Delta t = 20$ kHz and a piston circumference of $2\pi a = 10c\Delta t$ has been assumed. This latter choice, for the piston circumference, is both an approximate upper limit for

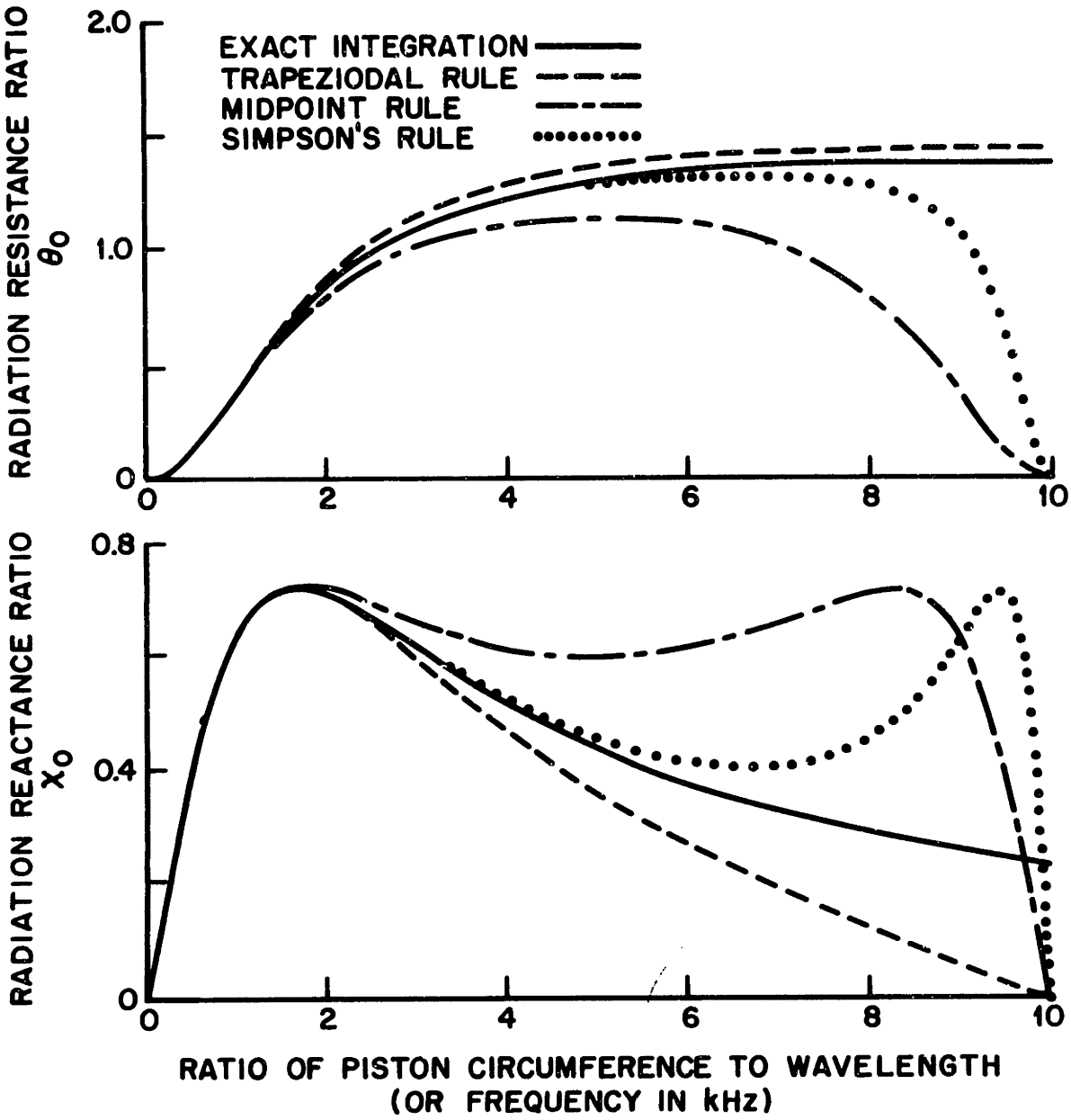


FIGURE 6.1

RADIATION IMPEDANCE RATIO ($Z_n = \theta_0 + jX_0$) FOR DIGITAL SIMULATIONS OF RADIATION LOAD BASED ON VARIOUS NUMERICAL INTEGRATION RULES

the human mouth and also allows the abscissas in Fig. 6.1 to be interpreted both as frequency in kilohertz and as the ratio of the piston circumference to wavelength.

The system functions used to compute these plots were easily obtained from the radiation impedance, Eq. (6.1) by letting $s = j\omega$ and using the mappings derived in Appendix A. The first integration rule considered is the midpoint rule, which corresponds to the mapping

$$s = \frac{1}{2\Delta t}\left(z - \frac{1}{z}\right) .$$

Referring to Fig. 6.1, it is clear that the midpoint rule does not give an adequate simulation of the radiation load except, perhaps, at very low frequencies. An important observation to be made from this plot is that the character of the error introduced by using the midpoint rule is such that the radiation resistance, and hence the loss, decreases erroneously at high frequencies thus amplifying the effects of high-frequency roundoff noise.

The next integration scheme considered is the trapezoid rule. This technique corresponds the mapping

$$s = \frac{2}{\Delta t}\,\frac{z-1}{z+1}$$

which is recognized as the bilinear transformation. The properties of this transformation are well known and it is commonly used in the design of recursive digital filters. In particular, it has the property of mapping the entire

analog frequency axis $(-\infty,\infty)$ in the s plane onto the unit circle $|z| = 1$ in the z plane by a tangential warping function. Thus, the digital frequency $\Omega = \pi^-$ corresponds to the analog frequency $\omega = +\infty$, $\Omega = -\pi^+$ corresponds to $\omega = -\infty$, and the frequency scale is warped according to the rule

$$\omega = \frac{2}{\Delta t} \tan\left(\frac{\Omega}{2}\right) .$$

Since the real part, the imaginary part, the magnitude, and the phase of the radiation impedance all monotonically approach constant values, the trapezoid rule appears well suited for simulating the radiation load.

Both the midpoint rule and the trapezoid rule are based on piecewise linear approximations to the integrand. It might be expected that better results can be obtained using a higher-order approximation to the integrand. Thus, we are led to investigate a simulation based on Simpson's (parabolic) rule. Simpson's rule can be shown to be equivalent to the mapping

$$s = \frac{3}{\Delta t} \frac{z^2-1}{z^2+4z+1}$$

Referring again to Fig. 6.1, it is seen that although Simpson's rule is a better approximation to an ideal integrator than either the midpoint rule or the trapezoid rule for frequencies less than about 7.5 kHz, the larger error introduced as the frequency approaches 10 kHz makes Simpson's rule undesirable for use in a simulation of the radiation load. As with the

midpoint rule, Simpson's rule also results in an erroneous decrease in the radiation resistance at high frequencies.

Based on the preceding analysis and the fact that the trapezoid rule tends to suppress high-frequency noise (whereas other common integration rules tend to amplify it),[37] the trapezoid rule was chosen for the digital simulation of the radiation load. The difference equation for this simulation is the following:

$$\rho_o c \left(\frac{L_{rad}}{A_o}\right)_o^n U_o^n - \left(\frac{L_{rad}}{R_{rad}} + \frac{\Delta t}{2}\right)_o^n p_o^n$$

$$= \rho_o c \left(\frac{L_{rad}}{A_o}\right)_o^{n-1} U_o^{n-1} - \left(\frac{L_{rad}}{R_{rad}} - \frac{\Delta t}{2}\right)_o^{n-1} p_o^{n-1} \qquad (6.2)$$

where

$$(\cdot)_1^n = (\cdot)_{x=i\Delta x}^{t=n\Delta t}$$

$$R_{rad} = 128/9\pi^2$$

$$L_{rad} = 8a/3\pi c$$

$$A_o = \pi a^2.$$

6.2 The Glottis

The vocal cords, together with the subglottal system, constitute the energy source for the vocal tract and thus provide the require boundary condition at $x = -\ell$ necessary to complete the model of the vocal tract. In recent years, there has been considerable effort in modeling

this system.[7-11] The model chosen for this simulation is the latest version of the Flanagan, Dudgeon, and Ishizaka Two-Mass Model. This model is discussed in detail in a recent paper by Flanagan and Ishizaka[11] and is formulated as the analogous electrical network shown in Fig. 6.2. The values of the equivalent circuit elements in the model are given by

$$R_{k1} = 0.19\ \rho_o / A_{g1}$$

$$R_{k2} = \frac{\rho_o \left[0.5 - \frac{A_{g2}}{A_1}\left(1 - \frac{A_{g2}}{A_1}\right)\right]}{A_{g2}^2}$$

$$R_{vi} = \frac{\mu \ell_g^2 d_i}{A_{gi}^3} \qquad (6.3)$$

$$i = 1,2$$

$$L_{gi} = \frac{\rho_o d_i}{A_{gi}}$$

where ℓ_g, d_1, and d_2 are the physical parameters shown in Fig. 6.3. A_1 (Flanagan's notation) is the cross-sectional area at the input ($x = -\ell$) to the vocal tract. A_{g1} and A_{g2} are the repective cross-sectional area specified by the displacements of the masses m_1 and m_2. These displacements are determined by solving the set of mechanical forcing relations for the oscillations of the two masses.[8,9,11]

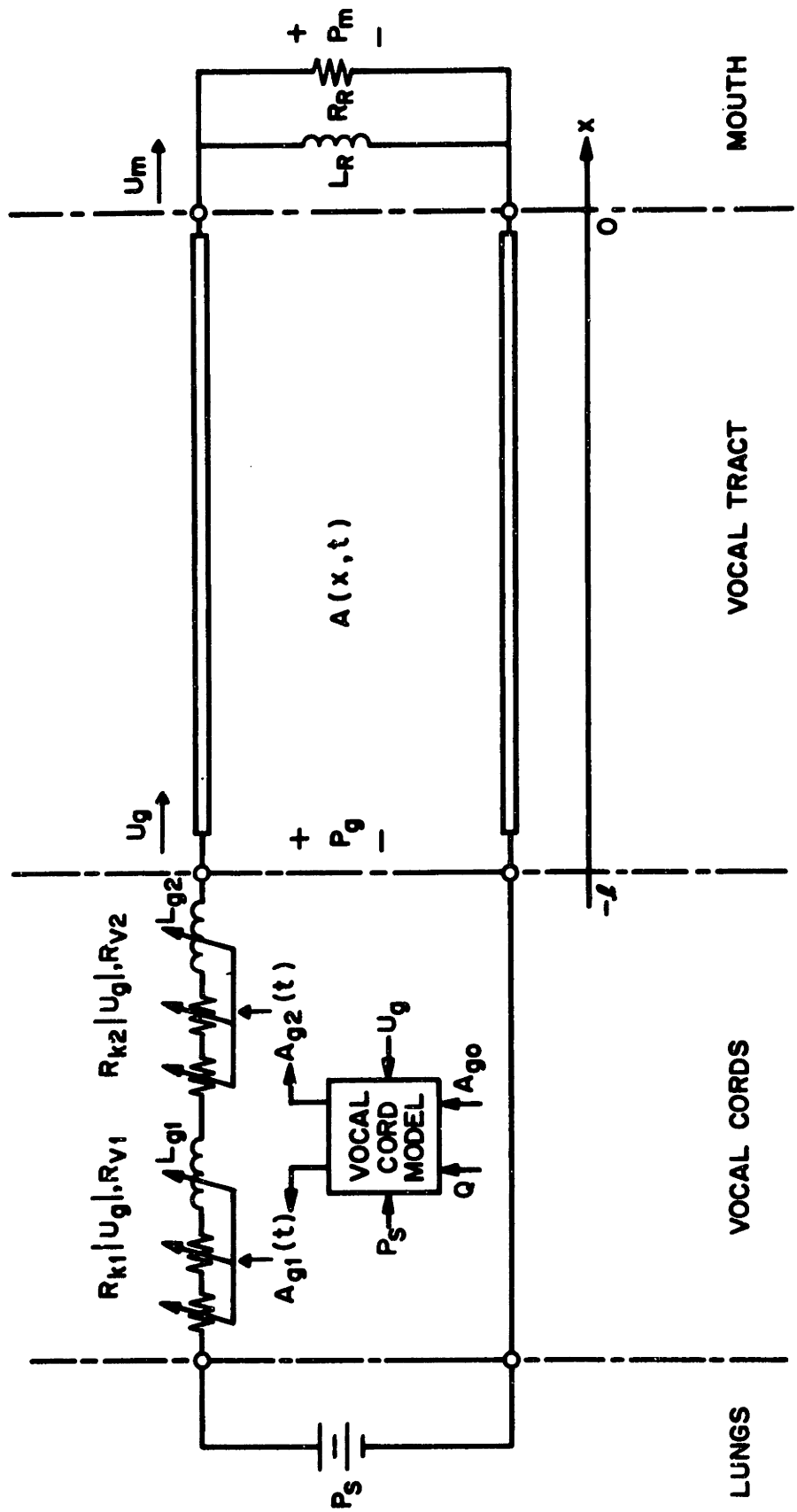


FIGURE 6.2
EQUIVALENT GLOTTAL NETWORK DRIVING VOCAL TRACT ANALOG

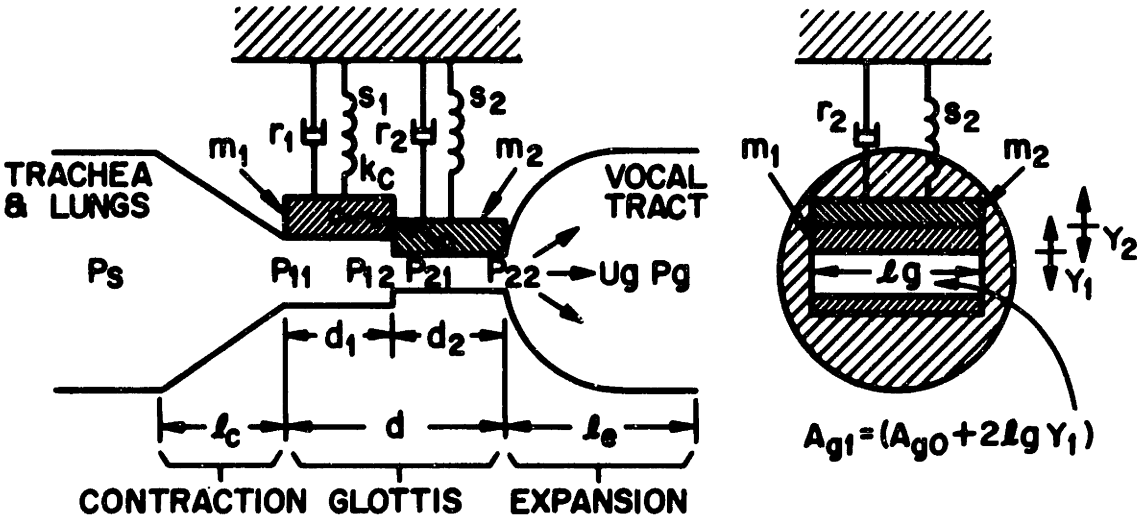


FIGURE 6.3
TWO-MASS MODEL

Computationally, this model may be separated into two distinct steps. As the first step, the mechanical forcing relations are solved, using the current state of the model to compute new values of the analogous circuit elements for the next iteration. The equations used in this step are identical to those given by Ishizaka and Flanagan.[11] As the second step of the computation, the equations for the analogous glottal network are solved to obtain the acoustic volume velocity U_g and pressure p_g at the "input" to the vocal tract. For this step, the equations given by Ishizaka and Flanagan must be modified slightly, because they model the vocal tract as a concatenation of T sections as shown in Fig. 6.4.

In the following chapter, it will be shown that from the vocal-tract transmission Eqs. (5.6) and (5.7), a linear constraint on U_g and p_g can be derived

$$p_g = R_{vt}U_g + p_{vt}$$

where the parameters R_{vt} and p_{vt} may be interpreted as an equivalent (time-varying) resistance and pressure, respectively, for the vocal tract. Thus, U_g is determined by solving the loop equation for the circuit shown in Fig. 6.5 rather than the one in Fig. 6.4. The differential equation for this circuit is

$$(R_{k1}+R_{k2})|U_g|U_g + (R_{v1}+R_{v2})U_g + (L_{g1}+L_{g2})\frac{dU_g}{dt} + R_{vt}U_g + p_{vt} - P_s = 0 \tag{6.4}$$

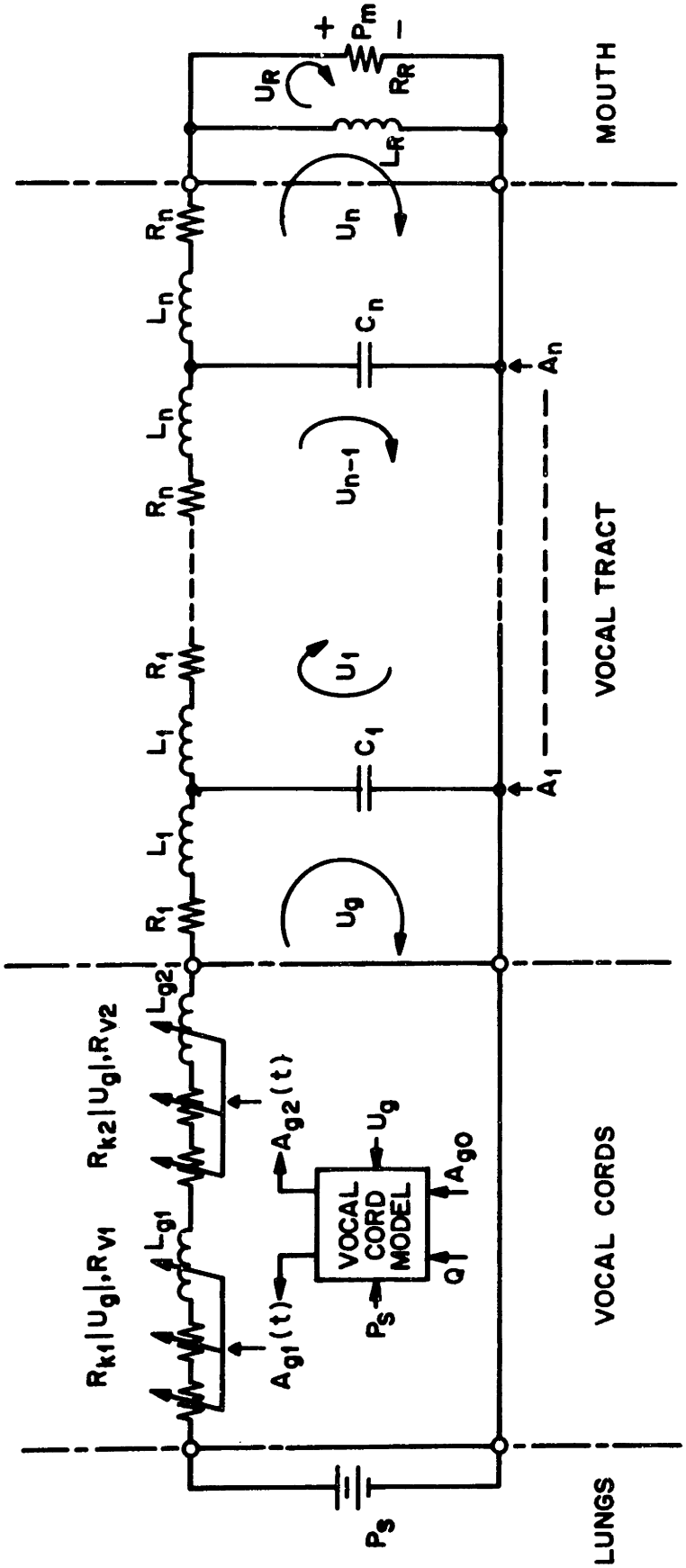


FIGURE 6.4
EQUIVALENT GLOTTAL NETWORK DRIVING T-SECTION LUMPED PARAMETER ANALOG OF VOCAL TRACT

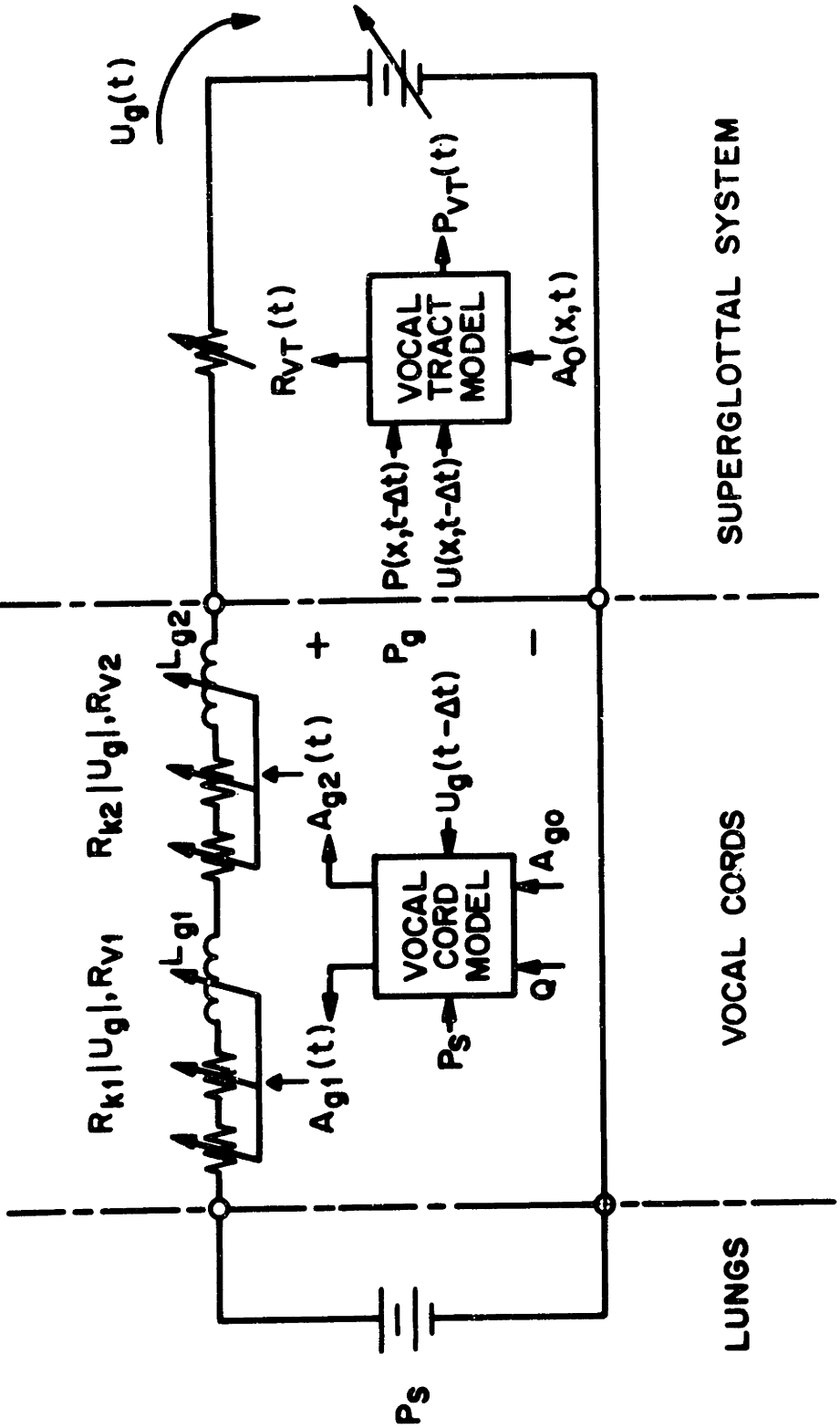


FIGURE 6.5
EQUIVALENT "GLOTTAL LOOP" NETWORK FOR CALCULATING $U_g(t)$

and the finite-difference simulation is

$$(R_{k1}^n+R_{k2}^n)|U_g^n|U_g^n + (R_{v1}^n+R_{v2}^n)U_g^n + \frac{L_{g1}^n+L_{g2}^n}{\Delta t}(U_g^n-U_g^{n-1}) + R_{vt}^n U_g^n + p_{vt}^n - P_s^n = 0 \tag{6.5}$$

This finite-difference formulation was chosen so that the modification required to adapt Ishizaka's model as a boundary condition for Eqs. (5.6) and (5.7) was minimized. Therefore, the formulation of Eq. (6.5) is consistent with the formulation used by Flanagan and Landgraf[8] and Ishizaka and Flanagan[11] based on replacing derivatives with backward differences.

It is important to note that the glottal loop Eq. (6.5) is nonlinear and must be solved simultaneously with partial-difference equations for the vocal tract, because the equivalent resistance R_{vt} and pressure p_{vt} are not known a priori. Exactly how this is accomplished will be discussed in the following chapter.

CHAPTER VII

Implementation of the Finite-Difference Equations

In the foregoing chapters, a set of finite-difference equations has been formulated which models acoustic wave propagation in the vocal tract and accounts for the effects of radiation at the lips and a self-oscillating source at the glottis. We now proceed to write down these equations in a systematic manner and give a procedure for solving the entire system.

Referring to the net shown in Fig. 7.1, assume that p and U are known at all net points for which n is less than some integer n_o and p and U are both zero for $n < 0$ (initial rest). We will now determine p and U at all points along the row $n = n_o$.

From the forcing relations given by Ishizaka and Flanagan[11] the values of the elements in the analogous network for the glottis, at time $n = n_o$, can be determined from the past values of the volume velocity, U_g, through the glottis and the past state of the model. The perturbation to the cross-sectional area of the vocal tract, $\delta A = S_o \xi$, at each net point in the row $n = n_o$ can be determined from Eq. (5.14):

$$\xi_i^n = c_{oi} p_i^{n-1} + c_{1i} \xi_i^{n-1} + c_{2i} \xi_i^{n-2} \qquad (5.14)$$

where the c_{vi} are given by Eqs. (5.15).

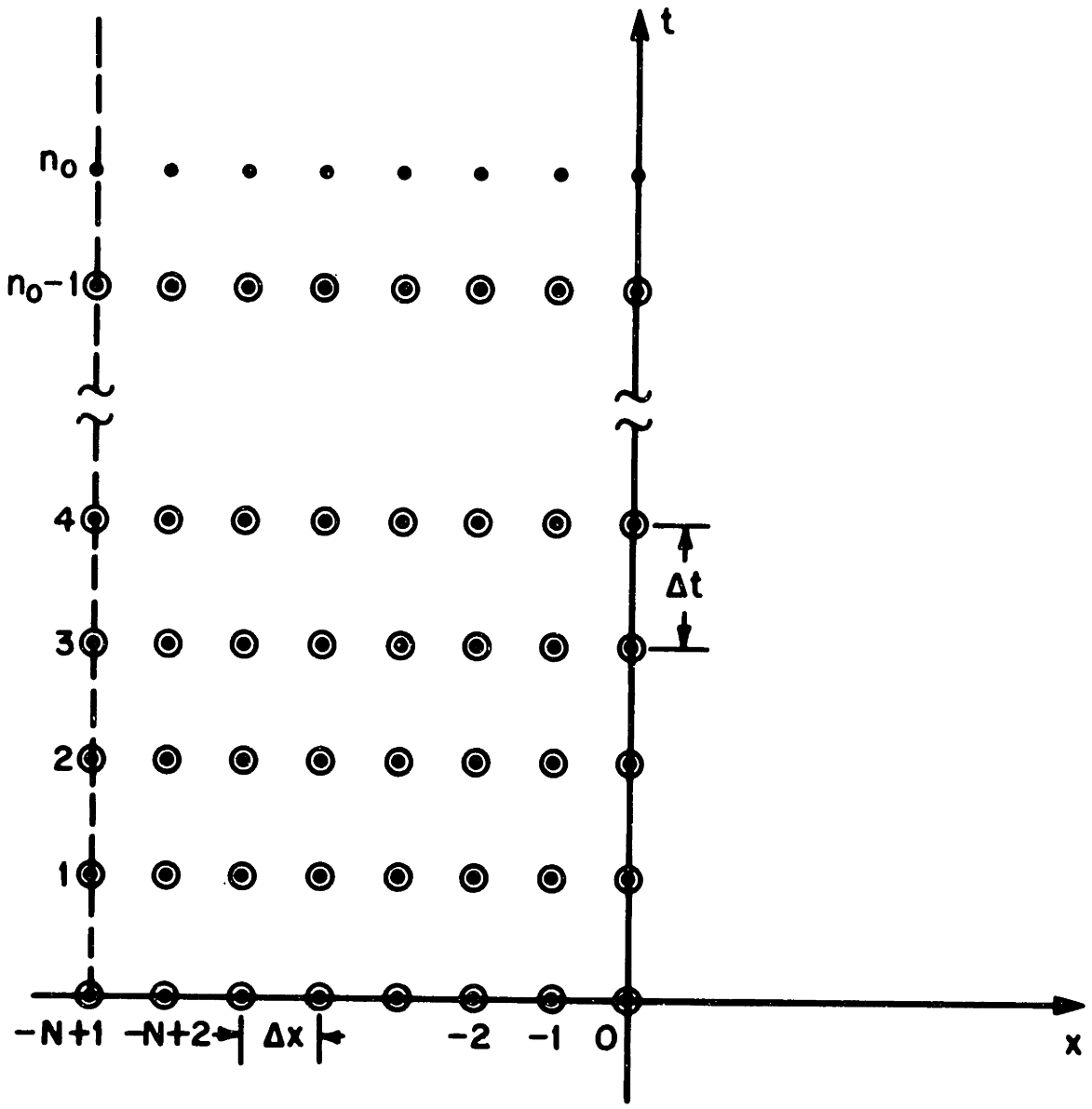


FIGURE 7.1

APPROXIMATION TO CHARACTERISTIC NET FOR FINITE DIFFERENCE SIMULATION. CIRCLED POINTS INDICATE POINTS AT WHICH SOLUTION IS ASSUMED TO BE KNOWN.

The system of linear algebraic equations to be solved for the values of U and p each of the net points on the row $n = n_o$ can now be written down systematically as follows. Starting at x = 0 we use Eq. (6.2) to write the boundary condition at the lips:

$$h_{11}^n U_o^n + h_{12}^n p_o^n = g_{11}^{n-1} U_o^{n-1} + g_{12}^{n-1} p_o^{n-1} = b_1^{n-1} \qquad (7.1)$$

where

$$h_{11}^n = \rho_o c \left(\frac{L_{rad}}{A_o}\right)_o^n \qquad g_{11}^n = h_{11}^n$$

$$h_{12}^n = -\left(\frac{L_{rad}}{R_{rad}} + \frac{\Delta t}{2}\right)_o^n \qquad g_{12}^n = -\left(\frac{L_{rad}}{R_{rad}} - \frac{\Delta t}{2}\right)_o^n$$

We next write down the (2N-2) partial-difference equations (5.6-5.7):

$$Z_i^n U_i^n + p_i^n + Z_{i-1}^n U_{i-1}^n - p_{i-1}^n$$

$$= Z_i^{n-1} U_i^{n-1} - p_i^{n-1} + Z_{i-1}^{n-1} U_{i-1}^{n-1} + p_{i-1}^{n-1}$$

and (7.2)

$$U_i^n + Y_i^n p_i^n - U_{i-1}^n + Y_{i-1}^n p_{i-1}^n$$

$$= -U_i^{n-1} + Y_i^{n-1} p_i^{n-1} + U_{i-1}^{n-1} + Y_{i-1}^{n-1} p_{i-1}^{n-1}$$

$$- \frac{\Delta x}{\Delta t}\left[A_i^n - A_i^{n-1} + A_{i-1}^n - A_{i-1}^{n-1}\right]$$

(7.3)

for

$$i = 0,-1,\ldots,-N+2$$

where the A_i^n are determined from Eqs. (5.13-5.14). The boundary condition at the glottis ($i = -N+1$) is the nonlinear loop equation (6.5) derived from Ishizaka's equivalent circuit for the glottis, which can be rewritten in the form

$$R_k^n|U_g^n|U_g^n + [R_v^n+L_g^n/\Delta t+R_{vt}^n]U_g^n + [p_{vt}^n-P_s^n-L_g^nU_g^{n-1}/\Delta t] = 0 \qquad (7.4)$$

where

$$U_g^n = U_{-N+1}^n$$

$$R_k = R_{k1} + R_{k2}$$

$$R_v = R_{v1} + R_{v2}$$

$$L_g = L_{g1} + L_{g2}$$

and R_{vt}^n and p_{vt}^n are the equivalent resistance and pressure looking into the vocal tract from the glottis at time $t = n\Delta t$. R_{vt} and p_{vt} are notknown a priori, but should depend only on the present state of the vocal tract. In fact, it will be shown shortly that R_{vt} and p_{vt} can determined from the first

2N-1 equations (7.1-7.3) which have been written down, i.e., the equations governing the superglottal system. For the present, however, we will assume a general linear condition at the glottis of the form

$$h^n_{2N,2N-1}U^n_{-N+1} + h^n_{2N,2N}p^n_{-N+1} = b^n_{2N} \tag{7.5}$$

and proceed with the formulation as if the coefficients of Eq. (7.5) were known.

Equations (7.1-7.5) can be written compactly in matrix form as

$$\underline{H}^{(n)}\underline{\zeta}^{(n)} = \underline{b}^{(n)} \tag{7.6}$$

where

$$\underline{H}^{(n)} = \begin{bmatrix}
h^n_{11} & h^n_{12} & & & & & & & & \\
Z^n_o & 1 & Z^n_{-1} & -1 & & & & & & \\
1 & Y^n_o & -1 & Y^n_{-1} & & & & & & \\
 & & Z^n_{-1} & 1 & Z^n_{-2} & -1 & & & & \\
 & & 1 & Y^n_{-1} & -1 & Y^n_{-2} & & & & \\
 & & & & \cdot & \cdot & & & & \\
 & & & & & \cdot & \cdot & & & \\
 & & & & & & \cdot & \cdot & & \\
 & & & & & & Z^n_{-N+2} & 1 & Z^n_{-N+1} & -1 \\
 & & & & & & 1 & Y^n_{-N+2} & -1 & Y^n_{-N+1} \\
 & & & & & & & & h_{2N,2N-1} & h_{2N,2N}
\end{bmatrix}$$

$$\underline{\zeta}^{(n)} = [U^n_o, p^n_o, U^n_{-1}, \ldots, U^n_{-N+1}, p^n_{-N+1}]^T$$

and $\underline{b}^{(n)}$ is given by

$$\underline{b}^{(n)} = \underline{G}^{(n-1)}\underline{\zeta}^{(n-1)} + \underline{D}[\underline{A}^{(n)} - \underline{A}^{(n-1)}] \tag{7.7}$$

where

$$\underline{G}^{(n)} = \begin{bmatrix} g^n_{11} & g^n_{12} & & & & & & & \\ Z^n_o & -1 & Z^n_{-1} & 1 & & & & & \\ -1 & Y^n_o & 1 & Y^n_{-1} & & & & & \\ & & Z^n_{-1} & -1 & Z^n_{-2} & 1 & & & \\ & & -1 & Y^n_{-1} & 1 & Y^n_{-2} & & & \\ & & & & \ddots & & \ddots & & \\ & & & & & & Z^n_{-N+2} & -1 & Z^n_{-N+1} & 1 \\ & & & & & & -1 & Y^n_{-N+2} & 1 & Y^n_{-N+1} \\ & & & & & & & & g^n_{2N,2N-1} & g^n_{2N,2N} \end{bmatrix}$$

$$\underline{A}^{(n)} = [A^n_o, A^n_{-1}, A^n_{-2}, \ldots, A^n_{-N+1}]^T$$

and $\underline{D}$ is the $(2N \times N)$ matrix $\underline{D} = \frac{\Delta x}{\Delta t}\,[d_{ij}]$ with elements

$$d_{ij} = \begin{cases} 1 & i=2k+1;\ j=k,k+1;\ k=1,2,\ldots,N-1 \\ 0 & \text{otherwise} \end{cases}$$

Equation (7.6), together with the nonlinear glottal-loop equation (7.4) can be solved efficiently using the algorithm described in Appendix B. This algorithm provides a method for solving the system of equations

$$\underline{H\zeta} = \underline{b} \qquad (7.8)$$

where $\underline{H}$ is a sparse matrix of the form given in Eq. (7.6). The solution is obtained by decomposing the (2N×2N) matrix $\underline{H}$ as the product of a lower triangular matrix $\underline{L}$ and an upper triangular matrix* $\underline{U}$. The matrices $\underline{L}$ and $\underline{U}$ are themselves sparse, with their only nonzero elements in triangular blocks along their main diagonals, and are of the form:

$$\underline{L} = \begin{bmatrix} 1 & & & & & & & \\ \ell_{21} & 1 & & & & & & \\ \ell_{31} & \ell_{32} & 1 & & & & & \\ & & \ell_{43} & 1 & & & & \\ & & \ell_{53} & \ell_{54} & 1 & & & \\ & & & & & \ddots & & \\ & & & & & & \ddots & \\ & & & & & & & \ddots \\ & & & & \ell_{2N-1,2N-3} & \ell_{2N-1,2N-2} & 1 & \\ & & & & & & \ell_{2N,2N-1} & 1 \end{bmatrix}$$

$$\underline{U} = \begin{bmatrix} u_{11} & u_{12} & & & & & & \\ & u_{22} & u_{23} & u_{24} & & & & \\ & & u_{33} & u_{34} & & & & \\ & & & u_{44} & u_{45} & u_{46} & & \\ & & & & u_{55} & u_{56} & & \\ & & & & & \ddots & & \\ & & & & & & \ddots & \\ & & & & & & & \ddots \\ & & & & & & u_{2N-1,2N-1} & u_{2N-1,2N} \\ & & & & & & & u_{2N,2N} \end{bmatrix}$$

*Not to be confused with the volume velocity U(x,t).

Once this decomposition is known, if we let

$$\underline{\eta} = \underline{U}\underline{\zeta} \tag{7.9}$$

then

$$\underline{H}\underline{\zeta} = (\underline{L}\ \underline{U})\underline{\zeta} = \underline{b} \tag{7.10}$$

becomes

$$\underline{L}\underline{\eta} = \underline{b} \tag{7.11}$$

Since $\underline{L}$ is lower triangular, Eq. (7.11) is solved trivially for $\underline{\eta}$; and since U is upper triangular, once $\underline{\eta}$ has been found, Eq. (7.9) is solved trivially for $\underline{\zeta}$.

The problem is, of course, that the (nonezero) elements in the last row of $\underline{H}$ and the last component of $\underline{b}$ are not known, since they define a linear boundary condition at the glottis. It will now be shown how to use the nonlinear condition given by Eq. (7.4) in lieu of such a linear condition. According to the algorithm of Appendix B, the elements of $\underline{L}$, $\underline{U}$, and $\underline{\eta}$ are computed "row-wise" so that any element in a particular row depends only on elements in the corresponding row of $\underline{H}$, and previously computed elements of $\underline{L}$, $\underline{U}$, and $\underline{\eta}$. Thus, the only elements of $\underline{L}$, $\underline{U}$, and $\underline{\eta}$ which depend on the last row of $\underline{H}$ and the last component of $\underline{b}$ are those in the last rows of $\underline{L}$ and $\underline{U}$ (viz., $\ell_{2N,2N-1}, u_{2N,2N}$) and the last component of $\underline{\eta}$ (viz., η_{2N}).

These elements appear only in the two equations

$$\ell_{2N,2N-1}\eta_{2N} = b_{2N}$$

and

$$u_{2N,2N}\zeta_{2N} = \eta_{2N}.$$

Once ζ_{2N} ($\zeta_{2N} = p_{-N+1} = p_g$) is determined, the remaining components of the solution vector $\underline{\zeta}$ can be found according to the algorithm of Appendix B.

The (2N-1) th equation in the system

$$\underline{U\zeta} = \underline{\eta} \tag{7.9}$$

is

$$u_{2N-1,2N-1}\zeta_{2N-1} + u_{2N-1,2N}\zeta_{2N} = \eta_{2N-1} \tag{7.12}$$

where $u_{2N-1,2N-1}$, $u_{2N-1,2N}$, and η_{2N-1} are determined from the algorithm of Appendix B. If $u_{2N-1,2N} \neq 0$, define

$$p_{vt}^{(n)} = \eta_{2N-1}/u_{2N-1,2N} \tag{7.13}$$

and

$$R_{vt}^{(n)} = -u_{2N-1,2N-1}/u_{2N-1,2N} \tag{7.14}$$

so that Eq. (7.12) becomes

$$\zeta_{2N} = R_{vt}^{(n)} \zeta_{2N-1} + p_{vt}^{(n)} \tag{7.15}$$

Finally, substituting

$$\zeta_{2N-1} = U_g^n$$

$$\zeta_{2N} = p_g^n$$

gives

$$p_g^n = R_{vt}^{(n)} U_g^n + p_{vt}^{(n)} \tag{7.16}$$

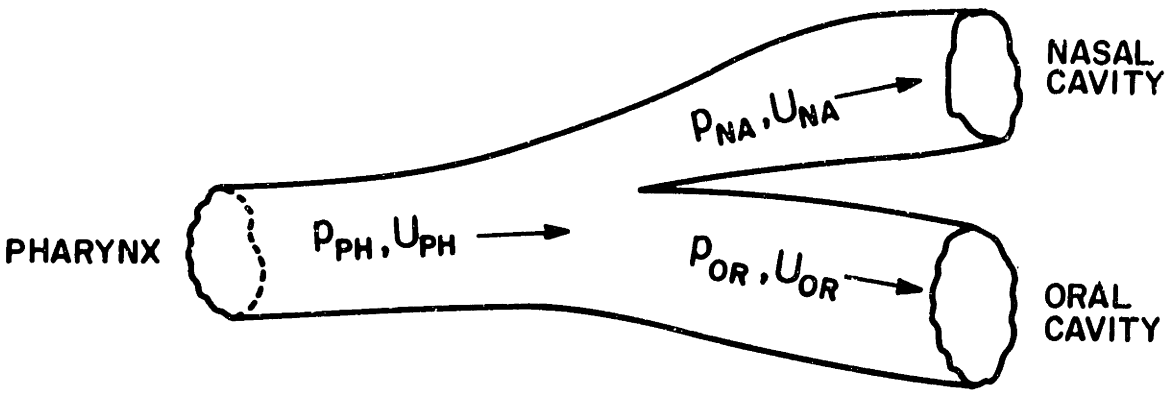
Equation (7.16) defines an equivalent linear circuit for the vocal tract as seen by the glottis at time $t = n\Delta t$. The equivalent pressure source is $p_{vt}^{(n)}$ and represents the stored energy in the vocal tract. The equivalent resistance is $R_{vt}^{(n)}$ and depends only on the vocal-tract geometry (Note, if $u_{2N-1,2N} = 0$ then the equivalent circuit for the vocal tract is simply the current source $U_g = \eta_{2N-1}/u_{2N,2N}$).

The equivalent network defined by Eq. (7.16) is combined with Ishizaka's equivalent network for the glottis to give the network shown in Fig. 6.5. The nonlinear loop equation (7.4) can then be solved for U_g; and p_g can be determined from Eq. (7.16). The remaining elements of the solution vector can then be determined according to the algorithm of Appendix B.

The entire procedure is repeated for succeeding time samples to determine recursively the state of the system for any $n > 0$. Notice that the matrices $\underline{H}$ and $\underline{G}$ depend only on the geometry of the vocal tract. If the geometry of the vocal tract at a particular time $t = n\Delta t$ remains unchanged from its configuration at time $t = (n-1)\Delta t$ then the matrices $\underline{H}$, $\underline{L}$, $\underline{U}$, and $\underline{G}$ also remain unchanged and it is therefore not necessary to recompute them for that time step (this property is especially useful when simulating the nasal tract, which is time invariant).

Thus far, only the implementation of the vocal-tract proper has been considered. We now describe an implementation for the nasal tract. The complete system (including the nasal tract) is modeled as three acoustic tubes joined in a "Y" configuration as shown in Fig. 7.2. The boundary conditions at the lips and glottis remain unchanged; and the boundary condition at the nostrils is again the radiation load as simulated by Eq. (7.1). The boundary condition at the junction of the three acoustic tubes follows directly from the physics of the situation viz., the pressure across the junction must be continuous and the net flow of mass into the junction must be zero. These conditions imply

$$p_{PH} = p_{OR} = p_{NA} \tag{7.17}$$



$$U_{PH} = U_{OR} + U_{NA}$$
$$P_{PH} = P_{OR} = P_{NA}$$

FIGURE 7.2

JUNCTION OF PHARYNGAL, ORAL, AND NASAL CAVITIES

and

$$U_{PH} = U_{OR} + U_{NA} \tag{7.18}$$

where the above quantities are measured at the junction of the pharyngal, oral, and nasal cavities as indicated in Fig. 7.2.

The solution to the system of three acoustic tubes follows in a similar manner to the previous case of a single tube. At each time step $t = n\Delta t$, two systems of equations, one for the oral cavity and the other for the nasal tract, are formulated as in Eq. (7.6). Again, for both of these systems the boundary condition at the left-hand ends of the tubes (in this case, the junction) are not known a priori. The decomposition algorithm is applied to both systems and equivalent circuits are determined as before. These circuits model the oral and nasal cavities as seen looking from the bifrication towards the mouth and nose; and together with the constraints at the junction provided by Eqs. (7.17-7.18) can be readily solved to yield a single equation of the form

$$h_{11}U_{PH} + h_{12}p_{PH} = b_1 \tag{7.19}$$

Equation (7.19) is now used as the boundary condition at the right-hand end of the pharynx. With Eq. (7.19) as the first equation, the system of equations for the pharynx

can be formulated and solved as in the previous case of a single acoustic tube with Ishizaka's equivalent circuit for the glottis used as the left-hand boundary condition. Once the solution to system of equations for the pharynx has been obtained, the equivalent circuits for the oral and nasal cavities together with Eqs. (7.17-7.18) can be used to determine p_{OR}, U_{OR}, p_{NA}, and U_{NA}. Finally, using these four quantities, the solution vectors for the oral and nasal cavities can be obtained in the same manner as the solution to the single-tube problem was obtained after the glottal-loop equation was solved.

The entire procedure is then repeated for succeeding time samples to determine recursively the state of the system for any $n > 0$. The same comment, concerning the necessity to recompute the factors of the coefficient matrices only when the geometry changes, still applies. However, notice that changing the geometry of the oral cavity changes the coefficient matrices for both the oral cavity and the pharynx but not for the nasal tract; whereas changing the geometry of the pharynx changes only the coefficient matrix for the pharynx. Also, changing the cross-sectional area of the nasal cavity at the junction does not require recomputing all of the elements in the matrix factors of the coefficient matrix for the nasal tract, but merely the last few elements of the factors.

CHAPTER VIII

Experimental Results and Concluding Remarks

8.1 The Computer Simulation

The system of difference equations formulated in Chapter VII was programmed in Fortran on a medium-size general-purpose computer. Specifically, the machine used here was one of the Honeywell DDP-516 computers in the Acoustics Research Department of the Bell Telephone Laboratories, Murray Hill, New Jersey.

Two series of experiments were performed to investigate the operation of the simulation. For the first series, the vocal-cord model was again replaced with an ideal (i.e., infinite impedance) volume-velocity source (as in Chapter IV). In this case, however, the source supplied a unit sample rather than a sinusoid, and the output of the simulation was taken as the acoustic pressure rather than the acoustic volume velocity. Thus, the unit-sample response of the model is obtained directly and the frequency response is obtained by taking a "large enough" discrete Fourier transform (DFT) of the unit-sample response.

The result of one such experiment is given in Fig. 8.1 which shows the unit-sample response and its DFT for a vocal-tract configuration 17.5 cm in length with a uniform cross section of 5 cm^2. In this case $\Delta t = 1/20$ msec and $\Delta x = 1.75$ cm, which satisfy the condition $\Delta x = c\Delta t$. The wall parameters

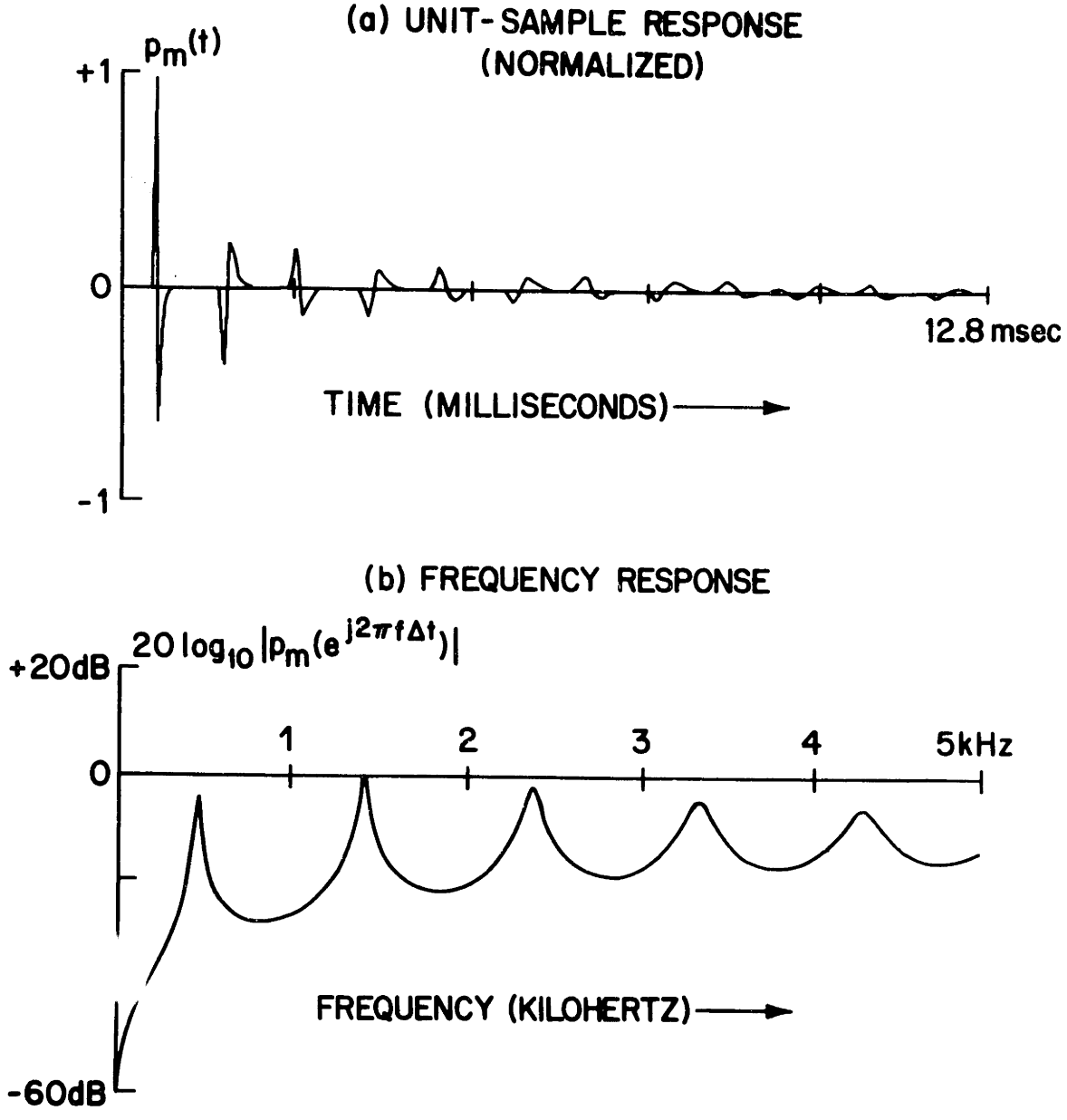


FIGURE 8.1

MOUTH SOUND PRESSURE AND ITS DFT FOR VOCAL-TRACT SIMULATION WITH UNIFORM CROSS SECTION ($A=5cm^2$) DRIVEN BY IDEAL VOLUME-VELOCITY SOURCE

are again those given by Flanagan.[19] The frequency response was obtained using a 1024 point (51.2 msec) DFT with a Hamming win' w. When comparing Fig. 8.1(b) with the analogous result obtained in the frequency domain, shown in Fig. 3.5, there are two points that must be kept in mind. First, Fig. 8.1(b) corresponds to the ratio $|P_m/U_g|$ rather than the ratio $|U_m/U_g|$. These two ratios are, of course, related by the radiation impedance at the lips. Second, Fig. 8.1(b) was obtained by taking the DFT of the unit-sample response of the discrete-variable simulation and is therefore an aliased version of the spectrum of the continuous variable model defined by the partial differential equation of Chapter II.

For the second series of experiments, the vocal-tract model was driven by the two-mass model of the vocal cords as described in Chapter VII and the simulation was run to obtained synthetic speech sounds. The data used to obtain the results discussed here were the same as those used in Chapter IV, viz., the area data obtained from Fant[20] for the Russian vowels /a/,/e/,/i/,/o/,/u/, quantized to the nearest $\sqrt[3]{2}$, and the wall parameters given by Flanagan.[19] The additional parameters required for the vocal-cord oscillator were those used by Ishizaka and Flanagan.[11]

Figures 8.2 through 8.7 show data obtained from the simulation of the five Russian vowels. Figures 8.2 through 8.6 each consist of four plots: plot (a) shows the acoustic pressure output at the lips when the vocal-tract

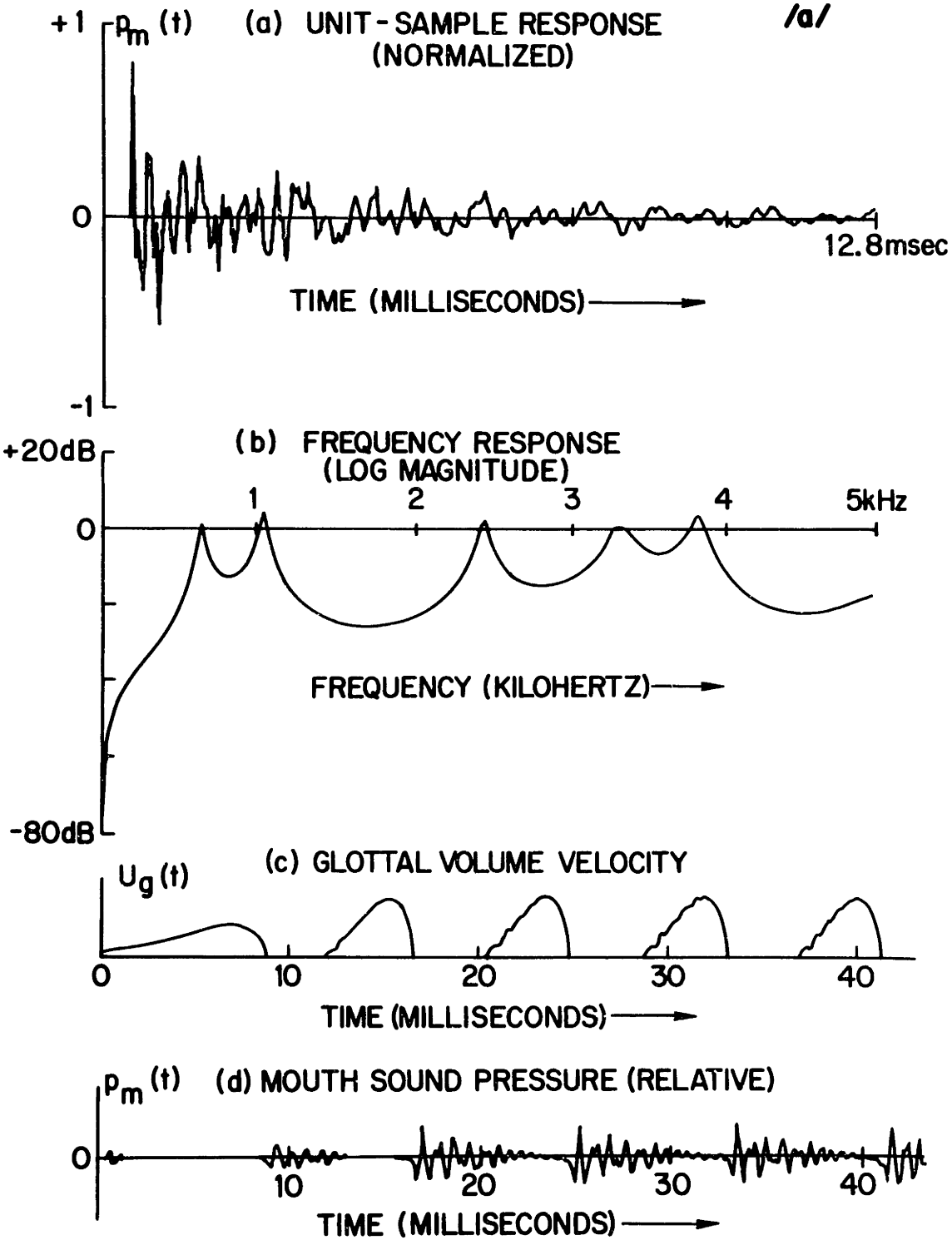

+1
p_m (t)
(a) UNIT-SAMPLE RESPONSE
(NORMALIZED)
/a/
0
12.8msec
TIME (MILLISECONDS)
-1
+20dB
(b) FREQUENCY RESPONSE
(LOG MAGNITUDE)
1
2
3
4
5kHz
0
FREQUENCY (KILOHERTZ)
-80dB
U_g (t)
(c) GLOTTAL VOLUME VELOCITY
0
10
20
30
40
TIME (MILLISECONDS)
p_m (t)
(d) MOUTH SOUND PRESSURE (RELATIVE)
0
10
20
30
40
TIME (MILLISECONDS)

FIGURE 8.2

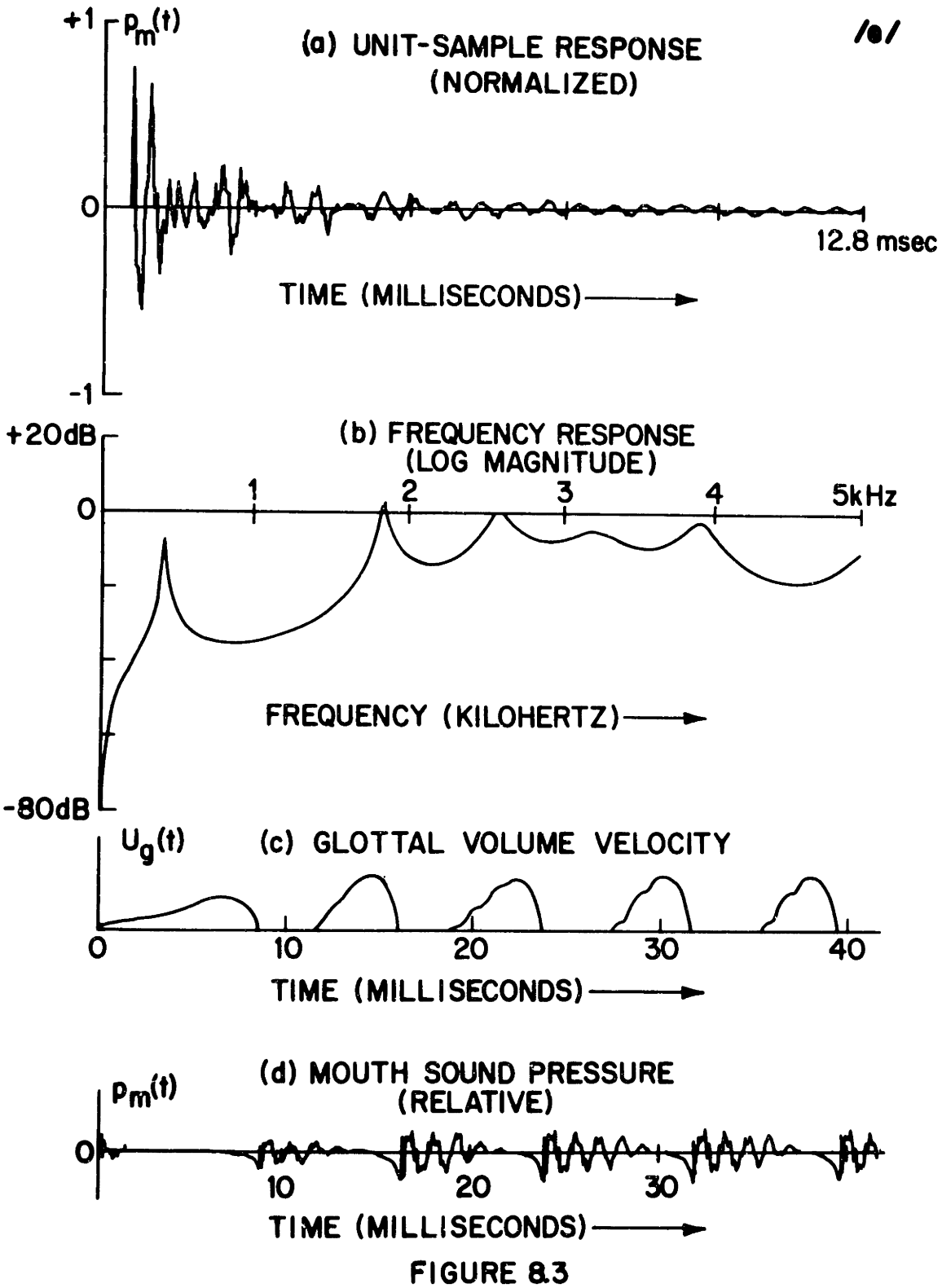

+1
$p_m(t)$
(a) UNIT-SAMPLE RESPONSE
(NORMALIZED)
/e/
0
12.8 msec
TIME (MILLISECONDS)
-1
+20dB
(b) FREQUENCY RESPONSE
(LOG MAGNITUDE)
0
1
2
3
4
5kHz
FREQUENCY (KILOHERTZ)
-80dB
$U_g(t)$
(c) GLOTTAL VOLUME VELOCITY
0
10
20
30
40
TIME (MILLISECONDS)
$p_m(t)$
(d) MOUTH SOUND PRESSURE
(RELATIVE)
0
10
20
30
TIME (MILLISECONDS)

FIGURE 8.3

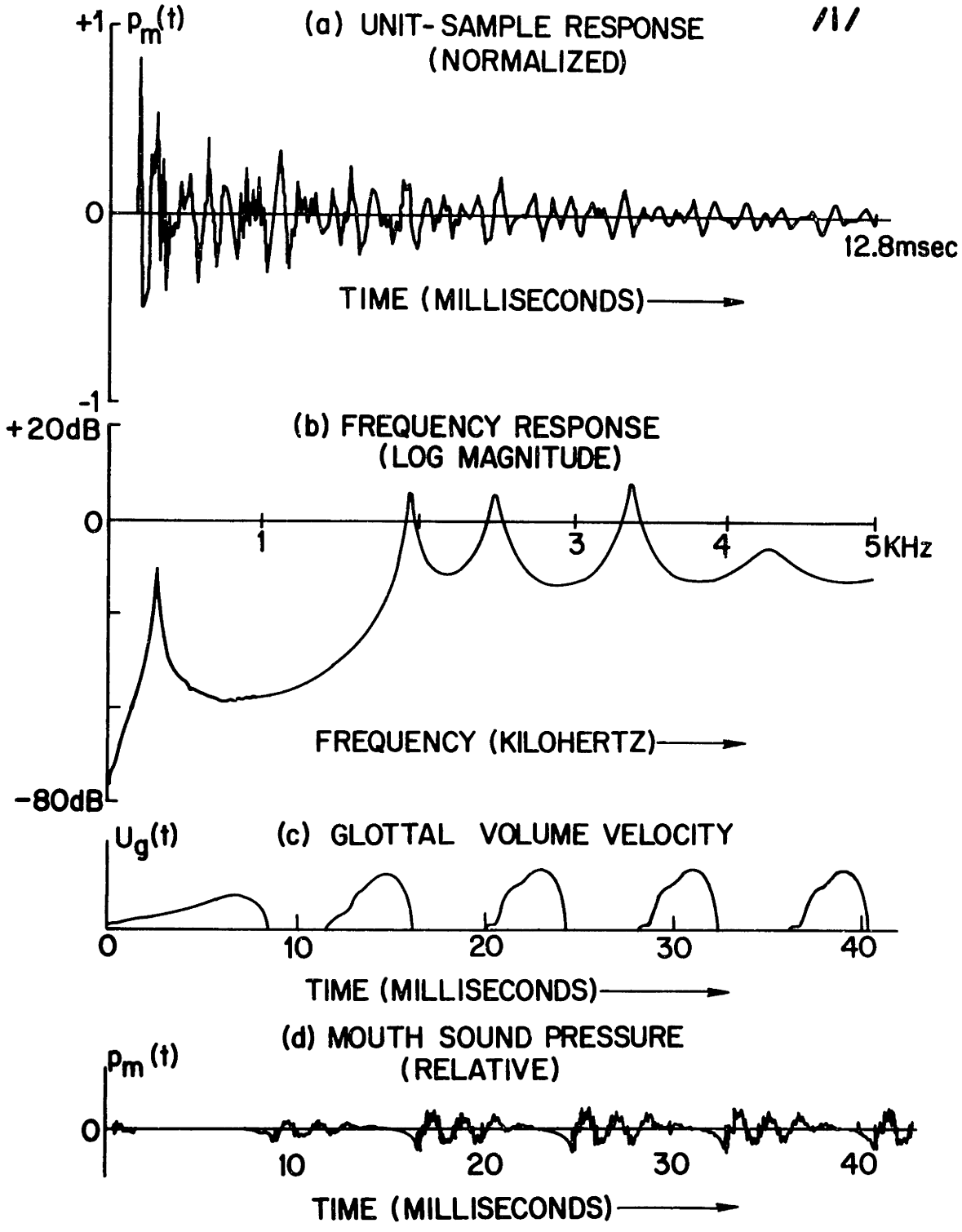

+1
$p_m(t)$
(a) UNIT-SAMPLE RESPONSE
(NORMALIZED)
0
12.8msec
TIME (MILLISECONDS)
-1
+20dB
(b) FREQUENCY RESPONSE
(LOG MAGNITUDE)
0
1
3
4
5KHz
FREQUENCY (KILOHERTZ)
-80dB
$U_g(t)$
(c) GLOTTAL VOLUME VELOCITY
0
10
20
30
40
TIME (MILLISECONDS)
(d) MOUTH SOUND PRESSURE
(RELATIVE)
$p_m(t)$
0
10
20
30
40
TIME (MILLISECONDS)

FIGURE 8.4

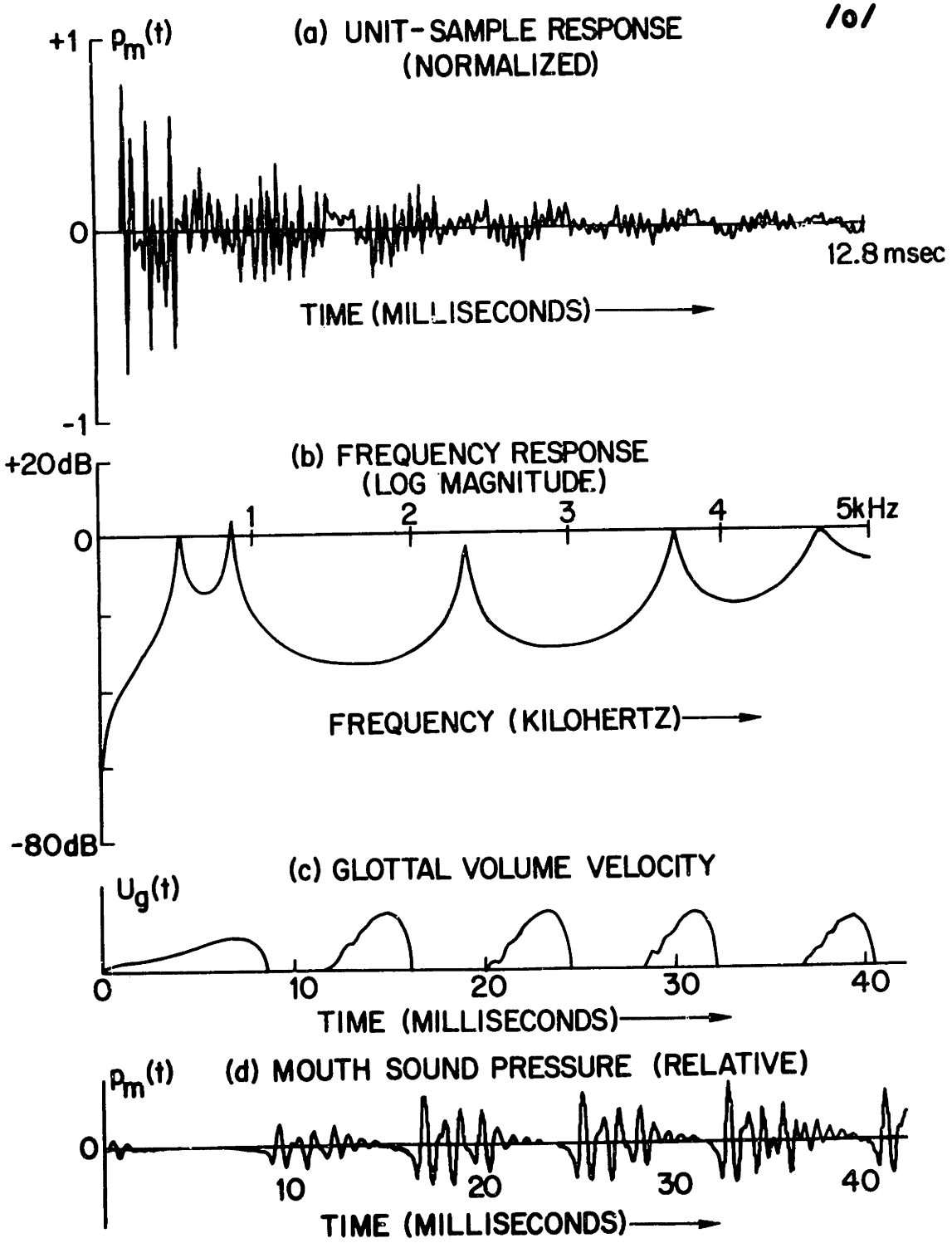

101
+1
$p_m(t)$
(a) UNIT-SAMPLE RESPONSE
(NORMALIZED)
0
12.8 msec
TIME (MILLISECONDS)
-1
+20dB
(b) FREQUENCY RESPONSE
(LOG MAGNITUDE)
1
2
3
4
5kHz
0
FREQUENCY (KILOHERTZ)
-80dB
(c) GLOTTAL VOLUME VELOCITY
$U_g(t)$
0
10
20
30
40
TIME (MILLISECONDS)
$p_m(t)$
(d) MOUTH SOUND PRESSURE (RELATIVE)
0
10
20
30
40
TIME (MILLISECONDS)

FIGURE 8.5

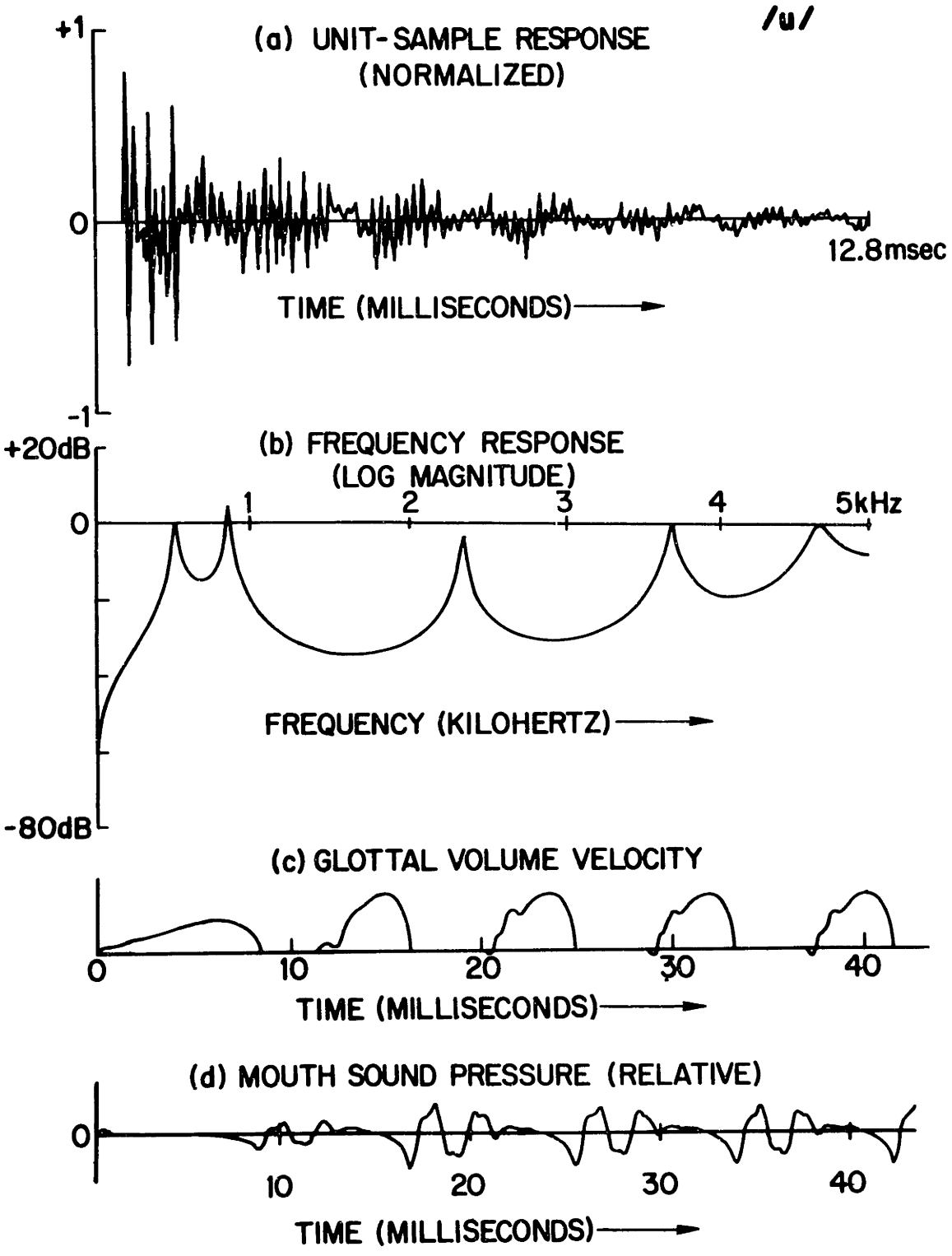

/u/
+1
(a) UNIT-SAMPLE RESPONSE
(NORMALIZED)
0
12.8msec
TIME (MILLISECONDS)
-1
+20dB
(b) FREQUENCY RESPONSE
(LOG MAGNITUDE)
1
2
3
4
5kHz
0
FREQUENCY (KILOHERTZ)
-80dB
(c) GLOTTAL VOLUME VELOCITY
0
10
20
30
40
TIME (MILLISECONDS)
(d) MOUTH SOUND PRESSURE (RELATIVE)
0
10
20
30
40
TIME (MILLISECONDS)

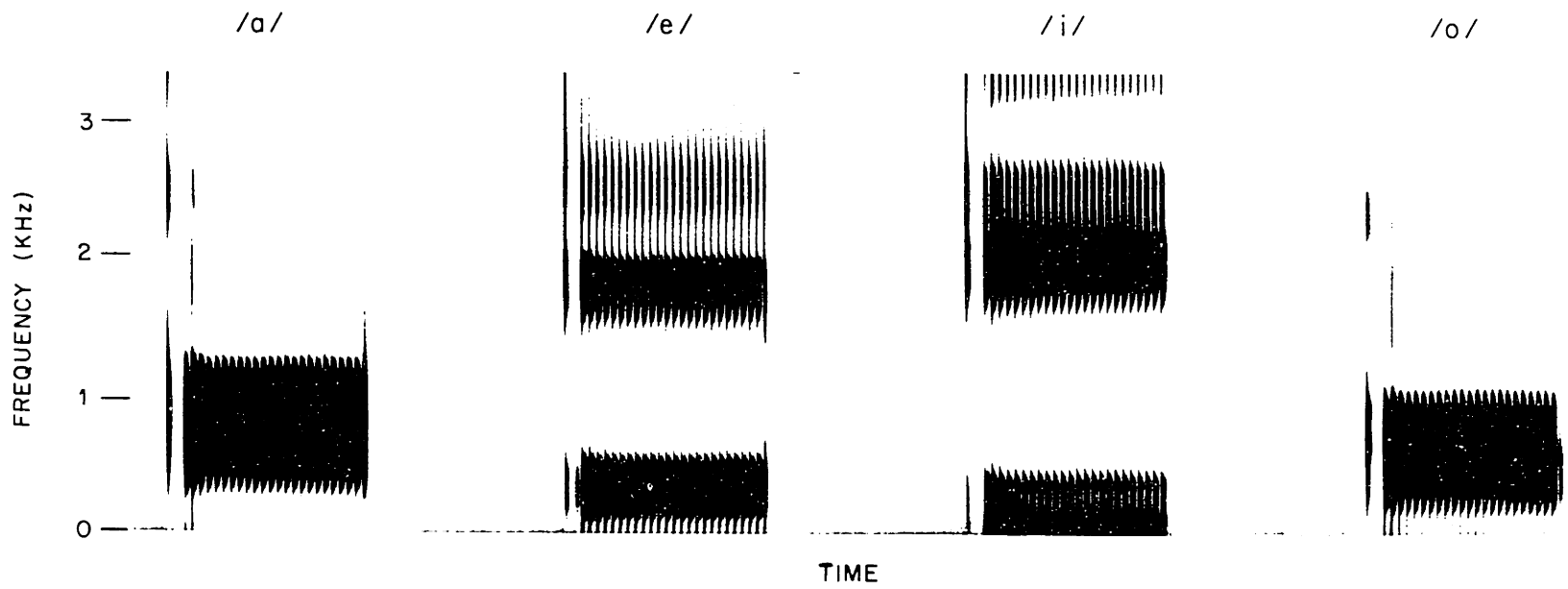
FIGURE 8.6

model is excited by a unit sample from an ideal volume-velocity source at the glottis; plot (b) shows the log magnitude of the frequency response of vocal tract obtained by multiplying the unit-sample response (a) by a 51.2 msec Hamming window and computing its DFT; plots (c) and (d) show respectively the volume velocity $U_g(t)$ measured at the glottis and the acoustic pressure $p_m(t)$ measured at the lips for the complete simulation, driven by the two-mass model. Finally, Fig. 8.7 shows wide-band spectrograms of .3 sec segments of four of the five vowels synthesized.

For each of the above simulations the sampling intervals were chosen as $\Delta t = 1/20$ msec and $\Delta x = 1.75$ cm so that the condition

$$\Delta x = c\Delta t \tag{8.1}$$

is satisfied. It was found that although the simulation was stable for any choice of sampling intervals (as predicted), the ability of the finite-difference equations to accurately simulate the partial-differential equations of Chapter II declined as the ratio $\Delta x/\Delta t$ deviated from c. Though a decrease in the accuracy of the simulation might be expected if a sampling rate is decreased, this is a somewhat unusual situation since the accuracy can also decrease if a sampling rate is increased. Such behavior was predicted in Chapter V, where it was argued that Eq. (8.1) implies a nearly characteristic net and thus should result in a nearly optimum simulation. By choosing various sampling rates, so as to violate Eq. (8.1), this contention was experimentally verified.


/a/
/e/
/i/
/o/
FREQUENCY (KHz)
3
2
1
0
TIME

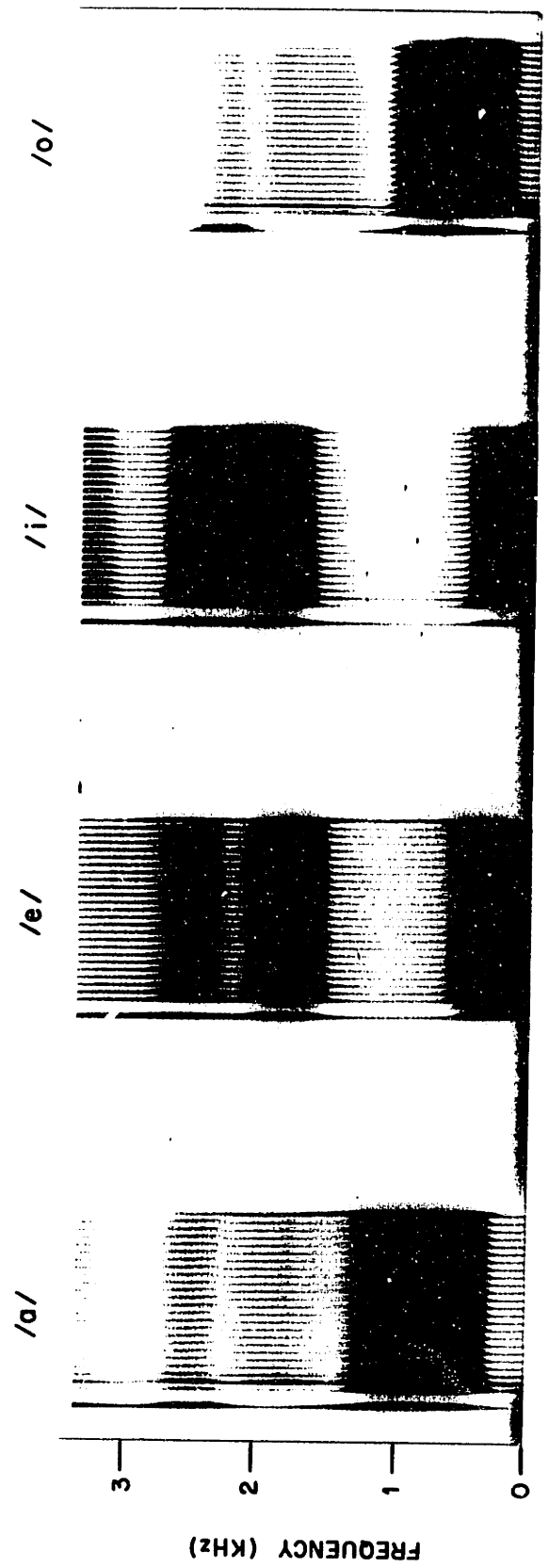


FIGURE 8.7

WIDE-BAND SPECTROGRAMS

8.2 Summary and Conclusions

It was the principal objective of this thesis to carefully formulate a digital simulation of the human vocal apparatus under a reasonable set of assumptions, valid for a class of speech-like sounds. In formulating this simulation, ideas and insights gained from the fields of digital signal processing and numerical analysis often proved valuable and provided novel approaches to some of the modeling problems. The utility of these techniques transcends the particular simulation treated here, and they deserve wider application.

The partial differential equations derived in Chapter II appear to describe adequately, sound propagation in the human vocal tract under the set of assumptions given in that chapter. In particular, the assumptions are valid for a class of speech-like sounds which include the vowels. However, they are not valid for the case of turbulent noise generated by high velocity flow through a constriction. Thus the model will not automatically generate fricative sounds without additional modifications.

Another important aspect of the simulation is that it is time varying. Although this is the correct treatment of the problem, it may be desirable, in a practical speech synthesis system, to increase computational efficiency by treating the problem as quasi static. If the dynamic equations are used in such a situation, it is expected that a rapid time change of the cross-sectional area function

$A_o(x,t)$ will generate an (unwanted) acoustic pulse. If, on the other hand, the term $\partial A_o/\partial t$ is eliminated from the dynamic equations, a discontinuity in the solution may still occur if the area changes rapidly because the solution before the change, which serves as the initial conditions for the equations after the change, no longer satisfies the new equations. Thus, if a quasi-static simulation be desired, the changes of the area function must be made with considerable care (e.g., immediately preceding a glottal pulse).

The centered-difference scheme, chosen here to formulate the finite difference simulation, is stable for any choice of sampling intervals, preserves the conservation-law character of partial-differential equations, has high (second order) accuracy, and yields an exact simulation under certain circumstances (see Appendix A). It was argued and then confirmed experimentally that an optimal simulation of the partial differential equations could be obtained when the condition $\Delta x = c\Delta t$ is satisfied. Since the simulation is stable for any choice of Δx and Δt, and since practical considerations make it desirable to choose $\Delta x \leq c\Delta t$ (the range in which explicit simulations are unstable), it might be interesting to investigate the perceptual effects of synthetic speech generated by a suboptimal simulation.

8.3 Additional Areas for Further Study

Several issues, relevant to the implementation of a speech synthesis system, have not been resolved in this thesis. As mentioned previously, the problem of turbulent

noise generation must be considered. Also, the case of a complete closure, occurring other than at the lips, where the cross-sectional area vanishes must be treated specially. Probably the most natural treatment of this case is to consider the closure as a boundary between two acoustic tubes at which the volume velocity vanishes.

The values of the parameters used here for the model of the yielding wall are at best crude estimates. Further study is necessary to determine the values of these parameters and their spatial variation more accurately.

An improvement on the Webster horn equation has been suggested by E. S. Weibel.[38] Weibel formulated a one-dimensional equation of the same form as the Webster equation, but with different coefficients. He obtained this equation for a lossfree hard-walled horn by exploiting a curvalinear coordinate system determined by the cophasic wavefronts for a given horn geometry. Although this approach was considered impractical for the pruposes of this thesis because it requires the solution of the multidimensional Laplace equation each time the horn geometry is changed, a similar approach might be investigated for a vocal-tract model, should applications arise requiring improved accuracy.

It was mentioned previously that choosing $\Delta x = c\Delta t$ is highly desirable. However, in order to obtain a "sufficiently fine" spatial sampling of the area function,

this condition often requires high temporal sampling frequencies. If it is assumed that the desired speech output has a low-pass character then it seems reasonable that sufficiently sharp changes in the area function, due to high spatial frequencies, should not be resolved by an acoustic wave with a low-pass character. Thus it may be possible to reduce the spatial sampling rate if the area function is in some way "smoothed" before it is sampled.

Finally, there is currently a growing interest in multidimensional signal processing and filter design. Once this field has been more thoroughly explored, it might be profitable to view the vocal tract as a two dimensional (one space and one time) linear time-varying filter, or as a three dimensional (one space and two time) linear time-invariant filter.

APPENDIX A

Numerical Integration and Differentiation: A Digital-Signal-Processing Approach.

One of the key results of this thesis is a demonstration of the efficacy of applying digital-signal-processing concepts and techniques to the simulation of a physical system. This approach is not only useful for formulating the simulation, but also for predicting and characterizing the error introduced by the simulation; and finally, for evaluating the simulation. With this in mind, it will be shown that a simulation of a linear time-invariant (or quasi time-invariant system obtained by replacing derivatives (or integrals) with finite-difference analogs can be characterized by a mapping of the s plane to the z plane. Moreover, this mapping is given by the transfer function of the digital network used to simulate the derivatives (or equivalently, the reciprocal of the transfer function of the digital network used to simulate the integrals). Several common numerical differentiation and integration techniques will be investigated and an extension of this approach, to a second-order system of linear partial-differential equations, will be illustrated.

Let H(s) be the transfer function of a linear time-invariant system with input u(t) and output y(t) satisfying the differential equation

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = \sum_{\ell=0}^{M} b_\ell \frac{d^\ell u(t)}{dt^\ell} \quad . \tag{A.1}$$

For simplicity, assume N > M and rewrite Eq. (A.1) as

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = \sum_{\ell=0}^{N-1} b_\ell \frac{d^\ell u(t)}{dt^\ell} \tag{A.2}$$

where $b_\ell = 0$ for any $\ell > M$. This equation can now be written in state-space form as

$$\begin{aligned} \dot{\underline{x}} &= \underline{A}\underline{x} + \underline{B}u \\ y &= \underline{C}\underline{x} \end{aligned} \tag{A.3}$$

where the components of the state vector $\underline{x}$ are given by

$$x_i(t) = \frac{d^{i-1} y(t)}{dt^{i-1}} \qquad i = 1,2,\ldots,N \tag{A.4}$$

and the matrices $\underline{A}$, $\underline{B}$, and $\underline{C}$ are given by

$$\underline{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\frac{a_0}{a_N} & -\frac{a_1}{a_N} & -\frac{a_2}{a_N} & \cdots & -\frac{a_{N-2}}{a_N} & -\frac{a_{N-1}}{a_N} \end{bmatrix}$$

$$\underline{B} = \begin{bmatrix} 0 & 0 & \cdots & 0 & \frac{1}{a_N} \end{bmatrix}^T \tag{A.5}$$

$$\underline{C} = \begin{bmatrix} b_0 & b_1 & \cdots & b_{N-1} \end{bmatrix}$$

(T signifies matrix transpose).

Using the representation of Eqs. (A.3) the transfer function H(s) may be written conveniently as

$$H(s) = Y(s)/U(s)$$

$$= \underline{C}[s\underline{I}-\underline{A}]^{-1}\underline{B} \qquad (A.6)$$

The state-space formulation is particularly convenient for our analysis for two reasons. First, in the time domain, there is only a single vector derivative operator in a conceptually simple network structure, rather than a complicated network of scalar derivative operators imbedded in an unspecified network structure. Second, in the frequency domain, the complex-frequency parameter s appears in a simple context in the expression for the transfer function H(s). By replacing each differentiator (or integrator) in H(s) by a finite-difference analog it will be shown that the transfer function $\tilde{H}(z)$ for the new (digital) system is given as

$$\tilde{H}(z) = \underline{C}[\psi(z)\underline{I}-\underline{A}]^{-1}\underline{B}$$

$$= H[\psi(z)]$$

where $\psi(z)$ is a transfer function associated with the finite-difference analog.

Consider the general finite-difference analog for the derivative whereby

$$g = \frac{df}{dt} \qquad (A.7)$$

is replaced by the difference equation*

$$\mathcal{D}_2\tilde{g} = \mathcal{D}_1\tilde{f} \tag{A.8}$$

where $\mathcal{D}_1$ and $\mathcal{D}_2$ are finite-difference operators. If $\tilde{F}(z)$ and $\tilde{G}(z)$ are the z transforms of $\tilde{f}$ and $\tilde{g}$, then define $D_1(z)$ and $D_2(z)$ so that Eq. (A.8) implies

$$D_2(z)\tilde{G}(z) = D_1(z)\tilde{F}(z). \tag{A.9}$$

From Eq. (A.9) a transfer function for the "digital differentiator" may be defined by

$$\begin{aligned} D(z) &= \tilde{G}(z)/\tilde{F}(z) \\ &= D_1(z)/D_2(z) \end{aligned} \tag{A.10}$$

Since $D_1(z)$ and $D_2(z)$ are polynomials in z and z^{-1}, $D(z)$ is a rational function of z.

If the finite-difference analog defined by Eq. (A.8) is now applied to the state Eqs. (A.3), the result is the state equations for the digital simulation:

$$\begin{aligned} \mathcal{D}_1\underline{\tilde{x}} &= \mathcal{D}_2\underline{\underline{A}}\underline{\tilde{x}} + \mathcal{D}_2\underline{B}\tilde{u} \\ \tilde{y} &= \underline{\underline{C}}\underline{\tilde{x}} \end{aligned} \tag{A.11}$$

---

* The tilde is used here to designate sequences so that $\tilde{f}(n) = f(n\Delta t)$.

Since $\underline{A}$,$\underline{B}$ and $\underline{C}$ are constants, the transfer function for the simulation is therefore

$$\tilde{H}(z) = \tilde{Y}(z)/\tilde{U}(z)$$

$$= \underline{C}[D(z)\underline{I}-\underline{A}]^{-1}\underline{B} \qquad (A.12)$$

which is recognized as H(s) by Eq. (A.6) with the transformtion

$$s = D(z) \qquad (A.13)$$

Thus, it has been shown that if a system H(s) is simulated by replacing each (analog) differentiator by a digital differentiator with transfer functions D(z), the resulting transfer function for the digital simulation is simply

$$\tilde{H}(z) = H[D(z)]$$

By a similar analysis, it can be shown that if a digital simulation is obtained by integrating Eq. (A.1), or equivalently Eq. (A.3), according to a numerical-integration formula for which

$$f(t) = \int_{t_0}^{t} g(\alpha)d\sigma + f(t_0) \qquad (A.14)$$

implies

$$J(z) = \tilde{F}(z)/\tilde{G}(z), \qquad (A.15)$$

then the transfer function of the resulting digital system is determined by the mapping

$$\frac{1}{s} = J(z) \qquad (A.16)$$

so that

$$\tilde{H}(z) = H[1/J(z)] \qquad (A.17)$$

The mapping functions for several common numerical differentiation and integration formulas are given in Table A-I. An interesting observation to be made from this table is that the four finite - difference analogs for the derivative are equivalent, respectively, to the first four numerical integration rules. Thus, for example, if it is desired to formulate a simulation by integrating a differential equation according to the trapezoid rule, it may be more convenient to implement the simulation using central differences with averaging. Furthermore, many important properties of the simulation can be predicted immediately from the properties of the particular mapping function used to generate the simulation.

Of particular interest is the region in the z plane which is the image of $Re(s) < 0$. For if this region is not wholly contained within the unit disk $|z| < 1$, there

RULE	DIFFERENCE EQUATION	MAPPING FUNCTION
FINITE - DIFFERENCE SCHEMES FOR $g(t) = df(t)/dt$		
FORWARD DIFFERENCE	$g(nT) = \frac{1}{T}\,[f(nT+T)-f(nT)]$	$s = \frac{1}{T}\,(z-1)$
BACKWARD DIFFERENCE	$g(nT) = \frac{1}{T}\,[f(nT)-f(nT-T)]$	$s = \frac{1}{T}\left(1-\frac{1}{z}\right)$
MEAN ODD CENTRAL DIFFERENCE	$g(nT) = \frac{1}{2T}\,[f(nT+T)-f(nT-T)]$	$s = \frac{1}{2T}\left(z-\frac{1}{z}\right)$
CENTRAL DIFFERENCE WITH AVERAGING	$\frac{1}{2}[g(nT)+g(nT-T)] = \frac{1}{T}\,[f(nT)-f(nT-T)]$	$s = \frac{2}{T}\left(\frac{z-1}{z+1}\right)$
FINITE - DIFFERENCE SCHEMES FOR $f(t) = f(t_o) + \int_{t_o}^{t} g(\tau)\,d\tau$		
STAIRCASE BOUNDED BY INTEGRAND	$f(nT) = f(nT-T) + Tg(nT-T)$	$s = \frac{1}{T}\,(z-1)$
STAIRCASE BOUNDING INTEGRAND	$f(nT) = f(nT-T) + Tg(nT)$	$s = \frac{1}{T}\left(1-\frac{1}{z}\right)$
MIDPOINT RULE	$f(nT) = f(nT-2T) + 2Tg(nT-T)$	$s = \frac{1}{2T}\left(z-\frac{1}{z}\right)$
TRAPEZOID RULE	$f(nT) = f(nT-T) + \frac{T}{2}\,[g(nT)+g(nT-T)]$	$s = \frac{2}{T}\left(\frac{z-1}{z+1}\right)$
SIMPSON'S PARABOLIC RULE	$f(nT) = f(nT-2T) + \frac{T}{3}\,[g(nT)+4g(nT-T)+g(nT-2T)]$	$s = \frac{3}{T}\left(\frac{z^2-1}{z^2+4z+1}\right)$

TABLE A-I

will be stable continuous-variable systems which are mapped to unstable discrete-variable systems. Moreover, it is generally desirable that the $j\omega$ axis ($\mathrm{Re}(s) = 0$) map to the unit circle $|z| = 1$ so that digital frequency corresponds to analog frequency. As an example, consider again the example given in Chapter I of an acoustic tube simulated with "T sections" and backward differences. The excessive damping can be understood immediately by observing that the mapping corresponding to backward differences maps $\mathrm{Re}(s) < 0$ onto $|z-\frac{1}{2}| < \frac{1}{2}$ and the $j\omega$ axis to $|z-\frac{1}{2}| = \frac{1}{2}$. If forward differences were used rather than backward differences, it would be found that the region $\mathrm{Re}(s) < 0$ maps onto $\mathrm{Re}(z) < 1$ and the $j\omega$ axis maps to $\mathrm{Re}(z) = 1$ so that the simulation, in general, would be unstable.

It will now be shown that a generalization of this analysis can be applied to the simulation of a system governed by a set of linear constant-coefficient partial-differential equations. Consider the case of a lossy acoustic tube or transmission line described by the set of equations

$$-\frac{\partial p}{\partial x} = L\frac{\partial U}{\partial t} + RU$$

$$-\frac{\partial U}{\partial x} = C\frac{\partial p}{\partial t} + Gp \qquad \text{(A.18)}$$

The solution to these equations is a summation (or integral) over modes of the form

$$\begin{bmatrix} p(x,t) \\ U(x,t) \end{bmatrix} = \begin{bmatrix} \hat{p}(\gamma,s) \\ \hat{U}(\gamma,s) \end{bmatrix} e^{\gamma x+st} \tag{A.19}$$

If Eq. (A.19) is substituted into Eqs. (A.18), the equation

$$\begin{bmatrix} \gamma & Ls + R \\ Cs + G & \gamma \end{bmatrix} \begin{bmatrix} \hat{p} \\ \hat{U} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{A.20}$$

results. Equation (A.20) has a nontrivial solution if and only if the determinant of the coefficient matrix vanishes, or equivalently, if and only if

$$\gamma^2 = (Ls+R)(Cs+G) \tag{A.21}$$

Equation (A.21) is the dispersion relation for the system and characterizes the nature of the wave propagation in the system. If $\gamma = \alpha + j\beta$ and $s = \sigma + j\omega$, then the phase velocity is ω/β and the group velocity is $d\omega/d\beta$.

Suppose the system governed by Eqs. (A.18) is simulated using the centered-difference scheme introduced in Chapter V. According to this scheme, Eqs. (A.18) transform to the partial-difference equations

$$\begin{aligned} -\mu_t\delta_x p &= L\mu_x\delta_t U + R\mu_x\mu_t U \\ -\mu_t\delta_x U &= C\mu_x\delta_t p + G\mu_x\mu_t p \end{aligned} \tag{A.22}$$

where μ_α is the averaging operator in the α direction and δ_α is the central-difference operator in the α direction.

The solution to these equations can be expressed as a summation (or integral) over modes of the form

$$\begin{bmatrix} p(k\Delta x, n\Delta t) \\ U(k\Delta x, n\Delta t) \end{bmatrix} = \begin{bmatrix} \hat{p}(v,z) \\ \hat{U}(v,z) \end{bmatrix} v^k z^n \qquad \text{(A.23)}$$

Substituting Eq. (A.23) into Eqs. (A.22) gives

$$-\frac{1}{2\Delta x}(1+z^{-1})(1-v^{-1})\hat{p} = \frac{L}{2\Delta t}(1+v^{-1})(1-z^{-1})\hat{U} + \frac{R}{4}(1+v^{-1})(1+z^{-1})\hat{U}$$

$$-\frac{1}{2\Delta x}(1+z^{-1})(1-v^{-1})\hat{U} = \frac{C}{2\Delta t}(1+v^{-1})(1-z^{-1})\hat{p} + \frac{G}{4}(1+v^{-1})(1+z^{-1})\hat{p} \qquad \text{(A.24)}$$

or,

$$\begin{bmatrix} \frac{2}{\Delta x}\left(\frac{v-1}{v+1}\right) & \frac{2}{\Delta t}\left(\frac{z-1}{z+1}\right)L + R \\ \frac{2}{\Delta t}\left(\frac{z-1}{z+1}\right)C + G & \frac{2}{\Delta x}\left(\frac{v-1}{v+1}\right) \end{bmatrix} \begin{bmatrix} \hat{p} \\ \hat{U} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{(A.25)}$$

Equation (A.25) has a nontrivial solution if and only if the determinant of the coefficient matrix vanishes so that

$$\left[\frac{2}{\Delta x}\left(\frac{v-1}{v+1}\right)\right]^2 = \left[\frac{2}{\Delta t}\left(\frac{z-1}{z+1}\right)L + R\right]\left[\frac{2}{\Delta t}\left(\frac{z-1}{z+1}\right)C + G\right] \qquad \text{(A.26)}$$

Equation (A.26) is the dispersion relation for the difference Eqs. (A.22) and is identical to the dispersion relation (A.21) for the partial-differential Eqs. (A.18) under the bilinear transformations

$$\gamma = \frac{2}{\Delta x}\left(\frac{v-1}{v+1}\right)$$

$$s = \frac{2}{\Delta t}\left(\frac{z-1}{z+1}\right) \qquad (A.27)$$

The continuous-variable system (A.18) is therefore related to the digital system (A.22) by the mapping (A.27).

It will now be shown that in the limit of no losses, the partial-differential Eqs. (A.18) are simulated exactly by the partial-difference Eqs. (A.22) when the sampling rates are chosen such that

$$\Delta x = c\Delta t \qquad (A.29)$$

This condition on the sampling rates is equivalent to requiring that the sample points lie on a characteristic net as described in Chapter V.

The modes of the solution to the partial-differential Eqs. (A.18), evaluated at the net points, is given by Eqs. (A.19) and (A.21) with $x = k\Delta x$ and $t = n\Delta t$ i.e.,

$$\begin{bmatrix} p(k\Delta x, n\Delta t) \\ U(k\Delta x, n\Delta t) \end{bmatrix} = \begin{bmatrix} \hat{p}(\gamma, s) \\ \hat{U}(\gamma, s) \end{bmatrix} e^{\gamma \Delta x k + s \Delta t n} \qquad (A.30)$$

where

$$\gamma^2 = (Ls+R)(Cs+G). \tag{A.31}$$

In the limit of the loss-free case, R = G = 0 and $c = 1/\sqrt{LC}$ so that Eq. (A.31) reduces to

$$\gamma = \pm s/c. \tag{A.32}$$

Equation (A.32) is a linear dispersion relation implying nondispersive plane-wave propagation with both phase and group velocity given by c.

The modes of the solution to the loss-free partial-difference equations is given by Eqs. (A.23) and (A.26) with R = G = 0 and $c = 1/\sqrt{LC}$, i.e.,

$$\begin{bmatrix} p(k\Delta x, n\Delta t) \\ U(k\Delta x, n\Delta t) \end{bmatrix} = \begin{bmatrix} \hat{p}(v,z) \\ \hat{U}(v,z) \end{bmatrix} v^k z^n \tag{A.33}$$

where

$$\left(\frac{v-1}{v+1}\right) = \pm \frac{\Delta x}{c\Delta t}\left(\frac{z-1}{z+1}\right) \tag{A.34}$$

Letting

$$v = e^{\gamma' \Delta x} \tag{A.35}$$

and

$$z = e^{s'\Delta t}$$

allows Eq. (A.34) to be written as

$$\tanh \frac{\gamma' \Delta x}{2} = \pm \frac{\Delta x}{c\Delta t} \tanh \frac{s' \Delta t}{2} \qquad (A.36)$$

Invoking the condition of the characteristic net, $\Delta x = c\Delta t$, gives

$$\gamma' = \pm s'/c \qquad (A.37)$$

which is identical to the dispersion relation (A.32) for the original partial-differential equations for $-\pi < \text{Im}(\gamma\Delta x) \leq \pi$ and $-\pi < \text{Im}(s\Delta t) \leq \pi$. Hence the simulation is exact.

Intuitively, this result can be interpreted as follows. Letting $\gamma = \alpha + j\beta$ and $s = j\omega$, the real and imaginary parts of Eq. (A.32) are

$$\begin{aligned} \alpha &= 0 \\ \beta &= \pm \omega/c \end{aligned} \qquad (A.38)$$

and the real and imaginary parts of Eq. (A.35) are

$$\tanh \frac{\alpha' \Delta x}{2} = 0$$

$$\tan \frac{\beta' \Delta x}{2} = \pm \frac{\Delta x}{c\Delta t} \tan \frac{\omega' \Delta t}{2}. \qquad (A.39)$$

Using $\alpha = \alpha' = 0$ and substituting Eqs. (A.35) into Eqs. (A.27) gives

$$\beta = \frac{2}{\Delta x} \tan \frac{\beta' \Delta x}{2}$$

$$\omega = \frac{2}{\Delta t} \tan \frac{\omega' \Delta t}{2} \qquad \text{(A.40)}$$

so that the frequency and wave number of the continuous-variable system are related to the frequency and wavenumber of the discrete-variable system by the tangential-warping functions (A.40) introduced by the bilinear transformation of the frequency spaces. If $\Delta x = c\Delta t$ then Eq. (A.39) reduces to Eq. (A.38). In other words, the choice of a characteristic net results in tangential warping of the β and ω axes such that the dispersion relation, β vs. ω. remains invariant.

The analytical approach presented here provides a useful suppliment to the classical numerical analysis approach to constructing discrete-variable simulations of continuous systems. Whereas the classical numerical-analysis approach typically characterizes a simulation according to an error term which vanishes faster than some power of the sampling interval, the approach presented here allows direct comparison, in the frequency domain, of the continuous system and its discrete-variable simulation. Such a comparison is often more useful and provides more insight into the nature and consequences of the discrete-variable approximation than a simple bound on an error term.

APPENDIX B

An In-Place Algorithm for the Solution of a Class of Sparse Linear Algebraic Equations.

When the trapezoid rule is applied to a two-point boundary-value problem for a system of two coupled first-order ordinary linear differential equations, or when the centered-difference scheme is applied to a two-point boundary-value problem for a system of two coupled first-order linear (hyperbolic) partial-differential equations, a system of simultaneous linear algebraic equations

$$\underline{H}\underline{\zeta} = \underline{b} \tag{B.1}$$

results. The coefficient matrix $\underline{H}$ is a sparse matrix of the form

$$\underline{H} = \begin{bmatrix} h_{11} & h_{12} & & & & & & & & \\ h_{21} & h_{22} & h_{23} & h_{24} & & & & & & \\ h_{31} & h_{32} & h_{33} & h_{34} & & & & & & \\ & & h_{43} & h_{44} & h_{45} & h_{46} & & & & \\ & & h_{53} & h_{54} & h_{55} & h_{56} & & & & \\ & & & & \bullet & & & & & \\ & & & & & \bullet & & & & \\ & & & & & & \bullet & & & \\ & & & & & & h_{2N-2,2N-3} & h_{2N-2,2N-2} & h_{2N-2,2N-1} & h_{2N-2,2N} \\ & & & & & & h_{2N-1,2N-3} & h_{2N-1,2N-2} & h_{2N-1,2N-1} & h_{2N-1,2N} \\ & & & & & & & & h_{2N,2N-1} & h_{2N,2N} \end{bmatrix} \tag{B.2}$$

(where N is the number of sample points on the interval including the two boundary points).

An efficient algorithm based on an adaptation of Gauss's algorithm[39,40] will now be developed for solving Eq. (B.1). It is well known, that an arbitrary square matrix $\underline{H}$ can be decomposed as the product of a lower-triangular matrix $\underline{L}$ and an upper-triangular matrix $\underline{U}$. Moreover, if $\underline{H}$ is nonsingular, then the decomposition can be made unique by specifying the diagonal elements of either $\underline{L}$ or $\underline{U}$.[39] We show, by construction, that if $\underline{H}$ is nonsingular and of the form given by Eq. (B.2) then $\underline{H}$ can be factored as the product

$$\underline{H} = \underline{L}\underline{U} \tag{B.3}$$

where $\underline{L}$ and $\underline{U}$ are triangular matrices of the form

$$\underline{L} = \begin{bmatrix} 1 & & & & & & & & \\ \ell_{21} & 1 & & & & & & & \\ \ell_{31} & \ell_{32} & 1 & & & & & & \\ & & \ell_{43} & 1 & & & & & \\ & & \ell_{53} & \ell_{54} & 1 & & & & \\ & & & & \ell_{65} & 1 & & & \\ & & & & & \ddots & & & \\ & & & & & & \ddots & & \\ & & & & & & \ell_{2N-2,2N-3} & 1 & \\ & & & & & & \ell_{2N-1,2N-3} & \ell_{2N-1,2N-2} & 1 \\ & & & & & & & & \ell_{2N,2N-1} \quad 1 \end{bmatrix} \tag{B.4}$$

and

$$\underline{U} = \begin{bmatrix} u_{11} & u_{12} & & & & & & & & \\ & u_{22} & u_{23} & u_{24} & & & & & & \\ & & u_{33} & u_{34} & & & & & & \\ & & & u_{44} & u_{45} & u_{46} & & & & \\ & & & & u_{55} & u_{56} & & & & \\ & & & & & u_{66} & u_{67} & u_{68} & & \\ & & & & & & \cdot & & & \\ & & & & & & & \cdot & & \\ & & & & & & & & \cdot & \\ & & & & & & & & u_{2N-2,2N-2} & u_{2N-2,2N-1} & u_{2N-2,2N} \\ & & & & & & & & & u_{2N-1,2N-1} & u_{2N-1,2N} \\ & & & & & & & & & & u_{2N,2N} \end{bmatrix} \tag{B.5}$$

The elements of $\underline{L}$ and $\underline{U}$ can be determined by assuming that $\underline{L}$ and $\underline{U}$ are of the form given above and considering the nonzero elements in the matrix product $\underline{LU}$ given by

$$h_{ij} = \sum_{k} \ell_{ik} u_{kj} \ . \tag{B.6}$$

Starting with

$$h_{11} = u_{11}, \quad h_{12} = u_{12} \tag{B.7}$$

and proceeding row-wise, the nonzero ℓ_{ik} and u_{kj} can be computed sequentially from the h_{ij} as follows:

$$\ell_{k,k-1} = h_{k,k-1}/u_{k-1,k-1}$$

$$u_{kk} = h_{kk} - \ell_{k,k-1}\, u_{k-1,k}$$

$$u_{k,k+1} = h_{k,k+1}$$

$$u_{k,k+2} = h_{k,k+2}$$

$$\ell_{k+1,k-1} = h_{k+1,k-1}/u_{k-1,k-1}$$

$$\ell_{k+1,k} = (h_{k+1,k} - \ell_{k+1,k-1}\, u_{k-1,k})/u_{kk}$$

$$u_{k+1,k+1} = h_{k+1,k+1} - \ell_{k+1,k}\, u_{k,k+1}$$

$$u_{k+1,k+2} = h_{k+1,k+2} - \ell_{k+1,k}\, u_{k,k+2}$$

for

$$k = 2,4,6,\ldots\ 2N-2 \qquad \text{(B.8)}$$

and

$$\ell_{2N,2N-1} = h_{2N,2N-1}/u_{2N-1,2N-1}$$

$$u_{2N,2N} = h_{2N,2N} - \ell_{2N,2N-1}\, u_{2N-1,2N} \qquad \text{(B.9)}$$

Observe that this precedure can fail only if one or more of the diagonal elements of $\underline{U}$ vanish. Since $\underline{L}$ and $\underline{U}$ are triangular matrices

$$\det \underline{L} = 1$$

$$\det \underline{U} = \prod_{i=1}^{2N} u_{ii}$$

and

$$\det \underline{H} = \det(\underline{LU}) = \prod_{i=1}^{2N} u_{ii}$$

Therefore, the procedure can fail only if $\det \underline{H} = 0$ i.e., only if the matrix $\underline{H}$ is singular.

Once the elements of $\underline{L}$ and $\underline{U}$ have been determined, the solution vector $\underline{\zeta}$ can be found in two steps. Define an auxillary vector $\underline{\eta}$ such that

$$\underline{H\zeta} = (\underline{LU})\underline{\zeta} = b$$

or

$$\underline{L\eta} = B \qquad \text{(B.10)}$$

and

$$\underline{U\zeta} = \underline{\eta} \qquad \text{(B.11)}$$

Since $\underline{L}$ is lower triangular, the solution to Eq. (B.10) can be written by inspection:

$$\eta_1 = b_1$$

$$\eta_k = b_k - \ell_{k,k-1}\,\eta_{k-1}$$

$$\eta_{k+1} = b_{k+1} - \ell_{k+1,k-1}\,\eta_{k-1} - \ell_{k+1,k}\,\eta_k$$

$$\eta_{2N} = b_{2N} - \ell_{2N,2N-1}\,\eta_{2N-1} \qquad \text{(B.12)}$$

for

$$k = 2,4,6,\ldots,2N-2.$$

Similarly, since $\underline{U}$ is upper triangular, the solution to Eq. (B.11) can be easily written as

$$\zeta_{2N} = \eta_{2N}/u_{2N,2N}$$

$$\zeta_{k+1} = (\eta_{k+1} - u_{k+1,k+2}\zeta_{k+2})/u_{k+1,k+1}$$

$$\zeta_k = (\eta_k - u_{k,k+1}\zeta_{k+1} - u_{k,k+2}\zeta_{k+2})/u_{kk}$$

$$\zeta_1 = (\eta_1 - u_{12}\zeta_2)/u_{11} \qquad \text{(B.13)}$$

for

$$k = 2N-2,\ldots,6,4,2.$$

Equations (B.7-B.9) and (B.12-B.13) provide the desired algorithm for solving Eq. (B.1). Before implementing this algorithm on a machine, several observations can be made to facilitate the programming. First, the elements of the (2N×2N) matrix $\underline{H}$ can be stored compactly in a (2N×4) array. Since the diagonal elements of the matrix $\underline{L}$ are all unity, there are only a total of (8N-4) elements in $\underline{L}$ and $\underline{U}$ which must be stored. Moreover, the computation

of these elements can be carried out in place, i.e., as each ℓ_{ij} and u_{ij} is computed, in the order specified by the algorithm, it can be stored in the location occupied by the corresponding element h_{ij} of the coefficient matrix. The ith component, η_i, of the auxillary vector may be computed during the ith iteration of the matrix factorization loop, i.e., immediately following the computation of the ith row of $\underline{U}$. If it is desired to solve Eq. (B.1) only once using a particular choice of $\underline{H}$ and $\underline{b}$, then it is not necessary to save the elements of $\underline{L}$, provided the elements of $\underline{\eta}$ are computed during the factorization loop as mentioned above. However, if Eq. (B.1) is to be solved more than once, using a particular $\underline{H}$ for several choices of $\underline{b}$, then the elements of $\underline{L}$ and $\underline{U}$ are saved, so that for each $\underline{b}$ it is only necessary to recompute the auxillary vector $\underline{\eta}$ and the solution vector $\underline{\zeta}$. Finally, for the equations of interest in this thesis, viz., Eqs. (4.20 and 7.6), some additional computational savings may be achieved by exploiting the fact that half of the nonzero elements in these coefficient matrices are plus or minus unity.

REFERENCES

1. Russell, G. O., The Vowel, Columbus: Ohio State University Press (1928).

2. Kratzenstein, C. G., "Sur la Raissance de la Formation des Voyelles," J. Phys. 22: 358-380 (1798).

3. von Kempelen, W., Le Mechanisme de la Parole, suivi de la Description d'une Machine Parlante, Vienna: Degan (1791).

4. Wheatstone, Sir Charles, The Scientific Papers of Sir Charles Wheatstone, London: Taylor and Francis (1879).

5. Helmholz, H. L. F., On the Sensation of Tone, New York: Dover Publication Inc. (1954).

6. Bell, A. G., "Prehistoric Telephone Days," National Geographic Mag. 41: 223-242 (1922).

7. Flanagan, J. L., "Source-System Interaction in the Vocal-Tract," Annals of the N.Y. Academy of Sciences 155: 9-17 (1968).

8. ———, Landgraf, L., "Self-Oscillating Source for Vocal-Tract Synthesizers," IEEE Trans. Audio and Electroacous. AU-16: 57-64 (1968).

9. ———, Cherry, L. "Excitation of Vocal-Tract Synthesizers," J. Acoust. Soc. Am. 45: 764-769 (1969).

10. Dudgeon, D. E., Multi-Mass Simulation of the Vocal Cords. S.B./S.M. Thesis, Dept. of Elect. Eng., M.I.T. (1969).

11. Ishizaka, K., Flanagan, J. L., "Synthesis of Voiced Sound from a Two-Mass Model of the Vocal Cords," Bell Syst. Tech. J. 51: 1233-1268 (1972).

12. Hildebrand, F. B., Finite-Difference Equations and Simulations, Englewood Cliffs: Prentice Hall (1968).

13. Forsythe, G. E., Wasow, W. R., Finite-Difference Methods for Partial Differential Equations, New York: John Wiley and Sons, Inc. (1960).

14. Morse, P. and Ingard, U., Theoretical Acoustics, New York: McGraw-Hill Book Co. (1968).

15. Gold, B., Rader, C. M., Digital Processing of Signals, New York: McGraw-Hill Book Co. (1969).

16. Chiba, T., Kajiyama, M., The Vowel, Its Nature and Structure, Tokyo: Tokyo-Kaiseikan Pub. Co. (1941).

17. Dunn, H. K., "The Calculation of Vowel Resonances, an Electrical Vocal Tract," J. Acoust. Soc. Am. 22: 740-753 (1950).

18. Ungeheuer, G., Elemente einer akustischen Theorie der Vokalarticulation, Berlin-Gottingen-Heidelberg: Springer-Verlag (1962).

19. Flanagan, J. L., Speech Analysis, Synthesis, and Perception, New York, Berlin: Springer-Verlag, Second Edition, (1972).

20. Fant, C. G. M., Acoustic Theory of Speech Production, The Hague: Mouton and Co. (1960).

21. Flanagan, J. L., C. H. Coker, L. R. Rabiner, R. W. Schafer and N. Umeda, "Synthetic Voices for Computers," IEEE Spectrum 7: 22-45 (1970).

22. Lesser, M. B., Lewis, J. A., "Applications of Matched Asymptotic Expansion Methods to Acoustics; Par I: The Webster Horn Equation and the Stepped Duct," J. Acoust. Soc. Am. 51: 1664-1669 (1972).

23. Morse, P. Vibration and Sound, New York: McGraw-Hill Book Co. (1948).

24. Hildebrand, F. B., Advanced Calculus for Applications, Englewood Cliffs: Prentice Hall (1962).

25. Gradshteyn, I.S., Ryzhik, I.W., Table of Integrals, Series, and Products, New York: Academic Press (1965).

26. Fox, L., The Numerical Solution of Two-Point Boundary Problems, Oxford: The Clarendon Press (1957).

27. Henrici, P., Discrete Variable Methods in Ordinary Differential Equations, New York: John Wiley and Sons, Inc. (1962).

28. Bose, A. G., Stevens, K. N., Introductory Network Theory, New York: Harper and Row, Publishers, Inc. (1965).

29. Courant, R., Friedrichs, K. O., Lewy, H., "Über die Partiellen Differenzengleichungen der Mathematischen Physik," Math. Ann. 100: 32-74 (1928).

30. Birkhoff, G., Lynch, R., "Numerical Solution of the Telegraph and Related Equations," Numerical Solution of Partial Differential Equations, J. H. Bromble, Ed., New York: Academic Press (1966).

31. Wendroff, B., "On Centered-Difference Equations for Hyperbolic Systems," J. Soc. Indust. Appl. Math. 8: 549-555 (1960).

32. Lax, P., Wendroff, B., "Systems of Conservation Laws," Comm. Pure Appl. Math. 13: 217-237 (1960).

33. ———, "Difference Schemes for Hyperbolic Equations with High Order Accuracy," Comm. Pure Appl. Math. 17: 381-398 (1964).

34. Gourlay, A. R., Morris, J. L., "A Multistep Formulation of the Optimized Lax-Wendroff Method for Nonlinear Hyperbolic Systems in Two Space Variables," Math. Comp. 22: 715-719 (1968).

35. von Rosenberg, D. V., Methods for the Numerical Solution of Partial Differential Equations, New York: American Elsevier Publishing Co. (1969).

36. Mitchell, A. R., Computational Methods in Partial Differential Equations, New York, John Wiley and Sons, Inc. (1969).

37. Hamming, R. W., Numerical Methods for Scientists and Engineers, New York: McGraw-Hill Book Co. (1962).

38. Weibel, E. S., "On Webster's Horn Equation," J. Acoust. Soc. Amer. 27: 726-727 (1955).

39. Gantmacher, F. R., Matrix Theory, Vol. I, New York: Chelsea Publishing Co. (1959).

40. Faddeeva, V. N., Computional Methods of Linear Algebra, New York: Dover Publications, Inc. (1959).