INVERSE SCATTERING TO DETERMINE THE SHAPE OF A VOCAL TRACT

By

Tuncay Aktosun

IMA Preprint Series # 2032

(March 2005)

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

514 Vincent Hall
206 Church Street S.E.
Minneapolis, Minnesota 55455–0436

Phone: 612/624-6066 Fax: 612/626-7370
URL: http://www.ima.umn.edu

INVERSE SCATTERING TO DETERMINE THE SHAPE OF A VOCAL TRACT

Tuncay Aktosun
Department of Mathematics and Statistics
Mississippi State University
Mississippi State, MS 39762, USA

Abstract: The inverse scattering problem is reviewed for determining the cross sectional area of a human vocal tract. Various data sets are examined resulting from a unit-amplitude, monochromatic, sinusoidal volume velocity sent from the glottis towards the lips. In case of nonuniqueness from a given data set, additional information is indicated for the unique recovery.

PACS (2003): 02.30.Zz, 43.72.Ct
Mathematics Subject Classification (2000): 34A55, 35R30, 76Q05
Keywords: Inverse scattering, Speech acoustics, Shape of vocal tract
Short title: Inverse scattering in a vocal tract

1. INTRODUCTION

A fundamental inverse problem related to human speech is [6,7,18,19,22] to determine the cross sectional area of the human vocal tract from some data. The vocal tract can be visualized as a tube of 14-20 cm in length, with a pair of lips known as vocal cords at the glottal end and with another pair of lips at the mouth. In this review paper, we consider various types of frequency-domain scattering data resulting from a unit-amplitude, monochromatic, sinusoidal volume velocity inputted at the glottis (the opening between the vocal cords), and we examine whether each data set uniquely determines the vocal-tract area, or else, what additional information may be needed for the unique recovery.

Human speech consists of phonemes; for example, the word "book" consists of the three phonemes /b/, /u/, and /k/. The number of phonemes may vary from one language to another, and in fact the exact number itself of phonemes in a language is usually a subject of debate. In some sense, this is the analog of the number of colors in a rainbow. In American language the number of phonemes is 36, 39, 42, 45, or more or less, depending on the analyst and many other factors. The phonemes can be sorted into two main groups as vowels and consonants. The vowels can further be classified into monophthongs such as /e/ in "pet" and diphtongs such as the middle sound in "boat." The consonants can further be classified into approximants (also known as semivowels) such as /y/ in "yes," fricatives such as /sh/ in "ship," nasals such as /ng/ in "sing," plosives such as /p/ in "put," and affricates such as /ch/ "church."

The production of human speech occurs usually by inhaling air into the lungs and then sending it back through the vocal tract and out of the mouth. The air flow is partially controlled by the vocal cords, by the muscles surrounding the vocal tract, and by various articulators such as the tongue and jaw. As the air is pushed out of the mouth the pressure wave representing the sound is created. The effect of articulators in vowel production is less visible than in consonant production, and one can use compensatory articulation, especially in vowel production, by ignoring the articulators and by producing a phoneme solely by controlling the shape of the vocal tract with the help of surrounding muscles. Such a technique is often used by ventriloquists.

From a mathematical point of view, we can assume [18,19] that phonemes are basic units of speech, each phoneme lasts about 10-20 msec, the shape of the vocal tract does not change in time during the production of each phoneme, the vocal tract is a right cylinder whose cross sectional area A varies along the distance x from the glottis. We let $x = 0$ correspond to the glottis and $x = l$ to the lips. We can also assume that A is positive on $(0, l)$ and that both A and its derivative A' are continuous on $(0, l)$ and have finite limits at $x = 0^+$ and $x = l^-$. In fact, we will simply write $A(0)$ for the glottal area and $A(l)$ for the area of the opening at the lips because we will not use A or A' when $x \notin [0, l]$.

Besides $A(x)$, the primary quantities used in the vocal-tract acoustics are the time t, the pressure $p(x, t)$ representing the force per unit cross sectional area exerted by the moving air molecules, the volume velocity $v(x, t)$ representing the product of the cross sectional area and the average velocity of the air molecules crossing that area, the air density μ (about 1.2×10^{-3} gm/cm^3 at room temperature), and the speed of sound c (about $c = 3.43 \times 10^4$ cm/sec in air at room temperature). In our analysis, we assume that the values of μ and c are already known and we start the time at $t = 0$. It is reasonable [18,19] to assume that the propagation is lossless and planar and that the acoustics in the vocal tract is governed [6,7,18,19,22] by

$$\begin{cases} A(x)\, p_x(x, t) + \mu\, v_t(x, t) = 0, \\ A(x)\, p_t(x, t) + c^2 \mu\, v_x(x, t) = 0, \end{cases} \tag{1.1}$$

where the subscripts x and t denote the respective partial derivatives.

In order to recover A, we will consider various types of data for $k \in \mathbf{R}^+$ resulting from the glottal volume velocity

$$v(0, t) = e^{ickt}, \qquad t > 0. \tag{1.2}$$

Note that the quantity $\dfrac{ck}{2\pi}$ is the frequency measured in Hertz. Informally, we will refer to k as the frequency even though k is actually the angular wavenumber.

The recovery of A can be analyzed either as an inverse spectral problem or as an inverse scattering problem. When formulated as an inverse spectral problem, the boundary conditions are imposed both at the glottis and at the lips. The boundary conditions at both ends of the vocal tract result in standing waves whose frequencies form an infinite sequence. It is known [9,11,14,15,17-19] that A can be recovered from a data set consisting of two infinite sequences. Such sequences can be chosen as the zeros and poles [15,17] of the input impedance or the poles and residues [11] of the input impedance.

If the recovery of A is formulated as an inverse scattering problem, a boundary condition is imposed at only one end of the vocal tract—either at the glottis or at the lips. The data can be acquired either at the same end or at the opposite end. We have a reflection problem if the data acquisition and the imposed boundary condition occur at the same end of the vocal tract. If these occur at the opposite ends, we have a transmission problem. The inverse scattering problem may be solved either in the time domain or in the frequency domain, where the data set is a function of t in the former case and of k in the latter. The reader is referred to [4,18-21,23] for some approaches as time-domain reflection problems, to [16] for an approach as a time-domain transmission problem, and to [1] for an approach as a frequency-domain transmission problem.

Our paper is organized as follows. In Section 2 we relate (1.1) to the Schrödinger equation (2.3), introduce the selfadjoint boundary condition (2.7) involving $\cot\alpha$ given in (2.8), and present the Jost solution f, the Jost function F_α, and the scattering coefficients T, L, and R. In Section 3 we introduce the relative radius η of the vocal tract, relate it to the regular solution φ_α to the half-line Schrödinger equation, and also express η in terms of the Jost solution and the scattering coefficients. In Section 4 we present the expressions for the pressure and the volume velocity in the vocal tract in terms of the area, the Jost solution, and the Jost function; in that section we also introduce various data sets that will be used in Section 6. In Section 5 we review the recovery of the potential in the Schrödinger equation and the boundary parameter $\cot\alpha$ from the amplitude of the Jost function; for this purpose we outline the Gel'fand-Levitan method [3,12,13,10] and also the method of [13]. In Sections 6 we examine the recovery of the potential, relative radius of the tract, and the vocal-tract area from various data sets introduced in Section 4. Such data sets include the absolute values of the impedance at the lips and at the glottis, the absolute value of the pressure at the lips, the absolute value of a Green's function at the lips associated with (2.6), the reflectance at the glottis, and the absolute value of the transfer function from the glottis to the lips. Finally, in Section 7 we show that two data sets containing the absolute value of the same transfer function but with different logarithmic derivatives of the area at the lips correspond to two distinct potentials as well as distinct radii and areas for the vocal tract.

2. SCHRÖDINGER EQUATION AND JOST FUNCTION

In this section we relate the acoustic system in (1.1) to the Schrödinger equation, present the selfadjoint boundary condition (2.7) identified by the logarithmic derivative of the area function at the glottis, and introduce the Jost solution, Jost function, and the scattering coefficients associated with the Schrödinger equation.

Letting

$$P(k,x) := p(x,t)\, e^{-ickt}, \quad V(k,x) := v(x,t)\, e^{-ickt}, \tag{2.1}$$

we can write (1.1) as

$$\begin{cases} A(x)\, P'(k,x) + ic\mu k\, V(k,x) = 0, \\ c^2\mu\, V'(k,x) + ick\, A(x)\, P(k,x) = 0, \end{cases} \tag{2.2}$$

where the prime denotes the-x derivative. Eliminating V in (2.2), we get

$$[A(x)\, P'(k,x)]' + k^2 A(x)\, P(k,x) = 0, \qquad x \in (0,l),$$

or equivalently

$$\psi''(k,x) + k^2\psi(k,x) = Q(x)\,\psi(k,x), \tag{2.3}$$

with

$$\psi(k,x) =: \sqrt{A(x)}\, P(k,x), \quad Q(x) := \frac{[\sqrt{A(x)}]''}{\sqrt{A(x)}}. \tag{2.4}$$

Alternatively, letting

$$\Phi(x,t) := \sqrt{A(x)}\, p(x,t), \tag{2.5}$$

we find that Φ satisfies the plasma-wave equation

$$\Phi_{xx}(x,t) - \frac{1}{c^2}\,\Phi_{tt}(x,t) = Q(x)\,\Phi(x,t), \qquad x \in (0,l), \quad t>0. \tag{2.6}$$

We can analyze the Schrödinger equation in (2.3) on the full line $\mathbf{R}$ by using the extension $Q \equiv 0$ for $x<0$ and $x>l$. We can also analyze it on the half line $\mathbf{R}^+$ by using the extension $Q \equiv 0$ for $x>l$ and by imposing the selfadjoint boundary condition

$$\sin\alpha \cdot \varphi'(k,0) + \cos\alpha \cdot \varphi(k,0) = 0, \tag{2.7}$$

where

$$\cot\alpha := -\frac{A'(0)}{2\,A(0)} = -\frac{[\sqrt{A(x)}]'\big|_{x=0}}{\sqrt{A(0)}}. \tag{2.8}$$

As solutions to the half-line Schrödinger equation, we can consider the regular solution φ_α satisfying the initial conditions

$$\varphi_\alpha(k,0)=1, \quad \varphi'_\alpha(k,0) = -\cot\alpha, \tag{2.9}$$

and the Jost solution f satisfying the asymptotic conditions

$$f(k,x) = e^{ikx}[1+o(1)], \quad f'(k,x) = ik\,e^{ikx}[1+o(1)], \qquad x \to +\infty.$$

Since $Q \equiv 0$ for $x>0$, we have

$$f(k,l) = e^{ikl}, \quad f'(k,l) = ik\,e^{ikl}. \tag{2.10}$$

The Jost function F_α associated with the boundary condition (2.7) is defined [3,10,12,13] as

$$F_\alpha(k) := -i[f'(k,0) + \cot\alpha \cdot f(k,0)], \tag{2.11}$$

and it satisfies [3,10,12,13]

$$F_\alpha(k) = k + O(1), \qquad k \to \infty \text{ in } \overline{\mathbf{C}^+}, \tag{2.12}$$

$$F_\alpha(-k) = -F_\alpha(k)^*, \qquad k \in \mathbf{R}, \tag{2.13}$$

where $\mathbf{C}^+$ is the upper half complex plane, $\overline{\mathbf{C}^+} := \mathbf{C}^+ \cup \mathbf{R}$, and the asterisk denotes complex conjugation.

Associated with the full-line Schrödinger equation, we have the transmission coefficient T, the left reflection coefficient L, and the right reflection coefficient R that can be obtained from the Jost solution f via

$$f(k,0) = \frac{1+L(k)}{T(k)}, \quad f'(k,0) = ik\,\frac{1-L(k)}{T(k)}, \quad R(k) = -\frac{L(-k)\,T(k)}{T(-k)}. \tag{2.14}$$

The scattering coefficients satisfy [2,5,12,13]

$$T(-k) = T(k)^*, \quad R(-k) = R(k)^*, \quad L(-k) = L(k)^*, \qquad k \in \mathbf{R}. \tag{2.15}$$

In the inverse scattering problem of recovery of A, the bound states for the Schrödinger equation do not arise. The absence of bound states for the full-line Schrödinger equation is equivalent for $1/T(k)$ to be nonzero on $\mathbf{I}^+$, where $\mathbf{I}^+ := i(0,+\infty)$ is the positive imaginary axis in $\mathbf{C}^+$. For the half-line Schrödinger equation with the boundary condition (2.7) the absence of bound states is equivalent [3,10,12,13] for $F_\alpha(k)$ to be nonzero on $\mathbf{I}^+$. It is known [3,10,12,13] that either $F_\alpha(0) \neq 0$ or F_α has a simple zero at $k=0$; the former is known as the generic case and the latter as the exceptional case for the half-line Schrödinger equation. For the full-line Schrödinger equation we have $T(0)=0$ generically or $T(0) \neq 0$ in the exceptional case. The exceptional case corresponds to the threshold at which the number of bound states may change by one under a small perturbation. In general, the full-line generic case and the half-line generic case do not occur simultaneously because the former is solely determined by Q whereas the latter jointly by Q and $\cot\alpha$.

3. RELATIVE CONCAVITY, RELATIVE AREA, AND AREA

In this section we introduce the relative radius η of the vocal tract, relate it to the regular solution to the Schrödinger equation, and present various expressions for it involving the Jost solution and the scattering coefficients.

Let r denote the radius of the cross section of the vocal tract so that $A(x) = \pi\, r(x)^2$. Then, we can write the potential Q appearing in (2.4) as

$$Q(x) = \frac{r''(x)}{r(x)}, \qquad x \in (0, l),$$

and hence we can refer to Q also as the relative concavity of the vocal tract. Define

$$\eta(x) := \frac{\sqrt{A(x)}}{\sqrt{A(0)}}, \tag{3.1}$$

or equivalently

$$A(x) = A(0)\,[\eta(x)]^2, \qquad x \in (0, l). \tag{3.2}$$

We can refer to η as the relative radius of the vocal tract and η^2 as the relative area of the tract. From (2.4) we see that η satisfies

$$y'' = Q(x)\, y, \qquad x \in (0, l),$$

with the initial conditions

$$\eta(0) = 1, \quad \eta'(0) = -\cot\alpha.$$

A comparison with (2.9) shows that η is nothing but the zero-energy regular solution, i.e.

$$\eta(x) = \varphi_\alpha(0, x), \qquad x \in (0, l). \tag{3.3}$$

With the help of the expression [3,10,12,13]

$$\varphi_\alpha(k, x) = \frac{1}{2k}\left[F_\alpha(k)\, f(-k, x) - F_\alpha(-k)\, f(k, x)\right],$$

we can write (3.3) as

$$\eta(x) = \dot{F}_\alpha(0)\, f(0, x) - F_\alpha(0)\, \dot{f}(0, x), \qquad x \in (0, l),$$

where an overdot indicates the k-derivative. We can also express η with the help of the scattering coefficients. In the full-line generic case we get [1]

$$\eta(x) = \begin{vmatrix} 0 & -\frac{i}{2}\,\dot{T}(0)\, f(0, x) & i\,\dot{f}(0, x) - \frac{i}{2}\,\dot{R}(0)\, f(0, x) \\ 1 & f(0, 0) & 1 \\ -\cot\alpha & f'(0, 0) & 0 \end{vmatrix}, \qquad x \in (0, l),$$

and in the full-line exceptional case we have [1]

$$\eta(x)=f(0,x)\begin{vmatrix} 0 & 1 & \displaystyle\int_0^x \frac{dz}{[f(0,z)]^2} \\ -1 & f(0,0) & 0 \\ \cot\alpha & 0 & \dfrac{1}{f(0,0)} \end{vmatrix}, \qquad x\in(0,l). \tag{3.4}$$

It is also possible to express η in other useful forms. Let $g(k,x)$ denote the corresponding Jost solution when we replace the zero fragment of Q for $x\in(l,+\infty)$ by another piece which is integrable, has a finite first moment, and does not yield any bound states. Let $\tau(k)$, $\ell(k)$, and $\rho(k)$ be the corresponding transmission coefficient, the left reflection coefficient, and the right reflection coefficient, respectively. In the generic case, i.e. when $\tau(0)=0$, we have [1]

$$\eta(x)=\begin{vmatrix} 0 & -\dfrac{i}{2}\,\dot\tau(0)\,g(0,x) & i\,\dot g(0,x)-\dfrac{i}{2}\,\dot\rho(0)\,g(0,x) \\ 1 & g(0,0) & 1 \\ -\cot\alpha & g'(0,0) & 0 \end{vmatrix}, \qquad x\in(0,l),$$

and in the exceptional case, i.e. when $\tau(0)\neq 0$, we have [1]

$$\eta(x)=g(0,x)\begin{vmatrix} 0 & 1 & \displaystyle\int_0^x \frac{dz}{[g(0,z)]^2} \\ -1 & g(0,0) & 0 \\ \cot\alpha & 0 & \dfrac{1}{g(0,0)} \end{vmatrix}, \qquad x\in(0,l).$$

4. PRESSURE AND VOLUME VELOCITY IN THE VOCAL TRACT

In this section we present the expressions for the pressure and the volume velocity in the vocal tract in terms of the Jost function and the Jost solution. We also introduce various data sets associated with the values of the pressure and the volume velocity at the glottal end of the vocal tract or at the lips.

From (1.2) and (2.1), we see that $V(k,0)=1$. Further, under the reasonable assumption that there is no reflected pressure wave at the mouth and all the pressure wave is transmitted out of the mouth, we get [1]

$$P(k,x) = -\frac{c\mu k\, f(-k,x)}{\sqrt{A(0)}\,\sqrt{A(x)}\,F_\alpha(-k)}, \qquad x \in (0,l), \tag{4.1}$$

$$V(k,x) = -\frac{i\,\sqrt{A(x)}}{\sqrt{A(0)}\,F_\alpha(-k)}\left[f'(-k,x) - \frac{A'(x)}{2\,A(x)}\,f(-k,x)\right], \qquad x \in (0,l), \tag{4.2}$$

where we recall that f is the Jost solution to the Schrödinger equation and F_α is the Jost function appearing in (2.11). Then, the pressure $p(x,t)$ and the volume velocity $v(x,t)$ in the vocal tract are obtained by using (4.1) and (4.2) in (2.1).

The transfer function from the glottis to the point x is defined as $v(x,t)/v(0,t)$, and in our case, as seen from (1.2) and (2.1), that transfer function is nothing but $V(k,x)$. In particular, the transfer function $\mathbf{T}(k,l)$ at the lips is obtained by using (2.10) in (4.2), and we have

$$\mathbf{T}(k,l) = \frac{\sqrt{A(l)}\,e^{-ikl}}{\sqrt{A(0)}\,F_\alpha(-k)}\left[-k + \frac{i}{2}\frac{A'(l)}{A(l)}\right],$$

$$|\mathbf{T}(k,l)|^2 = \frac{A(l)}{A(0)\,|F_\alpha(k)|^2}\left[k^2 + \frac{[A'(l)]^2}{4[A(l)]^2}\right], \qquad k \in \mathbf{R}. \tag{4.3}$$

Using (2.12) in (4.3), we get

$$A(0) = \frac{A(l)}{\lim\limits_{k\to+\infty} |\mathbf{T}(k,l)|}, \tag{4.4}$$

and hence we can write (4.3) as

$$|F_\alpha(k)|^2 = \frac{\lim\limits_{k\to+\infty} |\mathbf{T}(k,l)|^2}{|\mathbf{T}(k,l)|^2}\left[k^2 + \frac{1}{4}\frac{A'(l)^2}{A(l)^2}\right], \qquad k \in \mathbf{R}. \tag{4.5}$$

The impedance at the point x is defined as $p(x,t)/v(x,t)$, which is seen to be equal to $P(k,x)/V(k,x)$ because of (2.1). In particular, with the help of (1.2) we see that the glottal impedance $Z(k,0)$ is equal to $P(k,0)$, and hence we have

$$Z(k,0) = -\frac{c\mu k\, f(-k,0)}{A(0)\,F_\alpha(-k)},$$

$$|Z(k,0)| = \frac{c\mu\,|k|\,|f(k,0)|}{A(0)\,|F_\alpha(k)|}, \qquad k \in \mathbf{R}. \tag{4.6}$$

Using (2.12) and the fact [2,5,10,12,13] that $f(k,0)=1+O(1/k)$ as $k\to\infty$ in $\overline{\mathbf{C}^+}$, from (4.6) we get

$$A(0)=\frac{c\mu}{\lim\limits_{k\to+\infty}|Z(k,0)|}, \tag{4.7}$$

and thus we can write (4.6) in the equivalent form

$$\left|\frac{k\,f(k,0)}{F_\alpha(k)}\right|=\frac{|Z(k,0)|}{\lim\limits_{k\to+\infty}|Z(k,0)|},\qquad k\in\mathbf{R}. \tag{4.8}$$

In a similar manner, we can evaluate $Z(k,l)$, the impedance at the lips, by using (2.10) in (4.1) and (4.2). We get

$$Z(k,l)=\frac{2ic\mu k}{2ik\,A(l)+A'(l)},$$

$$|Z(k,l)|^2=\frac{4c^2k^2\mu^2}{4k^2A(l)^2+A'(l)^2},\qquad k\in\mathbf{R}. \tag{4.9}$$

We have already seen that the pressure at the glottis is the same as the impedance at the glottis because $V(k,0)=1$. So, let us only analyze the pressure at the lips. Using (2.10) in (4.1) we get

$$P(k,l)=-\frac{c\mu k\,e^{-ikl}}{\sqrt{A(0)}\,\sqrt{A(x)}\,F_\alpha(-k)},$$

and hence

$$|P(k,l)|=\frac{c\mu\,|k|}{\sqrt{A(0)\,A(l)}\,|F_\alpha(k)|},\qquad k\in\mathbf{R}. \tag{4.10}$$

Using (2.12) in (4.10), we obtain

$$\sqrt{A(0)\,A(l)}=\frac{c\mu}{\lim\limits_{k\to+\infty}|P(k,l)|}, \tag{4.11}$$

and hence we can write (4.10) in the equivalent form

$$|F_\alpha(k)|=\frac{|k|}{|P(k,l)|}\left(\lim_{k\to+\infty}|P(k,l)|\right),\qquad k\in\mathbf{R}. \tag{4.12}$$

The reflectance at x is defined as the ratio of the left-moving pressure wave to the right-moving pressure wave. As seen from (2.10) and (4.1), there is no left-moving pressure wave at the lips and hence the reflectance at the lips is zero. Since $Q \equiv 0$ for $x < 0$, we have [2,5,12,13]

$$f(-k,x) = \frac{e^{-ikx}}{T(-k)} + \frac{L(-k)\, e^{ikx}}{T(-k)}, \qquad x \le 0,$$

and hence the reflectance at the glottis is equal to $L(-k)$.

The Green's function $\mathbf{G}(k,l;t)$ associated with the vocal-tract acoustics can be defined [8] as the solution to the plasma-wave equation (2.6) with the input (1.2). From (2.1), (2.5), and (4.1) we obtain

$$\mathbf{G}(k,l;t) = \frac{-c\mu k\, e^{ik(ct-l)}}{\sqrt{A(0)}\, F_\alpha(-k)},$$

$$|\mathbf{G}(k,l;t)| = \frac{c\mu\, |k|}{\sqrt{A(0)}\, |F_\alpha(k)|}, \qquad k \in \mathbf{R}. \tag{4.13}$$

Note that the expression $|\mathbf{G}(k,l;t)|$ is independent of t. Using (2.12) in (4.13) we get

$$A(0) = \frac{c^2\mu^2}{\lim_{k\to+\infty} |\mathbf{G}(k,l;t)|^2}, \tag{4.14}$$

and hence hence we can write (4.13) in the equivalent form

$$|F_\alpha(k)| = \frac{|k| \lim_{k\to+\infty} |\mathbf{G}(k,l;t)|}{|\mathbf{G}(k,l;t)|}, \qquad k \in \mathbf{R}. \tag{4.15}$$

5. RECOVERY OF FROM THE JOST FUNCTION

In this section we review the recovery of the potential Q and the boundary parameter $\cot\alpha$ from the absolute value of the Jost function given for $k \in \mathbf{R}^+$. The methods presented include the Gel'fand-Levitan method [3,10,12,13] and the method of [3].

In the absence of bound states, one can use the data $\{|F_\alpha(k)| : \; k \in \mathbf{R}\}$, or equivalently $\{|F_\alpha(k)| : \; k \in \mathbf{R}^+\}$ as a consequence of (2.13), to recover the potential Q and the boundary parameter $\cot\alpha$ in the half-line Schrödinger equation. For example, in the Gel'fand-Levitan method [3,10,12,13], the potential Q is obtained as

$$Q(x) = 2\, \frac{dh_\alpha(x,x^-)}{dx}, \qquad x > 0, \tag{5.1}$$

the boundary parameter $\cot\alpha$ appearing in (2.7) is recovered as

$$\cot\alpha = -h_\alpha(0,0), \tag{5.2}$$

and the regular solution $\varphi_\alpha(k,x)$ appearing in (2.9) is constructed as

$$\varphi_\alpha(k,x) = \cos kx + \int_0^x dy\, h_\alpha(x,y)\, \cos ky, \qquad x \geq 0, \tag{5.3}$$

where $h_\alpha(x,y)$ is obtained by solving the Gel'fand-Levitan integral equation

$$h_\alpha(x,y) + G_\alpha(x,y) + \int_0^x dz\, G_\alpha(y,z)\, h_\alpha(x,z) = 0, \qquad 0 \leq y < x,$$

with the kernel $G_\alpha(x,y)$ given by

$$G_\alpha(x,y) := \frac{2}{\pi}\int_0^\infty dk\, \left[\frac{k^2}{|F_\alpha(k)|^2} - 1\right] \cos(kx)\, \cos(ky).$$

Comparing (5.1)-(5.3) with (2.4), (2.8), (3.2), and (3.3) we see that

$$\frac{[\sqrt{A(x)}]''}{\sqrt{A(x)}} = 2\,\frac{dh_\alpha(x,x^-)}{dx}, \qquad x \in (0,l),$$

$$\frac{A'(0)}{A(0)} = 2\,h_\alpha(0,0),$$

$$\eta(x) = 1 + \int_0^x dy\, h_\alpha(x,y), \qquad x \in (0,l),$$

$$A(x) = A(0)\left[1 + \int_0^x dy\, h_\alpha(x,y)\right]^2, \qquad x \in (0,l).$$

Alternatively, we can proceed [1,3] as follows. Using the data $\{|F_\alpha(k)| : \ k \in \mathbf{R}\}$, we evaluate the integral on the right hand side of

$$\Lambda_\alpha(k) = \frac{1}{\pi i}\int_{-\infty}^\infty \frac{ds}{s-k-i0^+}\left[\frac{s^2}{|F_\alpha(s)|^2} - 1\right], \qquad k \in \overline{\mathbf{C}^+},$$

where the quantity $i0^+$ indicates that the values for real k should be obtained as limits from $\mathbf{C}^+$. The function $\Lambda_\alpha(k)$ is equivalent to

$$\Lambda_\alpha(k) = \frac{k\,f(k,0)}{F_\alpha(k)} - 1, \qquad k \in \overline{\mathbf{C}^+}. \tag{5.4}$$

Next, $F_\alpha(k)$ is obtained from $|F_\alpha(k)|$ by using

$$F_\alpha(k)=k\,\exp\left(\frac{-1}{\pi i}\int_{-\infty}^{\infty}ds\,\frac{\log|s/F_\alpha(s)|}{s-k-i0^+}\right),\qquad k\in\overline{\mathbf{C}^+}.$$

Then, we have

$$f(k,0)=\frac{1}{k}\,F_\alpha(k)\,[1+\Lambda_\alpha(k)],\qquad k\in\overline{\mathbf{C}^+},$$

$$f'(k,0)=i\,F_\alpha(k)\left[1+\frac{1+\Lambda_\alpha(k)}{k}\lim_{k\to\infty}[k\,\Lambda_\alpha(k)]\right],\qquad k\in\overline{\mathbf{C}^+}.$$

$$\cot\alpha=-i\lim_{k\to\infty}[k\,\Lambda_\alpha(k)]\,,\tag{5.5}$$

where the limit in (5.5) can be evaluated in any manner in $\overline{\mathbf{C}^+}$. Having both $f(k,0)$ and $f'(k,0)$ at hand, all the quantities relevant to the scattering for the Schrödinger equation can be constructed. For example, as seen from (2.14), we have

$$T(k)=\frac{2ik}{ik\,f(k,0)+f'(k,0)},\quad L(k)=\frac{ik\,f(k,0)-f'(k,0)}{ik\,f(k,0)+f'(k,0)},$$

$$R(k)=\frac{-ik\,f(-k,0)-f'(-k,0)}{ik\,f(k,0)+f'(k,0)}.$$

Having obtained such quantities, the potential can be constructed via any one of the methods [2,5,12,13] to solve the inverse scattering problem.

6. RECOVERY FROM OTHER DATA SETS

In this section we consider the recovery of Q, η, and A from various data sets introduced in Section 4 related to the values of the pressure and the volume velocity at the ends of the vocal tract. For this purpose we use the result of Section 5 that the data set $\{|F_\alpha(k)|:\ k\in\mathbf{R}^+\}$ uniquely determines the corresponding Q and η, and that the same set along with the value of $A(0)$ uniquely determines A.

(i) Impedance at the lips. If we use the data set $\{|Z(k,l)|:\ k\in\mathbf{R}^+\}$ coming from the impedance at the lips, as seen from (4.9), we can only recover $A(l)$ and $|A'(l)|$. Using (4.9) at two distinct real k-values, say k_1 and k_2, we obtain $A(l)$ and $|A'(l)|$ algebraically as

$$A(l)=\sqrt{\frac{c^2\mu^2}{k_1^2-k_2^2}\left[\frac{k_1^2}{|Z(k_1,l)|^2}-\frac{k_2^2}{|Z(k_2,l)|^2}\right]},\tag{6.1}$$

$$|A'(l)| = \sqrt{\frac{4c^2\mu^2 k_1^2 k_2^2}{k_1^2 - k_2^2}\left[\frac{1}{|Z(k_2,l)|^2} - \frac{1}{|Z(k_1,l)|^2}\right]}. \tag{6.2}$$

No other information can be extracted from $\{|Z(k,l)| : \ k \in \mathbf{R}^+\}$ related to Q, η, or A.

(ii) Impedance at the glottis. The information contained in $\{|Z(k,0)| : \ k \in \mathbf{R}^+\}$ enables us to uniquely construct Q, η, and A, which is seen as follows. From (4.8) and the evenness of $|f(k,0)/F_\alpha(k)|$ in $k \in \mathbf{R}$, we see that the recovery from $\{|Z(k,0)| : \ k \in \mathbf{R}^+\}$ is equivalent to the recovery from $\{|k\, f(k,0)/F_\alpha(k)| : \ k \in \mathbf{R}\}$. Since we assume that the half-line Schrödinger equation with the boundary condition (2.7) does not have any bound states, it is known [3] that $kf(k,0)/F_\alpha(k)$ is analytic in $\mathbf{C}^+$, continuous in $\overline{\mathbf{C}^+}$, nonzero in $\overline{\mathbf{C}^+} \setminus \{0\}$, either nonzero at $k=0$ or has a simple zero there, and

$$\frac{k\, f(k,0)}{F_\alpha(k)} = 1 + O(1/k), \qquad k \to \infty \text{ in } \overline{\mathbf{C}^+}.$$

As a result, we can recover $kf(k,0)/F_\alpha(k)$ for $k \in \overline{\mathbf{C}^+}$ from its amplitude known for $k \in \mathbf{R}$ via

$$\frac{k\, f(k,0)}{F_\alpha(k)} = \exp\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} dt\, \frac{\log|t\, f(t,0)/F_\alpha(t)|}{t-k-i0^+}\right), \qquad k \in \overline{\mathbf{C}^+}.$$

Having constructed $kf(k,0)/F_\alpha(k)$ for $k \in \mathbf{R}$, we can use (5.4) and obtain $\Lambda_\alpha(k)$ for $k \in \mathbf{R}$. Next, by taking the real part of $\Lambda_\alpha(k)$ and using

$$\text{Re}[\Lambda_\alpha(k)] = \frac{k^2}{|F_\alpha(k)|^2} - 1, \qquad k \in \mathbf{R},$$

we construct $|F_\alpha(k)|$ for $k \in \mathbf{R}$. Then, as described in Section 5, we can construct Q and η. Finally, since $A(0)$ is available via (4.7), we can also construct A.

(iii) Pressure at the lips. If we use the data set $\{|P(k,l)| : \ k \in \mathbf{R}^+\}$ coming from the pressure at the lips, as seen from (4.12) we have $\{|F_\alpha(k)| : \ k \in \mathbf{R}^+\}$ at hand, and hence Q and η are uniquely determined. Then, having the value $\eta(l)$ and using (3.1) and (4.11), we obtain $A(0)$ as

$$A(0) = \frac{1}{\eta(l)}\, \frac{c\mu}{\lim\limits_{k\to+\infty} |P(k,l)|}.$$

Thus, A is also uniquely determined.

(iv) Green's function at the lips. If we use the data set $\{|\mathbf{G}(k,l;t)| : \ k\in\mathbf{R}^+\}$ coming from the Green's function at the lips, as seen from (4.15) we have $\{|F_\alpha(k)| : \ k\in\mathbf{R}^+\}$ at hand, and hence Q and η are uniquely determined. From (4.14) we also have $A(0)$, and hence A is also uniquely determined.

(v) Reflectance at the glottis. We know from Section 4 that the reflectance at the glottis is given by $L(-k)$. Because of (2.15), the data sets $\{L(-k) : \ k\in\mathbf{R}^+\}$ and $\{L(k) : k\in\mathbf{R}\}$ are equivalent. The latter set uniquely determines Q, for example, via the Faddeev-Marchenko method [2,5,12,13]. Either of the real and imaginary parts of the reflectance at the glottis known for $k\in\mathbf{R}^+$ also enables us to uniquely construct Q. This is because $\text{Re}[L(k)]$ is an even function of k on $\mathbf{R}$ and $\text{Im}[L(k)]$ is an odd function, and L can be recovered from either its real or imaginary part via the Schwarz integral formula as

$$L(k)=\frac{1}{\pi i}\int_{-\infty}^{\infty}\frac{ds\,\text{Re}[L(s)]}{s-k-i0^+}=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{ds\,\text{Im}[L(s)]}{s-k-i0^+},\qquad k\in\overline{\mathbf{C}^+},$$

due to the fact [2,5] that $L(k)$ is analytic in $\mathbf{C}^+$, continuous in $\overline{\mathbf{C}^+}$, and $o(1/k)$ as $k\to\infty$ in $\overline{\mathbf{C}^+}$. The reflectance contains no information related to $A(0)$ or $\cot\alpha$ given in (2.8). As a result, η is not uniquely determined and we have the corresponding one-parameter family for η with $\cot\alpha$ being the parameter. We also have the corresponding two-parameter family for A, where the parameters can be chosen, for example, as $A(0)$ and $A'(0)$, or as $A(l)$ and $A'(l)$.

(vi) Transfer function at the lips. From (4.5) we see that the data set $\{|\mathbf{T}(k,l)| : k\in\mathbf{R}^+, |A'(l)|/A(l)\}$ uniquely determines $\{|F_\alpha(k)| : \ k\in\mathbf{R}^+\}$ and hence also Q and η. On the other hand, there is the corresponding one-parameter set for A where $A(0)$ can be viewed as a parameter. In view of (4.4) and (4.5), we see that the data set $\{|\mathbf{T}(k,l)| : \ k\in\mathbf{R}^+, A(l), |A'(l)|\}$ uniquely determines each of Q, η, and A. Corresponding to the data set $\{|\mathbf{T}(k,l)| : \ k\in\mathbf{R}^+, A(l)\}$, in general there exists a one-parameter family for each of Q, η, and A, where $|A'(l)|$ can be chosen as the parameter. Corresponding to the data set $\{|\mathbf{T}(k,l)| : \ k\in\mathbf{R}^+\}$, in general there exists a two-parameter family for each of Q, η, and A, where $A(l)$ and $|A'(l)|$ can be chosen as the parameters. Note that we assume that $A(l)$

and $|A'(l)|$ do not change with k, and hence they are constants. As indicated in (i), they can be obtained via (6.1) and (6.2) by measuring the absolute value of the impedance at the lips at two different frequencies.

7. TRANSFER FUNCTION AND NONUNIQUENESS

Let the data set $\{|\mathbf{T}(k,l)| : \ k \in \mathbf{R}^+, \theta\}$ with $\theta := |A'(l)|/[2A(l)]$ correspond to the potential Q, the boundary parameter $\cot\alpha$, the Jost function F_α, the relative radius η, and the area A. From Section 6, we know that all these quantities are uniquely determined with the exception of A, which is determined up to a multiplicative constant $A(0)$. Let the data set $\{|\mathbf{T}(k,l)| : \ k \in \mathbf{R}^+, \tilde\theta\}$ with $\tilde\theta := |\tilde A'(l)|/[2\tilde A(l)]$ correspond to the potential $\tilde Q$, the boundary parameter $\cot\tilde\alpha$, the Jost function $F_{\tilde\alpha}$, the relative radius $\tilde\eta$, and the area $\tilde A$. We assume that $\theta \neq \tilde\theta$. In this section we show that we must have $Q \not\equiv \tilde Q$.

If we had $Q \equiv \tilde Q$, then from (4.5) we would get

$$\frac{|F_\alpha(k)|^2}{|F_{\tilde\alpha}(k)|^2} = \frac{k^2+\theta^2}{k^2+\tilde\theta^2}, \qquad k \in \mathbf{R}, \tag{7.1}$$

or equivalently, with the help of (2.13) we would have

$$\frac{(k+i\tilde\theta)\,F_\alpha(k)}{(k+i\theta)\,F_{\tilde\alpha}(k)} = \frac{(k-i\theta)\,F_{\tilde\alpha}(-k)}{(k-i\tilde\theta)\,F_\alpha(-k)}, \qquad k \in \mathbf{R}.$$

It is known [3,5,10,12,13] that F_α and $F_{\tilde\alpha}$ are analytic in $\mathbf{C}^+$, continuous in $\overline{\mathbf{C}^+}$, and satisfy the asymptotics given in (2.12). Furthermore, since there are no bound states, F_α and $F_{\tilde\alpha}$ are nonzero in $\overline{\mathbf{C}^+} \setminus \{0\}$, and each of them has either a simple zero at $k=0$ or is nonzero there. In fact, as seen from (7.1) we would have $F_\alpha(0)=0$ if and only if $\theta=0$, and similarly $F_{\tilde\alpha}(0)=0$ if and only if $\tilde\theta=0$. From Louiville's theorem it would follow that

$$\frac{(k+i\tilde\theta)\,F_\alpha(k)}{(k+i\theta)\,F_{\tilde\alpha}(k)} = 1, \qquad k \in \mathbf{C}. \tag{7.2}$$

From (3.18) of [3] we have

$$\mathrm{Re}\left[\frac{i\,F_{\tilde\alpha}(k)}{F_\alpha(k)}\right] = \frac{k\,(\cot\tilde\alpha - \cot\alpha)}{|F_\alpha(k)|^2}, \qquad k \in \mathbf{R}. \tag{7.3}$$

Using (7.2) in (7.3), with the help of (2.12) we would get

$$\theta - \tilde{\theta} = \cot \tilde{\alpha} - \cot \alpha,$$

$$|F_\alpha(k)|^2 = k^2 + \theta^2, \qquad k \in \mathbf{R},$$

implying $F_\alpha(k) = k + i\theta$ for $k \in \overline{\mathbf{C}^+}$. The unique potential corresponding to this particular $F_\alpha(k)$ is given by $Q \equiv 0$ and the boundary parameter $\cot \alpha$ is given by $\cot \alpha = -\theta$. Similarly, we would also have $\cot \tilde{\alpha} = -\tilde{\theta}$, $F_{\tilde{\alpha}}(k) = k + i\tilde{\theta}$, $\tilde{Q} \equiv 0$. Then, using (3.4) we would get

$$\eta(x) = 1 + \theta x, \quad \tilde{\eta}(x) = 1 + \tilde{\theta} x, \qquad x \in (0, l),$$

$$A(x) = A(0)\,[1 + \theta x]^2, \quad \tilde{A}(x) = \tilde{A}(0)\,[1 + \tilde{\theta} x]^2, \qquad x \in (0, l), \tag{7.4}$$

where $A(0)$ and $\tilde{A}(0)$ are arbitrary positive constants. Now, using $\theta = |A'(l)|/[2A(l)]$ in (7.4), we would obtain $\theta = 0$, and similarly we would get $\tilde{\theta} = 0$. However, this contradicts our assumption $\theta \neq \tilde{\theta}$, and hence we must have $Q \not\equiv \tilde{Q}$.

Acknowledgments. The research leading to this article was supported in part by the National Science Foundation under grant DMS-0204437 and the Department of Energy under grant DE-FG02-01ER45951.

REFERENCES

[1] T. Aktosun, *Inverse scattering for vowel articulation with frequency-domain data,* IMA Preprint #1996, October 2004.

[2] T. Aktosun and M. Klaus, *Chapter 2.2.4, Inverse theory: problem on the line,* in: E. R. Pike and P. C. Sabatier (eds.), *Scattering,* Academic Press, London, 2001, pp. 770–785.

[3] T. Aktosun, and R. Weder, *Inverse spectral-scattering problem with two sets of discrete spectra for the radial Schrödinger equation,* IMA preprint #1960, (2004).

[4] R. Burridge, *The Gelfand-Levitan, the Marchenko, and the Gopinath-Sondhi integral equations of inverse scattering theory, regarded in the context of inverse impulse-response problems,* Wave Motion **2**, 305–323 (1980).

[5] K. Chadan and P. C. Sabatier, *Inverse problems in quantum scattering theory,* 2nd ed., Springer, New York, 1989.

[6] G. Fant, *Acoustic theory of speech production,* Mouton, The Hague, 1970.

[7] J. L. Flanagan, *Speech analysis synthesis and perception,* 2nd ed., New York, Springer, 1972.

[8] B. J. Forbes, E. R. Pike, and D. B. Sharp, *The acoustical Klein-Gordon equation: The wave-mechanical step and barrier potential functions,* J. Acoust. Soc. Am. **114**, 1291–1302 (2003).

[9] L. Gårding, *The inverse of vowel articulation,* Ark. Mat. **15**, 63–86 (1977).

[10] I. M. Gel'fand and B. M. Levitan, *On the determination of a differential equation from its spectral function,* Am. Math. Soc. Transl. (ser. 2) **1**, 253–304 (1955).

[11] B. Gopinath and M. M. Sondhi, *Determination of the shape of the human vocal tract shape from acoustical measurements,* Bell Sys. Tech. J. **49**, 1195–1214 (1970).

[12] B. M. Levitan, *Inverse Sturm-Liouville problems,* VNU Science Press, Utrecht, 1987.

[13] V. A. Marchenko, *Sturm-Liouville operators and applications,* Birkhäuser, Basel, 1986.

[14] J. R. McLaughlin, *Analytical methods for recovering coefficients in differential equations from spectral data,* SIAM Rev. **28**, 53–72 (1986).

[15] P. Mermelstein, *Determination of the vocal-tract shape from measured formant frequencies,* J. Acoust. Soc. Am. **41**, 1283–1294 (1967).

[16] Rakesh, *Characterization of transmission data for Webster's horn equation,* Inverse Problems **16**, L9–L24 (2000).

[17] M. R. Schroeder, *Determination of the geometry of the human vocal tract by acoustic*

measurements, J. Acoust. Soc. Am. **41**, 1002–1010 (1967).

[18] J. Schroeter and M. M. Sondhi, *Techniques for estimating vocal-tract shapes from the speech signal,* IEEE Trans. Speech Audio Process. **2**, 133–149 (1994).

[19] M. M. Sondhi, *A survey of the vocal tract inverse problem: theory, computations and experiments,* in: F. Santosa, Y. H. Pao, W. W. Symes, and C. Holland (eds.), *Inverse problems of acoustic and elastic waves,* SIAM, Philadelphia, 1984, pp. 1–19.

[20] M. M. Sondhi and B. Gopinath, *Determination of vocal-tract shape from impulse response at the lips,* J. Acoust. Soc. Am. **49**, 1867–1873 (1971).

[21] M. M. Sondhi and J. R. Resnick, *The inverse problem for the vocal tract: numerical methods, acoustical experiments, and speech synthesis,* J. Acoust. Soc. Am. **73**, 985–1002 (1983).

[22] K. N. Stevens, *Acoustic phonetics,* MIT Press, Cambridge, MA, 1998.

[23] W. W. Symes, *On the relation between coefficient and boundary values for solutions of Webster's Horn equation,* SIAM J. Math. Anal. **17**, 1400–1420 (1986).