Determination of the Shape of the Human Vocal Tract from Acoustical Measurements

By B. GOPINATH and M. M. SONDHI

(Manuscript received January 14, 1970)

In this paper we describe methods for determining the cross-sectional area function of the human vocal tract from acoustical measurements made at one end. The pressure and volume velocity are assumed to obey Webster's horn equation, which is valid for frequencies below 3.5 kHz. Acoustical properties below 3.5 kHz do not uniquely specify the area function. This paper shows how high frequency information may be incorporated into the mathematical model in a manner consistent with a priori information about the vocal tract. Some results of application of the methods by computer simulation are presented. It is interesting to see from the figures that nine numbers (namely, length, four formants, and four residues) specify the area function quite well for practical purposes.

I. INTRODUCTION

In recent years there has been considerable interest in the modelling of speech production in terms of the motion of the articulators. This interest has stimulated work on the determination of the shape of the human vocal tract as a function of the utterance. For frequencies less than 3500 Hz, wave motion in the vocal tract is essentially planar, so that the shape is effectively specified by the cross-sectional area as a function of distance from one end of the tract (say from the glottis).

During the past two decades X-ray techniques have been used to determine these area functions. These techniques suffer from two major drawbacks: (*i*) In order to keep the exposure to X-rays within safe dosage limits, only a small number of measurements can be made on any one subject; (*ii*) The interpretation of X-ray data is a complex and difficult art, and a number of assumptions must be made in order to convert this data to area functions. The accuracy with which area functions are reconstructed is rather limited.

In 1965, Mermelstein and Schroeder suggested the new approach of inferring the area functions from acoustic information.[1] Under the usual assumptions of lossless plane wave propagation, they showed that if the area function $A(x)$ of a vocal tract of length l is of the form

$$\log A(x) = \log A_0 + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{l} \tag{1}$$

then in the limit as $a_n \to 0$ for all n, the nth eigenfrequency (with the tract closed at $x = 0$ and open at $x = 1$), is given by

$$\lambda_n = \lambda_{0n}(1 - \tfrac{1}{2}a_{2n-1}) \tag{2}$$

where λ_{0n} is the nth eigenfrequency of the uniform tract ($a_n \equiv 0$, $n = 1, \cdots$). Likewise, to the same approximation the nth eigenfrequency with the tract closed at both ends, is given by

$$\mu_n = \mu_{0n}(1 - \tfrac{1}{2}a_{2n}). \tag{3}$$

Using equation (2) to obtain a_{2n+1} from ω_n, Mermelstein and Schroeder obtained antisymmetric approximations to area functions from a knowledge of formant frequencies alone. This work was extended by Schroeder[2] and Mermelstein[3] to include the even-order coefficients [using measured values of the poles and zeroes of the input admittance at the lips, which correspond respectively to the λ's and μ's in equations (2) and (3)]. An extension was also made by Mermelstein[3] who devised an iterative algorithm to compute the first $2m$ coefficients in equation (1) from a knowledge of ($\lambda_1 \cdots \lambda_m$, $\mu_1 \cdots \mu_m$), when the perturbations of the eigenvalues are too large for first order perturbation theory to be accurate. Another iterative scheme has been obtained by J. Heinz, by applying perturbation theory to tracts of arbitrary shape.[4]

These methods are applications of very general techniques (namely, perturbation theory and steepest descents) which do not make use of the special characteristics of the problem at hand. They also leave unanswered certain mathematical questions such as the convergence of the iterative procedures and uniqueness of the solution.

In this paper, we describe two (noniterative) methods for computing the area function from acoustical data. Apart from clarifying the physical and mathematical aspects of the problem, these methods provide solutions in a form suitable for analyzing the sensitivity of the reconstructed area functions to inaccuracies of the data. They also enable us to answer such basic questions as: "What tube has all but a finite number of eigenvalues identical to those of a uniform tube?"

In Section II we introduce the wave equation and Webster's horn equation and list the basic properties of the solutions and eigenfunctions of the horn equation under homogeneous boundary conditions.

In Section III we present a method for computing the area function, based upon the factorization of the kernel of an integral operator which transforms solutions of the horn equation to the solutions of the equation for a uniform tract. The existence of such a transform was proved by Marchenko,[5] and the transform has been used in the solution of the inverse-Sturm Liouville problem by Gelfand and Levitan.[6]

In Section IV we present an alternative method for computing the area function based upon the solution of an integral equation whose kernel is the driving point response to an impulse at one end of the tract. This integral equation was introduced without derivation by Krein in a paper on an application of his theory of extensions of positive definite kernels.[7,8] Our derivation is physically motivated and uses only elementary theory of forced motion of a second order system.

In Section V we present preliminary results of an application of our methods to the determination of vocal tract shapes and a comparison with X-ray derived data. Figures 3 through 9 show the results of these computations.

II. MATHEMATICAL PRELIMINARIES

For a tube of variable cross-sectional area $A(x)$ the equations relating acoustical pressure p and volume velocity V are

$$\frac{\partial p}{\partial x} = -\frac{\rho}{A(x)} \frac{\partial V}{\partial t}, \tag{4a}$$

$$\frac{\partial V}{\partial x} = -\frac{A(x)}{\rho c^2} \frac{\partial p}{\partial t}, \tag{4b}$$

under the assumption of lossless plane wave propagation in the tube. These assumptions are accurate for the vocal tract for frequencies up to about 4 kHz. For convenience we will choose units such that the velocity of sound $c = 1$, the density of air $\rho = 1$ and the length of the tube is π. Then elimination of V in equations (4a) and (4b) gives

$$\frac{\partial}{\partial x} A(x) \frac{\partial p}{\partial x} = A(x) \frac{\partial^2 p}{\partial t^2} \qquad 0 \leqq x \leqq \pi; \tag{5}$$

and for sinusoidal time dependence, such that $p = \phi(x, \lambda)e^{i\lambda t}$, the function $\phi(x, \lambda)$ satisfies

$$\frac{\partial}{\partial x} A(x) \frac{\partial \phi(x, \lambda)}{\partial x} + \lambda^2 A(x)\phi(x, \lambda) = 0 \qquad 0 \leqq x \leqq \pi \tag{6}$$

which is Webster's horn equation. Throughout this paper it will be assumed that $A(x) > 0$ $(0 \leqq x \leqq \pi)$, $A(0) = 1$, and that $A(x)$ has continuous first and second derivatives except at a finite number of points in $[0, \pi]$. At the points of discontinuity $A(x)$ and its first two derivatives are assumed to have finite right and left limits. Under these conditions on $A(x)$ the following lemma holds.

Lemma 1: The solution of equation (6) *satisfying the initial conditions*

$$\phi(0, \lambda) = 1, \qquad \phi'(0, \lambda) = 0 \tag{7}$$

exists, and

$$| [A(x)]^{\frac{1}{2}}\phi(x, \lambda) - [A_0(x)]^{\frac{1}{2}}\psi(x, \lambda) | \leqq \frac{K}{\lambda} \tag{8}$$

where

$$0 \leqq K < \infty, \qquad 0 \leqq x \leqq \pi,$$

and $\psi(x, \lambda)$ *is the solution of equation* (6) *with the same initial conditions and* $A(x)$ *replaced by a canonical shape* $A_0(x)$. *The function* $A_0(x)$ *is such that* $A_0(0) = A(0) = 1$, $A_0(x)$ *is constant everywhere in* $[0, \pi]$ *except at points of discontinuity of* $A(x)$ *where it jumps by the same factor as does* $A(x)$.

The proof of this lemma is given in Appendix A.

The solutions of equation (6) satisfying the initial conditions (7) become eigenfunctions if they satisfy some homogeneous boundary condition at $x = \pi$. These eigenfunctions and eigenvalues have well known properties which for the specific case $\phi(\pi, \lambda) = 0$ we summarize in the following lemma.

Lemma 2: If $A(x)$ *satisfies the conditions described above, then there exists a sequence* λ_i *(the eigenvalues) satisfying*

(*i*) $\lambda_1 > 0$, $\lambda_i \to \infty$ as $i \to \infty$;

(*ii*) $\phi(x, \lambda_i)$ *are solutions of equation* (6) *satisfying the initial conditions* (7) *and the condition* $\phi(\pi, \lambda_i) = 0$; (8a)

(*iii*)
$$\int_0^\pi A(x)\phi(x, \lambda_i)\phi(x, \lambda_j)\, dx = 0, \qquad i \neq j,$$
$$= \alpha_i^2, \qquad i = j, \tag{9}$$

with

$$0 < \alpha_i^2 < \infty;$$

(iv) $\phi(x, \lambda_i)$ *are complete in the space* $L_2(0, \pi)$ *of square integrable functions.*

An immediate consequence of Lemmas 1 and 2 is the following corollary.

Corollary 2.1: If μ_i *is the sequence of eigenvalues for the canonical tube* $A_0(x)$, *with the conditions at* $x = 0$ *and* $x = \pi$ *as in Lemmas* 1 *and* 2, *then*

$$\lambda_i = \mu_i + 0\left(\frac{1}{i}\right) \tag{10}$$

and the α_i^2 *of Lemma* 2 *satisfy*

$$\begin{aligned} \alpha_i^2 &= \int_0^\pi A_0(x)\psi^2(x, \mu_i)\,dx + 0\left(\frac{1}{i^2}\right) \\ &= \gamma_i^2 + 0\left(\frac{1}{i^2}\right) \end{aligned} \tag{11}$$

where ψ *is as defined in Lemma* 1.

Finally we will require the following lemma.

Lemma 3: There exists a function $H(x, t)$ *such that*

$$[A_0(x)]^{\frac{1}{2}}\psi(x, \lambda) = [A(x)]^{\frac{1}{2}}\phi(x, \lambda) + \int_0^x H(x, t)[A(t)]^{\frac{1}{2}}\phi(t, \lambda)\,dt. \tag{12}$$

This can be proved by substituting equation (12) into equation (6). After trivial, but involved, algebra (see Appendix B) it turns out that for (12) to be true, $H(x, t)$ must satisfy the following:

$$\frac{\partial^2 H(x, t)}{\partial x^2} - \frac{\partial^2 H(x, t)}{\partial t^2} + \frac{\{[A(t)]^{\frac{1}{2}}\}''}{[A(t)]^{\frac{1}{2}}} H(x, t) = 0 \tag{13}$$

$$H(x, x) = -\frac{1}{2}\int_0^x \frac{\{[A(t)]^{\frac{1}{2}}\}''}{[A(t)]^{\frac{1}{2}}}\,dt - \{[A(x)]^{\frac{1}{2}}\}'\,\big|_{x=0} \tag{14}$$

$$\{[A(t)]^{\frac{1}{2}}\}'H(x, t)\,\big|_{t=0} - \left.\frac{\partial H(x, t)}{\partial t}\right|_{t=0} = 0. \tag{15}$$

The theory of partial differential equations guarantees the existence of a solution to equation (13) under the boundary conditions (14) and (15).

III. DERIVATION OF $A(x)$ FROM THE SPECTRAL FUNCTION OR TWO SETS OF EIGENVALUES

The spectral function is defined as a staircase function of λ with jumps of α_i^2 at λ_i ($i = 1, \cdots$). Thus to say that the spectral function is known is equivalent to saying that the pairs (λ_i, α_i^2) $i = 1, \cdots$ are known. Appendix C shows that if λ_i' ($i = 1, \cdots$) are the eigenvalues of the same tube for the conditions $\phi'(0, \lambda) = 0, \phi(0, \lambda) = 1, a\phi(\pi, \lambda) + b\phi'(\pi, \lambda) = 0$, ($b \neq 0$), then a knowledge of the pairs (λ_i, λ_i'), $i = 1, \cdots$ specifies the spectral function. Also, in Section IV it will turn out [see equation (27), with $x = 0$] that λ_i is the ith pole of the driving point impedance, and $1/2\alpha_i^2$ the corresponding residue.

We now derive $A(x)$, given the spectral function. In cases where $A(x)$ has continuous first and second derivatives the spectral function suffices to uniquely determine $A(x)$. In cases where $A(x)$ has a finite number of discontinuities, the locations and magnitudes of the jumps are also assumed to be known. Note that equation (12) may be written in symbolic form as

$$[A_0(x)]^{\frac{1}{2}}\psi(x, \lambda) = (I + H)\phi(x, \lambda)[A(x)]^{\frac{1}{2}} \tag{16}$$

where I is the identity on $(L^2[0, \pi])$ and H the integral operator such that

$$g = Hf \Leftrightarrow g(x) = \int_0^x H(x, t)f(t)\, dt. \tag{17}$$

Define the operator U which takes a square summable sequence of real numbers f_i, $i = 1, 2, \cdots$ to a square integrable function $f(x)$ [on $(0, \pi)$] defined as

$$f(x) = \sum_{i=1}^{\infty} f_i[A(x)]^{\frac{1}{2}}\phi(x, \lambda_i)/\alpha_i^2 . \tag{18}$$

Define the adjoint operator U^* which takes a square integrable function $f(x)$ to a square summable sequence f_i given by

$$f_i = \frac{1}{\alpha_i^2}\int_0^\pi f(x)\phi(x, \lambda_i)[A(x)]^{\frac{1}{2}}\, dx \qquad i = 1, \cdots . \tag{19}$$

Let R and R^* be defined analogously to U and U^*, with $[A(x)]^{\frac{1}{2}}\phi(x, \lambda_i)$ replaced by $\psi(x, \lambda_i)(A_0)^{\frac{1}{2}}$ in equations (18) and (19).

Then

$$R = (I + H)U$$

and

$$RR^* = (I + H)UU^*(I + H)^*. \tag{20}$$

However, from the completeness and orthogonality of the $[A(x)]^{\frac{1}{2}}\phi(x, \lambda_i)$, it follows that $UU^* = I$. Thus

$$RR^* = (I + H)(I + H)^*. \tag{21}$$

Note that

$$[(RR^* - I)f](x) = \int_0^\pi f(t) \sum_{i=1}^\infty \left(\frac{\psi(x, \lambda_i)\psi(t, \lambda_i)}{\alpha_i^2} - \frac{\psi(x, \mu_i)\psi(t, \mu_i)}{\gamma_i^2} \right) \cdot [A_0(t)A_0(x)]^{\frac{1}{2}}\, dt \tag{22}$$

(because $[A_0(x)]^{\frac{1}{2}}\psi(x, \mu_i)/\gamma_i$ is an orthonormal and complete sequence).

From the asymptotic formulae of Lemma 1, it is seen that the kernel of the operator $RR^* - I$ is of the Hilbert-Schmidt type [that is, square integrable on the square $(0 \leqq x, t \leqq \pi)$.] Therefore, if the λ_i, α_i correspond to those of a tube with appropriate boundary conditions then the factorization of equation (21) is always possible.

This essentially completes our derivation. For the kernel of RR^* can be constructed if the λ_i's and α_i's are known [and in the case of discontinuous $A(x)$, the positions and magnitudes of the jumps are also known]. The factorization (21) then gives $H(x, t)$. Finally, since $\phi(x, 0) = 1$ and $\psi(x, 0) = 1$ $(0 \leqq x \leqq \pi)$, equation (12) gives

$$[A(x)]^{\frac{1}{2}} + \int_0^x H(x, t)[A(t)]^{\frac{1}{2}}\, dt = [A_0(x)]^{\frac{1}{2}} \tag{23}$$

which can be solved for $[A(x)]^{\frac{1}{2}}$.

Although, in general, the factorization of equation (18) is difficult, we will show in Section V an effective method of computation when all but a finite number of λ_i's and α_i's are identical to the corresponding μ_i's and γ_i's.

IV. DERIVATION OF THE AREA FUNCTION FROM THE INPUT IMPEDANCE

In this section, for simplicity, the area function and its first two derivatives will be assumed continuous. Consider the forced pressure response $y(x, t)$ in the tube, due to a unit ramp $r(t)$ of volume velocity at $x = 0$. This may be obtained by including a term $\delta(x)r(t)$ on the right hand side of equation (4b). The resulting equation for $y(x, t)$ is

$$\frac{\partial}{\partial x} A(x) \frac{\partial y(x, t)}{\partial x} - A(x) \frac{\partial^2 y(x, t)}{\partial t^2} = -\delta(x)u(t) \tag{24}$$

where $u(t)$ is the unit step. Integrating equation (24) over x from 0 to a gives

$$A(a) \frac{\partial y}{\partial x}\bigg|_{x=a} - A(0) \frac{\partial y}{\partial x}\bigg|_{x=0} - \int_0^a A(x)Z(x, t)\, dx = -1 \qquad t \geqq 0. \quad (25)$$

In this equation we have put $u(t)\ \partial^2 y/\partial t^2 = Z(x, t)$; it is the transfer impedance (in the time domain) from the input end to the point x. However $(\partial y/\partial x) = 0$ at $x = 0$ because of the boundary condition. Also, since the velocity of sound has been normalized to unity, for $t \leqq a$ the region of the tube beyond $x = a$ is undisturbed. Thus for $t \leqq a$, $A(x)\ \partial y/\partial x = 0$ for $x \geqq a$. Thus equation (25) becomes

$$\int_0^a A(x)Z(x, t)\, dx = 1 \qquad 0 \leqq t \leqq a. \quad (26)$$

By expansion in terms of $\phi(x, \lambda_i)$ it can be verified that

$$Z(x, t) = \frac{\partial^2 y(x, t)}{\partial t^2} = \sum_{i=1}^{\infty} \frac{\phi(x, \lambda_i) \cos \lambda_i t}{\alpha_i^2} \qquad t \geqq 0 \quad (27)$$

where the convergence is assumed to be in the sense of distributions, and ϕ_i, α_i, λ_i are as defined in Section II.

Let $f(t)$ be a function such that

$$\int_0^a f(t)Z(x, t)\, dt = 1 \qquad x \leqq a. \quad (28)$$

Then by substitution into equation (26) it follows that

$$\int_0^a f(t)\, dt = \int_0^a A(x)\, dx. \quad (29)$$

The interesting duality in equations (26), (28), and (29) enables determination of $A(x)$ in terms of $Z(0, t)$ rather than $Z(x, t)$. Multiplying equation (28) by $A(x)Z(x, s)$, integrating over x and changing order of integration on the left-hand side we get

$$\int_0^a f(t)\, dt \int_0^a A(x)Z(x, t)Z(x, s)\, dx = \int_0^a A(x)Z(x, s)\, dx \qquad t < a. \quad (30)$$

For $s < a$, the right-hand side equals unity by virtue of equation (27). On the left-hand side the integration limits on x can be changed to $(0, \pi)$ since $Z(x, t) = 0$ for $x > t$. Then substituting for $Z(x, t)$, $Z(x, s)$ from equation (27) and using the orthogonality equation (9) we get

$$\int_0^a f(t) \sum \frac{\cos \lambda_i s \cos \lambda_i t}{\alpha_i^2}\, dt = 1, \qquad s \leqq a. \quad (31)$$

[Equation (31) may also be obtained by multiplying equation (28) by $[A(x)]^{\frac{1}{2}}$ and using the linear transformation of Section III.] Defining $\hat{f}(t) = f(|t|)$, $|t| \leqq a$ we note that $\hat{f}(t)$ satisfies

$$\int_{-a}^{a} \hat{f}(t) \sum \frac{\cos \lambda_i s \cos \lambda_i t + \sin \lambda_i s \sin \lambda_i t}{\alpha_i^2} dt = 1 \qquad |s| \leqq a \quad (32)$$

since $\sin \lambda_i t$ is odd. From an elementary trigonometric identity it then follows that

$$\int_{-a}^{a} \hat{f}(t) Z(0, |t - s|) \, dt = 1 \qquad s \leqq a \quad (33)$$

and, from equation (29)

$$\int_0^a \hat{f}(t) \, dt = \int_0^a A(x) \, dx. \quad (34)$$

Thus, if $Z(0, t)$ (which is the driving point impedance function at $x = 0$) is known, or measured, then solution of equation (33) for each a gives the area function. [Note that to get $A(x)$ for $x \leqq a$, $Z(0, t)$ is required for $t \leqq 2a$, as expected from physical considerations.]*

We close this section by noting that although we have discussed the method in terms of measurements made at $x = 0$, where the boundary condition corresponds to a closed end, trivial modifications are needed if measurements are to be made at an open end. In the latter case since the pressure vanishes, $\partial v / \partial x = 0$. Therefore, the same method is applicable to the horn equation for volume velocity, with a measurement of driving point admittance (instead of impedance). The driving point impedance (admittance) may, of course, be evaluated from measurements at any end with an arbitrary, known, termination.

V. APPLICATION OF THE METHODS TO DETERMINING VOCAL TRACT AREA FUNCTIONS

As noted in the introduction, the one dimensional Webster's horn equation is an accurate description of wave propagation in the vocal tract, only for frequencies less than about 3.5 kHz. Hence the λ_i of Section II have no physical counterpart whenever they exceed 3.5. We therefore start with the λ_i and α_i $(i = 1, \cdots n)$ as measured data, and assume that for $i > n$, the λ_i and α_i are identical with those of some canonical tube. In view of the asymptotic formulae of Section II, this assumption is reasonable.

* For another derivation of equation (33) see Appendix D where it is further shown that $f(a) = [A(a)]^{1/2}$.

We know of no *a priori* method for choosing the canonical shape so as to give the best match between the computed area functions and those of the actual vocal tract. For simplicity one might assume the canonical tube to be uniform. However the experimental area functions published by Fant[9] all show a sharp discontinuity at the epiglottis, which suggests choosing a canonical tube with such a discontinuity. We have tried both these canonical shapes.

Once a canonical shape has been chosen, α_i, λ_i, $\psi(x, \lambda_i)$ and $\psi(x, \mu_i)$, $i = 1, \cdots n$ may be computed. Under the assumptions of this section, the factorization of equation (21) can then be carried out in the following manner.

We use the vector notation

$$\mathbf{k}_1^T(x) = [\psi(x, \lambda_i)/\alpha_1, \cdots, \psi(x, \lambda_n)/\alpha_n, \quad \psi(x, \mu_i)/\gamma_1, \cdots, \psi(x, \mu_n)/\gamma_n]$$

$$\mathbf{k}_2^T(x) = [\psi(x, \lambda_i)/\alpha_1, \cdots, \psi(x, \lambda_n)/\alpha_n, \quad -\psi(x, \mu_1)/\gamma_1, \cdots, -\psi(x, \mu_n)/\gamma_n]$$

where $\mathbf{k}_1(x)$ and $\mathbf{k}_2(x)$ are n-dimensional column vectors and the superscript T denotes transposition. Then the kernel of equation (22) becomes $k_1^T(x)k_2(t)$, and it is easily seen that $H(x, t)$ has the form

$$H(x, t) = \mathbf{k}_1^T(x)\mathbf{h}(t) \qquad x > t \tag{35}$$

where $h(t)$ is some vector function of t. Then equation (18) becomes

$$\mathbf{k}_1^T(x)\mathbf{k}_2(t) = \mathbf{k}_1^T(x)\mathbf{h}(t) + \mathbf{k}_1^T(x)\left[\int_0^t \mathbf{h}(\sigma)\mathbf{h}^T(\sigma)\, d\sigma\right]\mathbf{k}_1(t) \tag{36}$$

as long as $\lambda_i \neq \mu_i$, $i = 1, \cdots n$. (If $\lambda_i = \mu_i$ for some i, a slight modification is necessary.) However, from the linear independence of the components of $\mathbf{k}_1(x)$ it follows that

$$\mathbf{k}_2(t) = \mathbf{h}(t) + \left[\int_0^t \mathbf{h}(\sigma)\mathbf{h}^T(\sigma)\, d\sigma\right]\mathbf{k}_1(t). \tag{37}$$

Equation (37) can be solved for $\mathbf{h}(t)$ by the analog circuit shown in Fig. 1, or by an equivalent computer simulation. Also, since equation (23) now becomes

$$[A(x)]^{\frac{1}{2}} + \mathbf{k}_1^T(x) \int_0^x \mathbf{h}(t)[A(t)]^{\frac{1}{2}}\, dt = [A_0(x)]^{\frac{1}{2}} \tag{38}$$

the analog circuit of Fig. 2 yields $[A(x)]^{\frac{1}{2}}$.

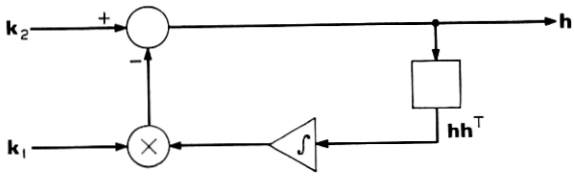


Fig. 1—Analog computer circuit for computing h.

The results of such computations for some area functions published by Fant[9] are shown in Figs. 3 through 9.

We close this section by noting that if the experimental data is in the form of a driving point impulse response, then the simplest procedure is to use the method of Section IV [that is, to solve equation (33) for various values of a]. We have not computed area functions by this method so far, but propose to do so, using impedance tube or other experimental data. The limitations due to the inapplicability of the horn equation at high frequencies apply to this method as well. The effect of low-pass filtering the driving point response is being investigated.

VI. CONCLUSIONS AND DISCUSSION

The comparison between measured and computed area functions of Figs. 3 through 9, indicates that knowledge of the first few (λ, α) pairs is sufficient to get reasonable estimates of the area function. The λ's may be obtained directly from the speech output, since they can be computed with reasonable accuracy from the formant frequencies. The α's on the other hand cannot be computed directly from the speech waveform, and impedance tube or other equivalent measurement would appear to be necessary. However, the vocal tract has physical constraints which might be reflected in a functional dependence of the α's on the λ's. The possibility of such functional dependence is being

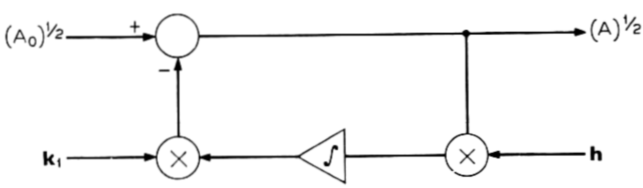


Fig. 2—Analog computer circuit for computing A.

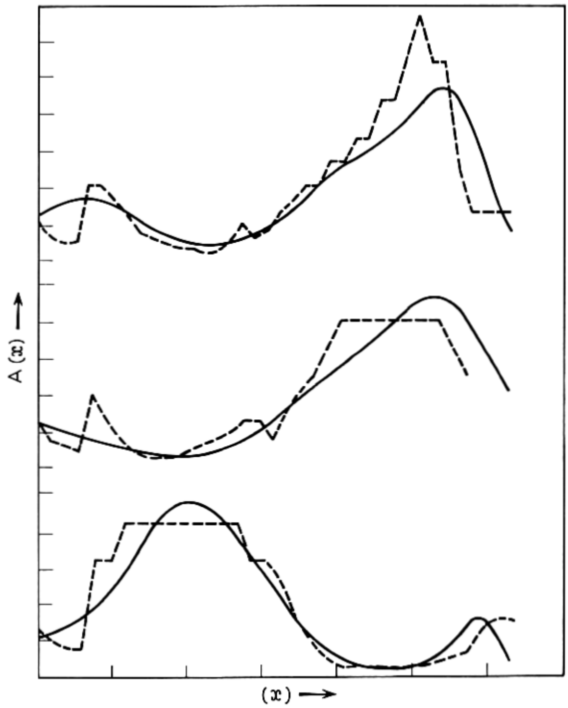


Fig. 3—Area functions reconstructed from the first *three* poles and residues of input impedance by the method of Section V using a uniform canonical tube. Dashed curves are the X-ray measurements.

investigated. The sensitivity of the computed area functions to changes in the α's is also being investigated.

If one is willing to make acoustical measurements at the lips, then the method of Section IV is the most direct way of computing the area function. It has the added advantage that the length of the vocal tract need not be known. Some preliminary results on the effect of band-limiting the impulse response have been obtained and will be reported in a later paper.

APPENDIX A

Proof of Lemma 1

Under the assumptions of this lemma, equation (6) may be transformed to:

$$\{[A(x)]^{\frac{1}{2}}\varphi(x, y)\}'' + \lambda^2[A(x)]^{\frac{1}{2}}\varphi(x, \lambda) = \{[A(x)]^{\frac{1}{2}}\}''\varphi(x, \lambda) \qquad (39)$$

except at the points $x_1, \cdots x_k$, where $A(x)$ is discontinuous. With $x_0 = 0$ and $x_{k+1} = \pi$, equation (39) gives

$$[A(x)]^{\frac{1}{2}}\varphi(x, \lambda) = f_i(x, \lambda) + \int_{x_i}^{x} \frac{\sin \lambda(x - t)}{\lambda} \frac{\{[A(t)]^{\frac{1}{2}}\}''}{[A(t)]^{\frac{1}{2}}} [A(t)]^{\frac{1}{2}}\varphi(t, \lambda)\, dt,$$

$$x_i < x \leqq x_{i+1}, \quad i = 0, 1, \cdots, k \qquad (40)$$

where

$$f_i(x, \lambda) = a_i(\lambda) \cos \lambda x + b_i(\lambda) \sin \lambda x.$$

The coefficients $a_i(\lambda)$, $b_i(\lambda)$ are to be determined so as to make $\varphi(x, \lambda)$ and $A(x)\varphi'(x, \lambda)$ everywhere continuous. (The conditions at $x = 0$ give $a_0(\lambda) = 1$, $b_0(\lambda) = 0$.) Clearly for every λ there exists a bound $m_i(\lambda) = \sup |[A(x)]^{\frac{1}{2}}\varphi(x, \lambda)|$, $x_i \leqq x \leqq x_{i+1}$. Then from equation (40),

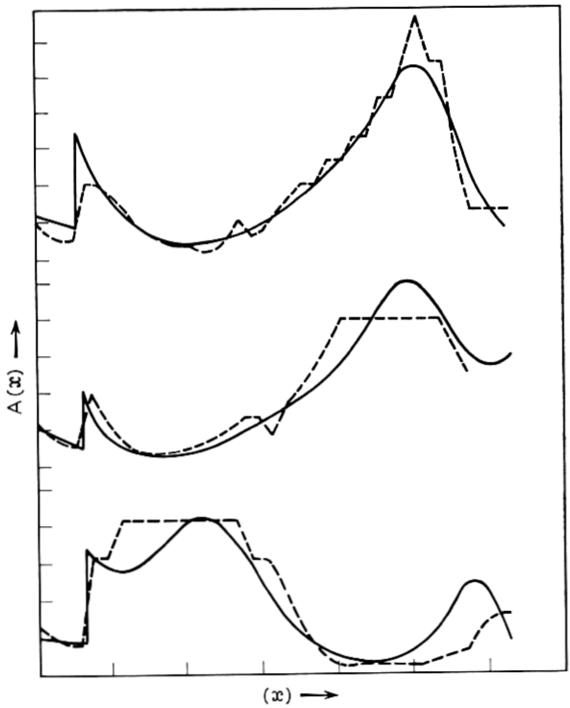
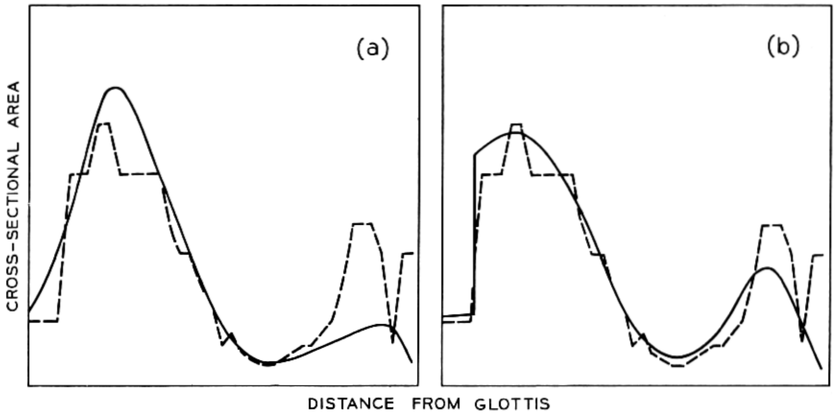


Fig. 4—Same as in Fig. 3, except the canonical tube was chosen with a discontinuity at the epiglottis.

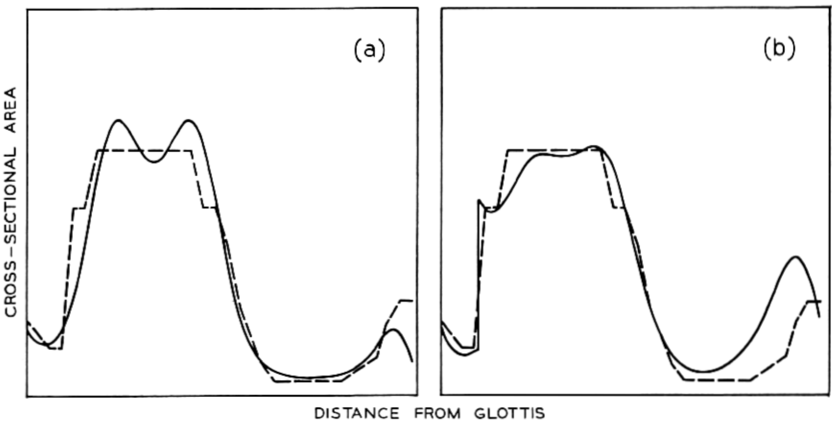
$$
\begin{aligned}
m_i(\lambda) &\leqq | f_i(x, \lambda) | + \frac{1}{\lambda} m_i(\lambda) \int_{x_i}^{x} | \sin \lambda(x - t) | \left| \frac{\{[A(t)]^{\frac{1}{2}}\}''}{[A(t)]^{\frac{1}{2}}} \right| dt \\
&\leqq [a_i^2(\lambda) + b_i^2(\lambda)]^{\frac{1}{2}} + M m_i(\lambda)/\lambda
\end{aligned}
\tag{41}
$$

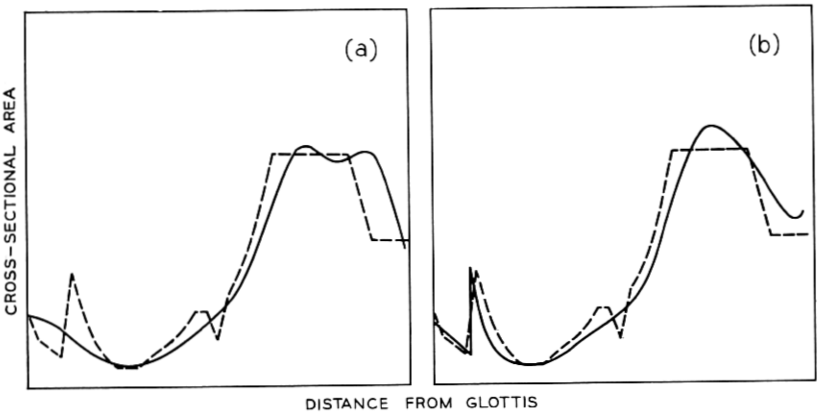
where M is a bound on the integral for all i. Thus for $\lambda > 2M$,

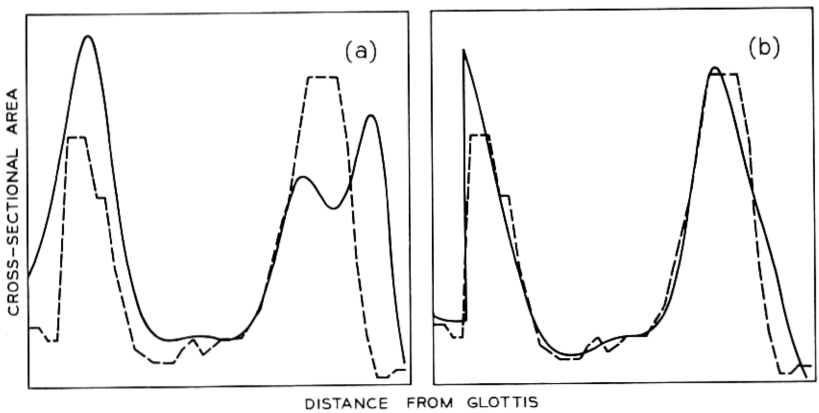
$$
\begin{aligned}
| [A(x)]^{\frac{1}{2}} \varphi(x, \lambda) | &\leqq 2[a_i^2(\lambda) + b_i^2(\lambda)]^{\frac{1}{2}} \\
&\triangleq 2\gamma_i(\lambda).
\end{aligned}
\tag{42}
$$

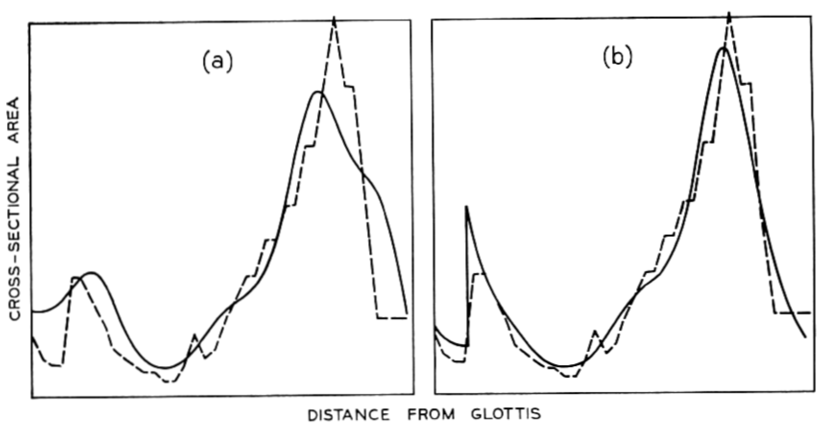


Figs. 5–9—Area functions reconstructed from the first *four* poles and residues: (a) the reconstruction using a uniform canonical tube, and (b) the reconstruction with a discontinuous canonical tube as in Section V. Dashed curves are X-ray measurements.




(a)
(b)
CROSS-SECTIONAL AREA
DISTANCE FROM GLOTTIS


(a)
(b)
CROSS-SECTIONAL AREA
DISTANCE FROM GLOTTIS


(a)
(b)
CROSS-SECTIONAL AREA
DISTANCE FROM GLOTTIS

Then from equation (40),

$$[A(x)]^{\frac{1}{2}}\varphi(x, \lambda) = f_i(x, \lambda) + \frac{\gamma_i(\lambda)c_i(x)}{\lambda}, \qquad x_i < x < x_{i+1} \tag{43}$$

where $c_i(x)$ is bounded. Differentiating equation (40) and using a similar argument gives,

$$\frac{[A(x)]^{\frac{1}{2}}}{\lambda}\frac{\partial\varphi(x, \lambda)}{\partial x} = \frac{1}{\lambda}\frac{\partial f_i(x, \lambda)}{\partial x} - \gamma_i(\lambda)\, d_i(x)/\lambda, \qquad x_i < x < x_{i+1} \tag{44}$$

with $d_i(x)$ bounded. Defining

$$\mathbf{k}_i \equiv \begin{bmatrix} [A(x_i+)/A(x_i-)]^{\frac{1}{2}} & 0 \\ 0 & [A(x_i-)/A(x_i+)]^{\frac{1}{2}} \end{bmatrix}$$

and

$$\mathbf{R}_i \equiv \begin{bmatrix} \cos \lambda x_i & \sin \lambda x_i \\ \sin \lambda x_i & -\cos \lambda x_i \end{bmatrix}$$

the continuity conditions at x_i give:

$$\begin{bmatrix} a_i(\lambda) \\ b_i(\lambda) \end{bmatrix} = \mathbf{R}_i\mathbf{k}_i\mathbf{R}_i\begin{bmatrix} a_{i-1}(\lambda) \\ b_{i-1}(\lambda) \end{bmatrix} + \frac{\gamma_{i-1}(\lambda)}{\lambda}\cdot \mathbf{R}_i\mathbf{k}_i\begin{bmatrix} c_{i-1}(x_i) \\ d_{i-1}(x_i) \end{bmatrix} - \frac{\gamma_i(\lambda)}{\lambda}\mathbf{R}_i\begin{bmatrix} c_i(x_i) \\ d_i(x_i) \end{bmatrix}. \tag{45}$$

Since the norm of $\mathbf{R}_i$ is unity and of $\mathbf{k}_i$ finite it follows upon taking the lengths of the vectors on either side of equation (45) that for large enough λ, $\gamma_i(\lambda) \leqq K'\gamma_{i-1}(\lambda)$, for some finite constant K'. Since $\gamma_0(\lambda) = 1$, it follows that $\gamma_i(\lambda)$ is bounded for all i, as $\lambda \to \infty$. Then from equation (45)

$$\begin{bmatrix} a_i \\ b_i \end{bmatrix} = \mathbf{R}_i\mathbf{k}_i\mathbf{R}_i\begin{bmatrix} a_{i-1} \\ b_{i-1} \end{bmatrix} + 0\left(\frac{1}{\lambda}\right). \tag{46}$$

However, if a_i, b_i satisfy equation (46), then $f_i(x, \lambda) = [A_0(x)]^{\frac{1}{2}}\psi(x, \lambda) + 0(1/\lambda)$ for $x_i \leqq x \leqq x_{i+1}$, therefore from equation (43)

$$[A(x)]^{\frac{1}{2}}\varphi(x) = [A_0(x)]^{\frac{1}{2}}\psi(x) + \frac{c(x)}{\lambda} \tag{47}$$

with $c(x)$ bounded.

APPENDIX B

Existence of the Linear Transformation of Equation (12)

In this appendix we prove the existence of the linear transformation of equation (12). Consider first the region $0 \leqq x \leqq x_1$, where x_1 is the first point of discontinuity of $A(x)$. Then with $y(x) = [A(x)]^{\frac{1}{2}}\phi(x, \lambda)$ and $q(x) = \{[A(x)]^{\frac{1}{2}}\}''/[A(x)]^{\frac{1}{2}}$, equation (6) becomes

$$y''(x) = -\lambda^2 y(x) + q(x)y(x). \tag{48}$$

Consider the function

$$\mathfrak{N}(x) = \int_0^x H(x, t)y(t)\, dt. \tag{49}$$

Then

$$\mathfrak{N}''(x) = \frac{d}{dx}[H(x, x)y(x)] + \frac{\partial H(x, t)}{\partial x} y(t) \Big|_{t=x} + \int_0^x \frac{\partial^2 H(x, t)}{\partial x^2} y(t)\, dt. \tag{50}$$

After integrating twice by parts we arrive at the identity

$$\int_0^x \frac{\partial^2 H(x, t)}{\partial t^2} y(t)\, dt = \frac{\partial H(x, t)}{\partial t} y(t) \Big|_{t=0}^{t=x} - H(x, t)\frac{dy}{dt}\Big|_{t=0}^{t=x} + \int_0^x H(x, t)\frac{d^2 y}{dt^2}\, dt. \tag{51}$$

Substituting for d^2y/dt^2 from equation (48) into equation (51) and adding equation (50) we get

$$\mathfrak{N}''(x) + \lambda^2\mathfrak{N} = \int_0^x \left[\frac{\partial^2 H(x, t)}{\partial x^2} - \frac{\partial^2 H(x, t)}{\partial t^2} + q(t)H(x, t)\right] \cdot y(t)\, dt + 2y(x)\frac{d}{dx}H(x, x) - y(t)\frac{\partial H(x, t)}{\partial t}\Big|_{t=0} + H(x, t)\frac{dy(t)}{dt}\Big|_{t=0}. \tag{52}$$

If now $H(x, t)$ satisfies the differential equation (13) with boundary conditions (14) and (15), then equation (52) shows that $\mathfrak{N}(x) + y(x)$ is some linear combination of $\cos \lambda x$ and $\sin \lambda x$. Matching of boundary conditions at $x = 0$ then shows that

$$\mathfrak{N}(x) + y(x) = \cos \lambda x = [A_0(x)]^{\frac{1}{2}}\psi(x, \lambda), \qquad 0 \leqq x \leqq x_1. \tag{53}$$

The proof may be extended to $x > x_1$ by a similar procedure. Thus, for example, in the range $x_1 \leqq x \leqq x_2$, $H(x, t)$ must satisfy the differential equation (13), the boundary conditions (14) and (15) for $x > x_1$, and boundary conditions at $x = x_1$ imposed by the continuity requirements on $\phi(x, \lambda)$, $\psi(x, \lambda)$, $A(x)\phi'(x, \lambda)$ and $A_0(x)\psi'(x, \lambda)$.

APPENDIX C

Spectral Function from Two Sets of Eigenvalues

We outline a method of getting a spectral function from two sets of eigenvalues. Let $\varphi(x, \lambda)$ be the solution of equation (6) such that

$$\varphi(\pi, \lambda) = \alpha, \qquad A(\pi)\varphi'(\pi, \lambda) = \beta \tag{54}$$

for every λ. Let $\psi(x, \lambda)$ be the solution such that

$$\psi(\pi, \lambda) = \gamma, \qquad A(\pi)\psi'(\pi, \lambda) = \delta. \tag{55}$$

Let $\lambda_1^2, \lambda_2^2, \cdots$ be the values of λ^2 for which $a\varphi(0, \lambda) + bA(0)\varphi'(0, \lambda) = 0$ and let $\mu_1^2, \mu_2^2, \cdots$ be the values of μ^2 for which $a\psi(0, \mu) + bA(0)\psi'(0, \mu) = 0$. Let

$$m(\lambda) = \frac{a\psi(0, \lambda) + bA(0)\psi'(0, \lambda)}{a\varphi(0, \lambda) + bA(0)\varphi'(0, \lambda)}. \tag{56}$$

Then the zeroes of $m(\lambda)$ are $\mu_1, \mu_2, \cdots$ and the poles are $\lambda_1, \lambda_2, \cdots$. If $X(x, \lambda)$ is any solution of equation (6), then it is easily shown that

$$(\lambda^2 - \lambda_n^2) \int_0^\pi A(x)X(x, \lambda)\varphi(x, \lambda_n)\, dx$$
$$= A(x)[X(x, \lambda)\varphi'(x, \lambda_n) - \varphi(x, \lambda_n)X'(x, \lambda)]_0^\pi . \tag{57}$$

Choosing $X(x, \lambda) = \psi(x, \lambda) - m(\lambda)\varphi(x, \lambda)$ in equation (57) and using the boundary conditions on $\psi(x, \lambda)$ and $\varphi(x, \lambda)$, we get

$$(\lambda^2 - \lambda_n^2) \int_0^\pi A(x)[\psi(x, \lambda) - m(\lambda)\varphi(x, \lambda)]\, dx = \beta\gamma - \alpha\delta \tag{58}$$

for all λ. As $\lambda \to \lambda_n$

$$\alpha_n^2 = \int_0^\pi A(x)\varphi^2(x, \lambda_n)\, dx = \lim_{\lambda \to \lambda_n} (\alpha\delta - \beta\gamma)/[(\lambda^2 - \lambda_n^2)m(\lambda)]. \tag{59}$$

Thus, given λ_1, λ_2, and μ_1, μ_2, one obtains $m(\lambda)$ and hence $\alpha_1, \alpha_2, \cdots$.

APPENDIX D

Derivation of Integral Equation (33)

We give here a derivation of the integral equation (33) based upon the results of Section III. For simplicity we will assume that $A(x)$ [and hence $A_0(x)$] has no discontinuities. Then from equation (23)

$$(I + H)A^{\frac{1}{2}}(x) = u(x) \tag{60}$$

where $u(x)$ is equal to 1 for all $x > 0$. Thus if $f(x)$ is a function such that

$$[(I + H)^{-1}u](x) = g(x) \tag{61}$$

then

$$\int_0^{\pi} g^2(x)\, dx = \int_0^{\pi} A(x)\, dx. \tag{62}$$

Notice that if

$$(I + H)(I + H)^* = I + K \tag{63}$$

then

$$\begin{aligned}\int_0^{\pi} A(x)\, dx &= \int_0^{\pi} [(I + K)^{-1}u](x)\, dx \\ &= \int_0^{\pi} f(x)\, dx\end{aligned} \tag{64}$$

where $f(x)$ satisfies the equation

$$[(I + K)f](x) = u(x). \tag{65}$$

The kernel of $I + K$ is recognized as that of equation (31) with $a = \pi$. Equation (33) therefore follows (for $a = \pi$) from the symmetrization of $f(x)$, exactly as in Section IV. However, the argument given here is independent of the length π, which may be replaced by a.

Using equations (60), (63) and (65), we have

$$(I + H)^*f(\cdot) = [A(\cdot)]^{\frac{1}{2}}. \tag{66}$$

Therefore from equation (17)

$$f(a) = [A(a)]^{\frac{1}{2}}. \tag{67}$$

REFERENCES

1. Mermelstein, P., and Schroeder, M. R., "Determination of Smoothed Cross-Sectional Area Functions of the Vocal Tract from Formant Frequencies," paper A-24, Proceedings of the Fifth International Congress on Acoustics,

1965, Liege, Belgium, D. E. Commins, editor (Imprimerie Georges Thone, Liege, 1965), Vol. 1a.
2. Schroeder, M. R., "Determination of the Geometry of the Human Vocal Tract," J. Acoust. Soc. of Amer., *41*, No. 4 (April 1967), pp. 1002–1010.
3. Mermelstein, P., "Determination of the Vocal-Tract Shape from Measured Formant Frequencies," J. Acoust. Soc. of Amer., *41*, No. 5 (May 1967), pp. 1283–1294.
4. Heinz, J., "Perturbation Functions for the Determination of Vocal Tract Area Functions from Vocal Tract Eigenvalues," Quarterly Progress and Status Report, April 15, 1967, Speech Transmission Laboratory, Royal Inst. of Tech., Stockholm, Sweden, pp. 1–14.
5. Marchenko, V. A., "Some Questions in the Theory of Second Order Differential Operators," Doklady Akademii Nauk, SSSR (N.S.), *72*, No. 3 (May 21, 1950), pp. 457–460.
6. Gelfand, I. M., and Levitan, B. M., "On the Determination of a Differential Equation from its Spectral Function," Izvestia Akademii Nauk, SSSR (Seria Matematicheskaya), *15*, No. 4 (July–August 1951), pp. 309–360. English Translation: Amer. Math. Soc. Translations, Ser. 2, *1* (1955), pp. 253–304.
7. Krein, M. G., "Solution of the Inverse Sturm-Liouville Problem," Doklady Akademii Nauk SSSR (N.S.), *76*, No. 1 (January 1, 1951), pp. 21–24.
8. Krein, M. G., "Determination of the Density of a Nonuniform Symmetric String from its Frequency Spectrum," Doklady Akademii Nauk SSSR (N.S.), *76*, No. 3 (January 21, 1951), pp. 345–348.
9. Fant, G. M., "Acoustic Theory of Speech Production," The Hague, Netherlands: Mouton and Co., p. 115.