Source-tract interaction with prescribed vocal fold motion

Richard S. McGowan^{a)}

CReSS LLC, 1 Seaborn Place, Lexington, Massachusetts 02420

Michael S. Howe

College of Engineering, Boston University, 110 Cummings Street, Boston, Massachusetts 02215

(Received 15 April 2011; revised 20 January 2012; accepted 26 January 2012)

An equation describing the time-evolution of glottal volume velocity with specified vocal fold motion is derived when the sub- and supra-glottal vocal tracts are present. The derivation of this Fant equation employs a property explicated in Howe and McGowan [(2011) *J. Fluid Mech.* **672**, 428–450] that the Fant equation is the adjoint to the equation characterizing the matching conditions of sub- and supra-glottal Green's functions segments with the glottal segment. The present aeroacoustic development shows that measurable quantities such as input impedances at the glottis, provide the coefficients for the Fant equation when source-tract interaction is included in the development. Explicit expressions for the Green's function are not required. With the poles and zeros of the input impedance functions specified, the Fant equation can be solved. After the general derivation of the Fant equation, the specific cases where plane wave acoustic propagation is described either by a Sturm-Liouville problem or concatenated cylindrical tubes is considered. Simulations show the expected skewing of the glottal volume velocity pulses depending on whether the fundamental frequency is below or above a sub- or supra-glottal formant. More complex glottal wave forms result when both the first supra-glottal fundamental frequencies are high and close to the first sub-glottal formant. © 2012 Acoustical Society of America. [DOI: 10.1121/1.3685824]

PACS number(s): 43.70.Bk [AL]

Pages: 2999–3016

NOMENCLATURE

a_0, a_1 = glottal area coefficients, Eq. (62)
 A_n = cross-sectional areas
 A_g = glottis cross-sectional area
 A_{sub} = tract area upstream of glottis
 A_{sup} = tract area downstream of glottis
 $\mathcal{A}_n, \mathcal{B}_n$ = amplitudes, Eq. (56)
 $B, \hat{B}$ = fluctuating total enthalpy
 $\bar{B}, \hat{\bar{B}}$ = fluctuating total enthalpy
 B_{sub} = total enthalpy due to lung
 B_{ω} = total entropy due to vorticity
 c_o = speed of sound
 $\mathcal{C}$ = correction factor, Eq. (54)
 f_0 = vocal fold frequency
 F_n = recursive function, Eq. (A9)
 $\mathcal{F}$ = aeroacoustic source function
 $\mathcal{F}_{\text{sub}}$ = sub-glottal source function
 $\hat{g}_0 = \hat{G}_0$ phase factor removed
 G = Green's function, Eq. (17)
 $G_0, \hat{G}_0 = G$ for homogeneous boundary conditions
 $G_M, \hat{G}_M = G$ for inhomogeneous boundary conditions
 $\mathcal{G}$ = Sturm-Liouville Green's function
 $\hat{h}_0$ = adjoint of $\hat{g}_0$
 $\mathcal{H}, \mathcal{K}$ = coefficients, Eq. (A1)
 $\mathcal{I}_n$ = constants, Eq. (53)
 J_n, K_n = integration functions, Eqs. (59) and (60)
 $\mathcal{J}_n$ = constants, Eq. (53)

K = order 1 constant, Eq. (1)
 l = cylindrical tube section length
 $\bar{\ell}(\tau)$ = glottal end correction, Eq. (10)
 $\bar{L}_{\text{sup}}$ = supra-glottal length
 $\bar{L}_{\text{sub}}$ = sub-glottal length
 $\mathcal{L}$ = glottal inductance, Eq. (1)
 p = pressure
 p_f = far-field fluctuating pressure
 $p_{f\text{max}}$ = maximum p_f
 p_{sub} = pressure upstream of the glottis
 $Q, \bar{Q}$ = glottal volume velocity
 $\bar{Q}$ = dummy variable, Eq. (36)
 $Q_{\ell}, \bar{Q}_{\ell}$ = volume velocity at the lips
 Q_{norm} = normalized Q
 Q_{max} = maximum Q_{norm}
 r = distance from lips to far field point
 $r_n = \psi_n - i\zeta_n$ zeros of $\hat{Z}_{\text{sup}}$
 $R_n = \Psi_n - i\Xi_n$ zeros of $\hat{Z}_{\text{sub}}$
 $\mathcal{R}_n$ = constants, Eq. (53)
 R' = coefficient for viscous loss, Eq. (1)
 S = control surface
 S_L = S through air of bronchial tubes
 $s_n = \omega_n - i\gamma_n$ poles of $\hat{Z}_{\text{sup}}$
 $S_n = \Omega_n - i\Gamma_n$ poles of $\hat{Z}_{\text{sub}}$
 S_g = glottal region surface
 S_{sub} = control surface orthogonal to the tract axis below the glottis
 S_{subwall} = S along solid sub-glottal walls
 S_{sup} = control surface orthogonal to the tract axis above the glottis
 S_{supwall} = S along solid supra-glottal walls

^{a)}Author to whom correspondence should be addressed. Electronic mail: rsmcgowan@cressllc.net.

t = time
 $\hat{u}; \hat{v}$ = Sturm-Liouville solutions
 U_σ = glottal jet velocity
 v = magnitude of $\mathbf{v}$
 $\mathbf{v}$ = velocity
 v_n = normal velocity to surfaces
 V = control volume
 W = Wronskian
 $\mathbf{x} = (x_1, x_2, x_3)$ Cartesian coordinates
 $\mathbf{y} = (y_1, y_2, y_3)$ Cartesian coordinates
 $\mathbf{y}_M$ = matching region
 y_n = normal to surfaces
 $\mathbf{Y}$ = Kirchhoff vector, Eqs. (8) and (9)
 Z_n^{in} = input impedance
 Z_n = tube impedance
 $\hat{Z}_{\text{sup}}$ = supra-glottal input impedance
 $\hat{Z}_{\text{sub}}$ = sub-glottal input impedance
 $\mathcal{Z}_{w+l}$ = boundary condition operator, Eq. (3)
 $\alpha, \hat{\alpha}$ = Green's function coefficients, Eq. (8)
 $\beta, \hat{\beta}$ = Green's function coefficients, Eq. (8)
 κ = Sturm-Liouville parameter
 λ = eigenvalues for Sturm-Liouville
 λ_n = parameter, Eq. (A1)
 Λ = length scale, Eq. (50)
 ρ = air density
 ρ_0 = mean ρ
 σ = jet contraction ratio
 τ = time
 θ = independent variable
 $\phi, \hat{\phi}$ = velocity potential
 ϕ_n = phase, Eq. (56)
 ω = radian frequency
 $\boldsymbol{\omega}$ = vorticity, curl $\mathbf{v}$

I. INTRODUCTION

Source-filter theory was formulated as a first approximation describing acoustic air motion in the vocal tract, and it was a particularly compelling model for voiced vowel production. At the time of Fant's 1960 monograph it was known that the sub-glottal and supra-glottal vocal tracts' fluid mechanics would have an effect on the volume velocity pulse emanating from the glottis (Fant, 1960). For instance, it was known that the effect of the supra-glottal vocal tract is particularly strong during transition to or from obstruent consonants due to strong flow resistance. Glottal flow is affected by the fluid mechanics of the surrounding vocal tract because glottal flow depends on the pressures upstream and downstream of the glottis. The equation for the volume flow through the glottis, $Q = Q(t)$, derived by Fant did not account for the surrounding vocal tract. This equation is

$$\mathcal{L} \frac{dQ}{dt} + K \rho_0 \frac{Q^2}{2A_g^2} + R'Q = p_{\text{sub}}, \quad (1)$$

where $\mathcal{L}, A_g \equiv A_g(t)$ are respectively the "inductance" and cross-sectional area of the glottis at its narrowest section, ρ_0 is the mean air density, K is an order 1 constant, the coeffi-

cient R' represents the effects of viscous losses, and p_{sub} is the pressure just upstream of the glottis (Fant, 1960). This equation assumes that the supra-glottal pressure is zero, and in the case that it is not zero, the sub-glottal pressure, p_{sub} , can be replaced by the difference between sub- and supra-glottal pressures, or trans-glottal pressure.

Equation (1) actually is an approximation for the time evolution of a jet through a time-varying orifice, $A_g(t)$, into free space. But, for a jet exiting from a semi-infinite tube of cross-sectional area A_{sub} into a semi-infinite tube of area A_{sup} , there are pressures upstream and downstream of the glottis related to the glottal volume flow itself that need to be taken into account using impedances $\rho_0 c_0 / A_{\text{sub}}$ and $\rho_0 c_0 / A_{\text{sup}}$, where c_0 is the speed of sound, assuming that the tubes support only plane wave propagation. Titze showed that, under these conditions, a non-negligible term, $Q \rho_0 c_0 (1/A_{\text{sub}} + 1/A_{\text{sup}})$, should be added to the left-hand side of Eq. (1) (Titze, 1984).

Realistic results for finite tubes should involve the inverse Fourier transforms of input impedance functions of the sub- and supra-glottal vocal tract resonators. Such effects have been studied numerically. Flanagan and Landgraf (1968) and Flanagan and Cherry (1968) added an electrical line analog for the supra-glottal vocal tract to a one mass-model of the vocal folds to account for feedback at the glottis provided by a finite length supra-glottal tract. Thus, the effects of the supra-glottal impedances on the glottal volume velocity, which included changes in the vocal folds' oscillation, were simulated by these authors. Ishizaka and Flanagan also modeled the effects of supra-glottal impedances in their time domain numerical simulation of the two-mass model of vocal fold oscillation coupled with a vocal tract (Ishizaka and Flanagan, 1972).

The effect of the vocal tract on the voice source with prescribed vocal fold motion or glottal area time variation has been considered in order to simplify the problem for greater physical understanding. [Models of source-tract interaction with specified glottal area functions are termed level I models by Titze (2008).] According to Rothenberg (1981), when the fundamental frequency is well below the first formant, the supra-glottal vocal tract input impedance can be approximated by an inductive element. He showed that this has the effect of skewing the glottal pulse so that the glottal flow decreases faster than it increases for a glottal area pulse that is symmetric in time. This phenomenon was confirmed by Ananthapadmanabha and Fant (1982), while they also studied the ripples on the volume velocity source caused by source-tract interaction (see also Fant, 1986). From their numerical work they concluded that it is necessary to include only the effects of the lowest formant in each the sub-glottal and supra-glottal tracts to adequately account for source-tract interaction during speech. Titze has gone on to show how realistic sub- and supra-glottal resonant tube shapes affect the amplitudes of the harmonics of the glottal flow in level I interactions using time-domain simulations of acoustic propagation in the sub- and supra-glottal vocal tract, particularly when the fundamental frequency is close to a formant frequency (Titze, 2008).

The approach that we have adopted is to derive the equation for the volume flow through the glottis, $Q(t)$, in the

time domain (Howe and McGowan, 2007, 2011a), using integral solutions to the equations of air motion with Green's function kernels. This has two advantages. The first advantage is that a closed form ordinary differential equation for Q is obtained with known degrees of approximation. The resulting differential equation for Q is termed a Fant equation, and is a generalization of Eq. (1). The relative sizes of terms in the Fant equation can be estimated, and the equation approximated after the full equation for Q is derived. In particular, the importance of source-tract interaction can be judged. The second advantage of our approach is that a time-domain numerical solution of the differential equation for Q is possible without simulating wave propagation, while still including the effects the sub- and supra-glottal vocal tracts on Q .

In the previous work of Howe and McGowan (2007, 2011a) it was possible to obtain explicit expressions for the Green's functions in infinitely long sub- and supra-glottal tracts, in the limit that these tracts allow only plane wave propagation. A Green's function could also be derived for simplified sub-glottal systems (Howe and McGowan, 2009) and special supra-glottal vocal tracts, such as straight tubes (Howe and McGowan, 2011b). The Green's functions from these previous works were derived using asymptotic matching procedures between neighboring regions: between sub-glottal and glottal regions and between supra-glottal and glottal regions. Flow in the sub- and/or supra-glottal tracts was assumed to be acoustic, and the flow was assumed incompressible in the glottal region under the acoustic compactness approximation. However, it is not possible to write an explicit Green's function for human vocal tracts in the generality necessary for the great variety of speech sounds.

Mathematical progress has recently been made in understanding how to derive Fant equations that account for the effect of sub- and supra-glottal wave propagation on Q (Howe and McGowan, 2011b). As prelude to the discussion of this progress, an important fact is introduced: while the wave-operator is a self-adjoint operator, the problem is not self-adjoint because of the moving boundaries, i.e., the moving vocal folds (Morse and Feshbach, 1953). Thus, it is appropriate to use the Green's function for the adjoint problem, denoted G here, as the kernel in the integral solution (Morse and Feshbach, 1953; Howe and McGowan, 2009, 2011a). The mathematical progress is in showing that the integro-differential operator for the Fant equation is the adjoint to one derived from matching procedures for the Green's function between neighboring regions for a Helmholtz resonator or a straight-tube resonator (Howe and McGowan, 2011b). This result is easily generalized here. Thus, an explicit form of the Green's function need not be available, only the asymptotic matching conditions for the Green's function G are needed to derive the operator that appears in the Fant equation for Q .

The first-order integro-differential equation that the Green's function satisfies comes from considering its compact approximation, and matching segments of the Green's function between incompressible and acoustic regions. In other words, the Green's function is considered to be composed of segments valid in three different regions of the

vocal tract: sub-glottal, glottal, and supra-glottal. These segments are related by asymptotic matching conditions, which provide for continuity of pressure and volume velocity. Matching the Green's function segments between the sub-glottal region and the glottal region and between the supra-glottal region and the glottal region results in the desired first-order integro-differential equation.

In the following section a theory of source-tract interaction with prescribed vocal fold motion is presented. Initially, minimal assumptions are made about the nature of the Green's function. However, it is necessary to make the nature of the problem more specific only to finalize the form of the Fant equation. Further, the Fant equation derived here will be such that only measured quantities appear as coefficients or integration kernels. At the end of this section, functional forms for the input impedances are derived assuming a Sturm-Liouville problem or a series of cylindrical concatenated tubes for the sub- and supra-glottal vocal tracts so that the Fant equation can be solved for a great variety of vocal tracts. Simulations based on the Fant equation will be presented in Sec. III. Finally, a conclusion is provided in Sec. IV.

The Fant equation derived in this paper will be applied to the level I problem of source-tract interaction, where the glottal areas are specified. The application to the level I problems will show some of the expected trends of volume velocity pulse shape. However, these will not be realistic simulations of source-tract interaction, because fluid-structure interaction in the glottal region needs to be considered for realistic simulations. Those considerations are beyond the scope of this paper. Section IV indicates how the same Fant equation can be applied in solving full source-tract interaction problems.

II. THEORY

A. Preliminaries

The supra-glottal vocal tract, the glottal air channel, and the sub-glottal tract are partially bounded by solid surfaces about a common Cartesian axis in the x_1 direction (Fig. 1). The cross-sectional areas near the glottis are such that acoustic propagation occurs only along the x_1 direction for the frequencies considered. The area of the sub-glottal tube adjoining the glottal region is A_{sub} , the area of the supra-glottal vocal tract adjoining the glottal region is A_{sup} , and the area of the narrowest part of the glottis is $A_g = A_g(t)$. The coordinate system, (x_1, x_2, x_3) has its origin along the vocal tract axis at the anterior end of the glottis.

In this work we are concerned with disturbances to the air in the vocal tract caused by vocal fold vibration: voiced speech. While the Mach number of fluid flow can change from zero to the order of 0.1 in the glottal region, the resulting fluctuating pressures are small compared to the ambient atmospheric pressure. For the low Mach number, homeotropic flow considered here, pressure is a function of density alone, and the fluctuating quantity of interest is the total enthalpy, $B = \int \frac{dp}{\rho} + \frac{1}{2} v^2$, where $v = |\mathbf{v}|$ with $\mathbf{v}$ the air particle velocity. Further, dissipation is confined to the lateral boundaries with strong shear layers and wall vibration. For

FIG. 1. Schematic of the vocal tract. The control volume is bounded by the solid walls of the vocal tract and the surfaces between the lips and through the bronchial tubes, denoted S_L . The surfaces S_{sub} and S_{sup} are used to segment the control volume into three portions: the sub-glottal tract, glottal region, and supra-glottal tract.

such flows the vortex sound equation can be written (Howe, 1998, 2002; Howe and McGowan, 2007).

$$\left(\frac{1}{c_o^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_j^2} \right) B = \text{div}(\omega \wedge \mathbf{v}), \quad (2)$$

where $\omega = \nabla \wedge \mathbf{v}$ is vorticity, t is time, and the repeated subscript j implies summation.

A control volume $V(\tau)$, where τ is time, encompassing the supra-laryngeal vocal tract, glottal region, and a large portion of the sub-glottal vocal tract is now specified (Fig. 1). The control volume is bounded by the surface $S(\tau)$, which covers the opening at the lips and lines the vocal tract into the bronchial tree until a bronchial generation is encountered where acoustic disturbances created at the glottis become negligible due to viscous loss. The surface is completed with cross-sectional planes through each of the bronchial tubes in that generation. The part of S that span across the bronchial tubes is a disconnected surface denoted S_L .

The boundary condition for B on the surface $S(\tau)$ in the glottal region is $\frac{\partial B}{\partial y_n} = -\frac{\partial v_n}{\partial \tau}$ with y_n the normal direction to the solid surfaces (directed into the control volume), and v_n fluid particle velocity normal to the surface. It is assumed that a real-valued operator, $\mathcal{Z}_{\omega+l}(\mathbf{y}, \tau)$, exists such that

$$\frac{\partial v_n(\mathbf{y}, \tau)}{\partial \tau} + \int_{-\infty}^{\infty} \mathcal{Z}_{\omega+l}(\mathbf{y}, \tau') B(\mathbf{y}, \tau - \tau') d\tau' = 0, \quad (3)$$

on the solid surfaces of the sub- and supra glottal vocal tract walls and on the surface between the lips. $\mathbf{y} = (y_1, y_2, y_3)$ are Cartesian coordinates. The boundary condition on the surface S_L that cuts through the bronchi will be determined by the rate of contraction of the lungs and the resulting flow of air through that surface. This boundary condition will account for the source of B from the lungs. Here, the conception of the bronchial tree with an upper conductive region and lower transitional and respiratory region is adopted, as in Ishizaka *et al.* (1976). The surface S_L is supposed to be somewhere in the conductive region.

It is assumed that the Green's function, $G(\mathbf{x}, t; \mathbf{y}, \tau)$, for the adjoint problem exists for $\mathbf{x}$ and $\mathbf{y}$ in the control volume and for the boundary conditions specified below. The Green's function for the adjoint problem must be employed because the glottal configuration changes rapidly and nonli-

nearly with time (Howe and McGowan, 2011a,b). When the governing equations for B are in terms of the field variables $\mathbf{x}$ and t , the Green's function, G , is a solution to the adjoint problem in variables $\mathbf{y}$ and τ . Thus, the Green's function for the adjoint problem is the advanced potential Green's function $G(\mathbf{x}, t; \mathbf{y}, \tau)$, that satisfies

$$\left(\frac{1}{c_o^2} \frac{\partial^2}{\partial \tau^2} - \frac{\partial^2}{\partial y_j^2} \right) G = \delta(\mathbf{x} - \mathbf{y}) \delta(t - \tau),$$

$$G = 0 \text{ for } \tau > t, \quad (4)$$

where $\delta(\cdot)$ is the Dirac delta function. G is chosen to satisfy $\frac{\partial G}{\partial y_n} = 0$ on two surfaces: the solid surface in the glottal region and the surface S_L that cuts through the bronchi. Further, with consideration of the boundary conditions for B expressed by Eq. (3), it is required that

$$\frac{\partial G(\mathbf{x}, t; \mathbf{y}, \tau)}{\partial y_n} - \int_{-\infty}^{\infty} \mathcal{Z}_{\omega+l}(\mathbf{y}, \tau') G(\mathbf{x}, t; \mathbf{y}, \tau + \tau') d\tau' = 0, \quad (5)$$

on the solid surfaces of the sub- and supra-glottal vocal tract, and on the surface between the lips. G represents “incoming” waves as a function of $\mathbf{y}$ and τ that vanish after convergence onto the source $\delta(\mathbf{x} - \mathbf{y}) \delta(t - \tau)$ at $\tau = t$. The Green's function for the adjoint problem, G , is referred to here as *the* Green's function.

Equations (2) and (4) are combined (Morse and Feshbach, 1953; Howe and McGowan, 2007, 2011a; Howe, 2008) using Green's theorem, the radiation condition and the momentum equation, to supply the causal solution of Eq. (2) at low Mach numbers in the form

$$B(\mathbf{x}, t) = \int_{-\infty}^{\infty} \oint_{S(\tau)} \left[G \frac{\partial \mathbf{v}}{\partial \tau} + B \frac{\partial G}{\partial \mathbf{y}} - \mathbf{v} \frac{\partial G}{\partial \mathbf{y}} \wedge \omega \right] \cdot d\mathbf{S}(\mathbf{y}) d\tau$$

$$- \int_{-\infty}^{\infty} \int_{V(\tau)} \frac{\partial G}{\partial \mathbf{y}} \cdot \omega \wedge \mathbf{v} d^3 \mathbf{y} d\tau, \quad (6)$$

where $\mathbf{x}$ is in the control volume $V(\tau)$, ν is the kinematic viscosity of air, and $d\mathbf{S}(\mathbf{y})$ are vector surface elements of $S(\tau)$ directed into the control volume. The boundary conditions specified for B and G can be employed to simplify this formal representation of the B field. With boundary

conditions in Eqs. (3) and (5) applied to the solid surfaces, in the sub- and supra-glottal regions, denoted $S_{\text{subwall}}(\tau)$ and $S_{\text{supwall}}(\tau)$, respectively, the first two terms of the surface integral in Eq. (6) are zero. Because the surface segment between the lips is not solid and Eqs (3) and (5) apply to this section of surface, all three terms in the surface integral vanishes at that surface segment. Applying the remaining boundary conditions on the solid surface in the glottal region and on S_L , Eq. (6) reduces to,

$$\begin{aligned} B(\mathbf{x}, t) = & \int_{-\infty}^{\infty} \oint_{S_L} G \frac{\partial \mathbf{v}}{\partial \tau} \cdot d\mathbf{S}(\mathbf{y}) d\tau \\ & - \int_{-\infty}^{\infty} \int_{S_g(\tau)+} \nu \frac{\partial G}{\partial \mathbf{y}} \wedge \boldsymbol{\omega} \cdot d\mathbf{S}(\mathbf{y}) d\tau \\ & - \int_{-\infty}^{\infty} \int_{V(\tau)} \frac{\partial G}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v} d^3\mathbf{y} d\tau, \end{aligned} \quad (7)$$

where $S_g(\tau)$ is the solid surface in the glottal region.

Subsequent discussion of the Green's function is with the control volume divided into three contiguous regions: the sub-glottal region, the supra-glottal region, and the glottal region. This entails the addition of two additional surfaces bounding these regions. These are portions of planes orthogonal to y_1 , one portion just below the glottis, S_{sub} , and another just above the glottis S_{sup} (Fig. 1). Surface S_{sub} separates the sub-glottal and glottal regions, has directed normal into the sub-glottal tract, and area A_{sub} , while surface S_{sup} separates the supra-glottal and glottal regions, has directed normal into the supra-glottal tract, and area A_{sup} . Boundary conditions at these planes will result from matching conditions between adjoining regions. When the frequency is sufficiently low, the flow in the glottal region is approximately incompressible and the time derivatives in Eqs. (2) and (4) can be neglected. This is the acoustic compactness approximation.

In Sec. II B an equation analogous to Eq. (1), a Fant equation, but including the back-reaction of the sub- and supra-glottal vocal tracts, is derived. The subsequent development of the Fant equation does not require that an explicit Green's function be derived. This is important for two reasons. First, there is an infinite number of supra-glottal vocal tract shapes possible in speech and singing, and only a few are amenable to analytic solution for the supra-glottal Green's function segment. Also, the form of the sub-glottal segment of the Green's function is unknown because of the complicated nature of the sub-glottal system (e.g., [Habib et al., 1994](#)). Instead of the explicit expressions for the Green's functions, input impedances at the glottis are used to account for acoustic back-reaction in the glottal region.

The derivation of the Fant equation in Sec. II B generalizes the development for the straight tube in the Appendix of [Howe and McGowan \(2011b\)](#). In Sec. II C low-frequency approximations are applied in the sub- and supra-glottal vocal tracts in order to derive expressions for the input impedances in terms of their poles and zeros in the complex frequency plane. These expressions are used to compute the effect of the sub- and supra-glottal tracts on the voice source when glottal motion is specified in Sec. III.

B. The Fant equation

1. Green's function matching conditions

Under the acoustic compactness approximation with $\mathbf{y}$ in the glottal region,

$$G(\mathbf{x}, t; \mathbf{y}, \tau) = \alpha(\tau) + \beta(\tau)Y(\mathbf{y}, \tau), \quad (8)$$

where $\alpha(\tau), \beta(\tau)$ are functions to be determined, $\nabla^2 Y(\mathbf{y}, \tau) = 0$, and $Y(\mathbf{y}, \tau)$ denotes the velocity potential of flow at unit speed within the duct just upstream of the glottis such that the normal derivative $\partial Y / \partial y_n = 0$ on the instantaneous surface $S(\tau)$ of the glottis and the neighboring duct, so that

$$Y(\mathbf{y}, \tau) \sim \begin{cases} y_1 - \bar{\ell}(\tau), & |y_1| \gg \sqrt{A_{\text{sub}}} \text{ for } y_1 < 0 \\ \left(\frac{A_{\text{sub}}}{A_{\text{sup}}}\right)y_1, & |y_1| \gg \sqrt{A_{\text{sup}}} \text{ for } y_1 > 0. \end{cases} \quad (9)$$

The dependence of α and β on the field point $\mathbf{x}$ and time t is suppressed. The length $\bar{\ell}(\tau)$ is a time-dependent "end correction" ([Howe, 1998](#); [Howe and McGowan, 2007](#)),

$$\bar{\ell}(\tau) = \int_{-\infty}^{\infty} \left(\frac{\partial Y(\mathbf{y}, \tau)}{\partial y_1} - \frac{A_{\text{sub}}}{A_{\text{sup}}} H(y_1) - H(-y_1) \right) dy_1, \quad (10)$$

where $H(\cdot)$ is the Heaviside step function and the integration is along any path through the glottis. The asymptotic form of $Y(\mathbf{y}, \tau)$ implies that the tubes neighboring the sub- and supra-glottal tracts are approximately straight along the y_1 axis and do not support acoustic cross-modes.

Initially, we work in the frequency domain by taking Fourier transforms with respect to τ . For instance,

$$\hat{G}(\mathbf{x}, t; \mathbf{y}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\mathbf{x}, t; \mathbf{y}, \tau) e^{-i\omega\tau} d\tau, \quad (11)$$

Matching the Green's function segments between the sub-glottal and glottal regions, and between the supra-glottal and glottal regions, using Eqs. (8) and (9), provides

$$\begin{aligned} \lim_{y_1 \rightarrow -0} \hat{G}(\mathbf{x}, t; \mathbf{y}, \omega) & \approx \lim_{\substack{y \rightarrow 0 \\ y_1 < 0}} \hat{G}(\mathbf{x}, t; \mathbf{y}, \omega) \\ & \equiv \hat{G}(\mathbf{x}, t; -\mathbf{0}, \omega) \\ & = \hat{\alpha}(\omega) - (\hat{\bar{\ell}} * \hat{\beta})(\omega), \end{aligned} \quad (12)$$

$$\begin{aligned} \lim_{y_1 \rightarrow +0} \hat{G}(\mathbf{x}, t; \mathbf{y}, \omega) & \approx \lim_{\substack{y \rightarrow 0 \\ y_1 > 0}} \hat{G}(\mathbf{x}, t; \mathbf{y}, \omega) \\ & \equiv \hat{G}(\mathbf{x}, t; +\mathbf{0}, \omega) = \hat{\alpha}(\omega), \end{aligned} \quad (13)$$

where $(\hat{\bar{\ell}} * \hat{\beta})(\omega) = \int_{-\infty}^{\infty} \hat{\bar{\ell}}(\omega') \hat{\beta}(\omega - \omega') d\omega'$. It follows from Eq. (9) that the variation in $\hat{G}$ with y_2 and y_3 as $y_1 \rightarrow \pm 0$ is small, and thus, the first approximations made in Eqs. (12) and (13) are valid. Matching the spatial derivatives gives the relations

$$\lim_{y_1 \rightarrow -0} \frac{\partial \hat{G}(\mathbf{x}, t; \mathbf{y}, \omega)}{\partial y_1} \approx \frac{\partial \hat{G}(\mathbf{x}, t; -\mathbf{0}, \omega)}{\partial y_1} = \hat{\beta}(\omega), \quad (14)$$

$$\lim_{y_1 \rightarrow +0} \frac{\partial \hat{G}(\mathbf{x}, t; \mathbf{y}, \omega)}{\partial y_1} \approx \frac{\partial \hat{G}(\mathbf{x}, t; +\mathbf{0}, \omega)}{\partial y_1} = \left(\frac{A_{\text{sub}}}{A_{\text{sup}}} \right) \hat{\beta}(\omega). \quad (15)$$

2. Green's function decomposition

The Green's functions segments in the sub- and supra-glottal vocal tracts can be written as a sum of two functions, G_0 and G_M . G_0 denotes the Green's functions for a point source in either the sub- or supra-glottal regions [i.e., each G_0 satisfies Eq. (4)] with boundary conditions specified below. The G_0 functions satisfy the same homogeneous boundary conditions that G satisfies on the control surface S . The G_0 functions also satisfy homogeneous matching conditions where their regions join with the glottal region. The particular matching condition is the "hard-wall" condition that is familiar in the source-filter theory of speech production,

$$\frac{\partial G_0(\mathbf{x}, t; \mathbf{y} = \pm \mathbf{0}, \tau)}{\partial y_1} = 0. \quad (16)$$

The other functions, G_M , satisfy the wave equation [i.e., each G_M satisfies Eq. (4), but with the delta function replaced by 0], but account for boundary inhomogeneities on S_{sub} and S_{sup} that result from matching Green's function segments across neighboring regions [see Eqs. (21) and (22) below]. Otherwise, G_M satisfies the same homogeneous boundary conditions as G and G_0 on surface S . Thus,

$$G(\mathbf{x}, t; \mathbf{y}, \tau) = G_0(\mathbf{x}, t; \mathbf{y}, \tau) + G_M(\mathbf{y}_M | \mathbf{x}, t; \mathbf{y}, \tau), \quad (17)$$

where $\mathbf{y}_M$ denotes the region of matching: $\mathbf{y}_M = -\mathbf{0}$ denotes the matching region between the sub-glottal tract and glottal region, and $\mathbf{y}_M = +\mathbf{0}$ denotes the matching region between the supra-glottal tract and glottal region. $G_M(\mathbf{y}_M | \mathbf{x}, t; \mathbf{y}, \tau)$ is used only with $\mathbf{y}_M$, $\mathbf{x}$, and $\mathbf{y}$ in the same region. Each $G_0(\mathbf{x}, t; \mathbf{y}, \tau)$ segment is defined in both the sub-glottal and supra-glottal regions as long as it is understood that these functions are zero for $\mathbf{x}$ and $\mathbf{y}$ in different regions. It should be clear in subsequent discussion which segment is being referred to.

The functions G_0 in both the sub- and supra-glottal vocal tracts are assumed to be homogeneous in time. In particular, the sub- and supra-glottal tracts are quasi-steady when compared to the time scale of vocal fold vibration. This would permit G_0 to be written as a function of $t - \tau$ instead of two separate arguments t and τ . However, to explicitly note that the variable t remains in the Fourier transform of $G_0(\mathbf{x}, t; \mathbf{y}, \tau)$ with respect to τ , albeit as part of a simple phase factor, there will be no change in the argument list to denote homogeneity in time. A closely related function, $\hat{g}_0 = \hat{g}_0(\mathbf{x}, \mathbf{y}, \omega)$, can be defined without the phase factor involving the t variable,

$$\hat{g}_0(\mathbf{x}, \mathbf{y}, \omega) = e^{i\omega t} \hat{G}_0(\mathbf{x}, t; \mathbf{y}, \omega). \quad (18)$$

It follows from Eq. (4) that

$$\left(\left(\frac{\omega}{c_0} \right)^2 + \frac{\partial^2}{\partial y_j^2} \right) \hat{g}_0(\mathbf{y}, \mathbf{y}', \omega) = -\delta(\mathbf{y} - \mathbf{y}'). \quad (19)$$

The boundary conditions for the $\hat{g}_0$ segments are the Fourier transforms of the homogeneous boundary conditions for G_0 expressed in Eq. (5). Thus $\hat{g}_0$ is a Green's function for the Helmholtz equation.

Returning to the matching conditions for the $\hat{G}_M$ functions on S_{sub} and S_{sup} , it follows from Eqs. (14)–(17), that

$$\hat{\beta}(\omega) = \frac{\partial \hat{G}(\mathbf{x}, t; -\mathbf{0}, \omega)}{\partial y_1} = \frac{\partial \hat{G}_M(\mathbf{y}_M = -\mathbf{0} | \mathbf{x}, t; -\mathbf{0}, \omega)}{\partial y_1}, \quad (20)$$

$$\left(\frac{A_{\text{sub}}}{A_{\text{sup}}} \right) \hat{\beta}(\omega) = \frac{\partial \hat{G}_M(\mathbf{y}_M = +\mathbf{0} | \mathbf{x}, t; +\mathbf{0}, \omega)}{\partial y_1}. \quad (21)$$

From Eqs. (20) and (21) it is seen that $\hat{\beta}(\omega)$ and $\left(\frac{A_{\text{sub}}}{A_{\text{sup}}} \right) \hat{\beta}(\omega)$ are "velocity sources" for $\hat{G}_M(\mathbf{y}_M = -\mathbf{0} | \mathbf{x}, t; \mathbf{y}, \omega)$ and $\hat{G}_M(\mathbf{y}_M = +\mathbf{0} | \mathbf{x}, t; \mathbf{y}, \omega)$ at the glottal boundaries of their respective regions. Because the functions G_M satisfy the homogeneous wave equation, the functions $\hat{G}_M$ satisfy the Helmholtz equation

$$\left(\left(\frac{\omega}{c_0} \right)^2 + \frac{\partial^2}{\partial y_j^2} \right) \hat{G}_M(\mathbf{y}'_M = \pm \mathbf{0} | \mathbf{x}, t; \mathbf{y}', \omega) = 0, \quad (22)$$

where the $\hat{G}_M$ satisfy the same homogeneous boundary conditions that the $\hat{g}_0$ do, except on S_{sub} or S_{sup} , where they are given by Eqs. (20) and (21). Multiplying Eq. (22) by $\hat{g}_0$ and Eq. (19) by $\hat{G}_M$, subtracting, integrating over either the sub- or supra-glottal region, and applying Green's theorem, provides

$$\begin{aligned} \hat{G}_M(\mathbf{y}_M = -\mathbf{0} | \mathbf{x}, t; \mathbf{y}, \omega) &= -\hat{g}_0(\mathbf{y}, -\mathbf{0}, \omega) \hat{\beta}(\omega) A_{\text{sub}} \\ &= -\hat{g}_0(-\mathbf{0}, \mathbf{y}, \omega) \hat{\beta}(\omega) A_{\text{sub}} \end{aligned} \quad (23)$$

$$\begin{aligned} \hat{G}_M(\mathbf{y}_M = +\mathbf{0} | \mathbf{x}, t; \mathbf{y}, \omega) &= -\hat{g}_0(\mathbf{y}, +\mathbf{0}, \omega) \left(\frac{A_{\text{sub}}}{A_{\text{sup}}} \right) \hat{\beta}(\omega) A_{\text{sup}} \\ &= -\hat{g}_0(\mathbf{y}, +\mathbf{0}, \omega) \hat{\beta}(\omega) A_{\text{sub}} \\ &= -\hat{g}_0(+\mathbf{0}, \mathbf{y}, \omega) \hat{\beta}(\omega) A_{\text{sub}} \end{aligned} \quad (24)$$

where the final equalities in Eqs. (23) and (24) follow because the $\hat{g}_0$ are symmetric in the spatial variables (Morse and Feshbach, 1953).

3. Input impedances and relations to Green's function segments

Input impedance at the supra-glottal surface S_{sup} can be defined as the Fourier transform of the pressure signal at S_{sup} resulting from a delta function volume velocity pulse input to the supra-glottal tract at S_{sup} . Equivalently, let $\hat{\phi}(\mathbf{y}, \omega)$ be the velocity potential in a neighborhood of S_{sup} . Then the Fourier transform of a brief volume velocity input at S_{sup} is $A_{\text{sup}} \frac{\partial \hat{\phi}(+\mathbf{0}, \omega)}{\partial y_1}$ and the resulting pressure at S_{sup} is $-i\rho_0 \omega \hat{\phi}(+\mathbf{0}, \omega)$. The supra-glottal input impedance, $\hat{Z}_{\text{sup}}(\omega)$ can be defined as the ratio of the pressure to the volume velocity at S_{sup}

$$\hat{Z}_{\text{sup}}(\omega) \equiv \frac{-i\omega\rho_0\hat{\phi}(+\mathbf{0}, \omega)}{A_{\text{sup}}\left(\frac{\partial\hat{\phi}(+\mathbf{0}, \omega)}{\partial y_1}\right)}. \quad (25)$$

Here it is assumed that the flow in a neighborhood of S_{sup} is dominated by potential flow. Input impedance is not usually measured in this direct manner with brief volume velocity pulses, but indirectly with known impedances in series with the system (e.g., [Ishizaka et al., 1976](#)). A similar expression holds for sub-glottal input impedance, $\hat{Z}_{\text{sub}}$,

$$\hat{Z}_{\text{sub}}(\omega) \equiv -\frac{-i\omega\rho_0\hat{\phi}(-\mathbf{0}, \omega)}{A_{\text{sub}}\left(\frac{\partial\hat{\phi}(-\mathbf{0}, \omega)}{\partial y_1}\right)}. \quad (26)$$

The minus sign in Eq. (26) is because the sub-glottal tract in the $-y_1$ direction with respect to its surface S_{sub} at $y_1 = -0$.

Let functions $\hat{h}_0$ be the adjoints for $\hat{g}_0$, so that the functions $\hat{g}_0$ are the Fourier transforms of the Green's function for the direct problems in the sub- and supra-glottal vocal tracts. By the properties of adjoint pairs ([Morse and Feshbach, 1953](#)), $\hat{h}_0(\mathbf{y}, \mathbf{y}', \omega) = \hat{g}_0(\mathbf{y}', \mathbf{y}, -\omega)$ and $h_0(\mathbf{y}, \mathbf{y}', \tau) = g_0(\mathbf{y}', \mathbf{y}, \tau)$. It follows by taking the inverse Fourier transform of Eq. (19) that

$$\begin{aligned} \lim_{y_1 \rightarrow -0} \frac{\partial h_0(\mathbf{y}, -\mathbf{0}, \tau)}{\partial y_1} &= \lim_{y_1 \rightarrow -0} \frac{\partial g_0(-\mathbf{0}, \mathbf{y}, -\tau)}{\partial y_1} \\ &= -\frac{1}{A_{\text{sub}}} \delta(-\tau), \end{aligned} \quad (27)$$

$$\begin{aligned} \lim_{y_1 \rightarrow +0} \frac{\partial h_0(\mathbf{y}, +\mathbf{0}, \tau)}{\partial y_1} &= \lim_{y_1 \rightarrow +0} \frac{\partial g_0(+\mathbf{0}, \mathbf{y}, -\tau)}{\partial y_1} \\ &= -\frac{1}{A_{\text{sup}}} \delta(-\tau), \end{aligned} \quad (28)$$

so that

$$\lim_{y_1 \rightarrow +0} \left(-A_{\text{sup}} \frac{\partial h_0(\mathbf{y}, +\mathbf{0}, \tau)}{\partial y_1} \right)$$

and

$$\lim_{y_1 \rightarrow -0} \left(-A_{\text{sub}} \frac{\partial h_0(\mathbf{y}, -\mathbf{0}, \tau)}{\partial y_1} \right)$$

are delta function “volume velocity” inputs to the supra- and sub-glottal systems that have

$$\lim_{y_1 \rightarrow +0} \left(\rho_0 \frac{\partial h_0(\mathbf{y}, +\mathbf{0}, \tau)}{\partial \tau} \right)$$

and

$$\lim_{y_1 \rightarrow -0} \left(\rho_0 \frac{\partial h_0(\mathbf{y}, -\mathbf{0}, \tau)}{\partial \tau} \right)$$

“pressure” outputs. It follows from these considerations and Eqs. (25)–(28) that

$$\begin{aligned} \hat{Z}_{\text{sup}}(\omega) &= i\omega\rho_0 \lim_{y_1 \rightarrow +0} \hat{h}_0(\mathbf{y}, +\mathbf{0}, \omega) \\ &= i\omega\rho_0 \lim_{y_1 \rightarrow +0} \hat{g}_0(+\mathbf{0}, \mathbf{y}, -\omega), \end{aligned} \quad (29)$$

$$\begin{aligned} \hat{Z}_{\text{sub}}(\omega) &= -i\omega\rho_0 \lim_{y_1 \rightarrow -0} \hat{h}_0(\mathbf{y}, -\mathbf{0}, \omega) \\ &= -i\omega\rho_0 \lim_{y_1 \rightarrow -0} \hat{g}_0(-\mathbf{0}, \mathbf{y}, -\omega). \end{aligned} \quad (30)$$

Note that both $\hat{Z}_{\text{sup}}(\omega) = \hat{Z}_{\text{sup}}^*(-\omega)$ and $\hat{Z}_{\text{sub}}(\omega) = \hat{Z}_{\text{sub}}^*(-\omega)$. These identities follow from the reality of the inverse Fourier transform of $\hat{g}_0(\mathbf{x}, \mathbf{y}, \omega)$. Combining Eqs. (23), (24), (29), and (30) results in

$$\hat{G}_M(\mathbf{y}_m = +\mathbf{0}|\mathbf{x}, t; +\mathbf{0}, \omega) = -\frac{\hat{Z}_{\text{sub}}^*(\omega)}{i\omega\rho_0} \hat{\beta}(\omega) A_{\text{sub}}, \quad (31)$$

$$\hat{G}_M(\mathbf{y}_m = -\mathbf{0}|\mathbf{x}, t; -\mathbf{0}, \omega) = \frac{\hat{Z}_{\text{sub}}^*(\omega)}{i\omega\rho_0} \hat{\beta}(\omega) A_{\text{sub}}. \quad (32)$$

4. Combining results for the Fant equation

Subtracting Eq. (12) from Eq. (13), and using Eqs. (17), (31), and (32) gives

$$\begin{aligned} i\omega \left(\hat{\ell} * \hat{\beta}(\omega) \right) + \frac{A_{\text{sub}}}{\rho_0} \hat{\beta}(\omega) (\hat{Z}_{\text{sub}}(\omega) + \hat{Z}_{\text{sub}}^*(\omega)) \\ = i\omega (\hat{G}_0(\mathbf{x}, t; +\mathbf{0}, \omega) - \hat{G}_0(\mathbf{x}, t; -\mathbf{0}, \omega)). \end{aligned} \quad (33)$$

Taking the inverse Fourier transform,

$$\begin{aligned} \frac{\partial \bar{\ell}(\tau) \beta(\tau)}{\partial \tau} + \frac{A_{\text{sub}}}{2\pi\rho_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \beta(\tau') (\hat{Z}_{\text{sub}}^*(\omega) \\ + \hat{Z}_{\text{sub}}^*(\omega)) e^{i\omega(\tau-\tau')} d\omega d\tau' \\ = \frac{\partial G_0(\mathbf{x}, t; +\mathbf{0}, \tau)}{\partial \tau} - \frac{\partial G_0(\mathbf{x}, t; -\mathbf{0}, \tau)}{\partial \tau}. \end{aligned} \quad (34)$$

The Fant equation for the time evolution of glottal volume velocity Q is now derived from Eq. (34). First, it is assumed that the sum of all aeroacoustic source terms, $\mathcal{F}(\tau)$, can be related to $B(\mathbf{x}, t)$ for $\mathbf{x}$ in the supra-glottal tract by an equation of the form,

$$B(\mathbf{x}, t) = \int_{-\infty}^{\infty} \beta(\tau) \mathcal{F}(\tau) d\tau. \quad (35)$$

Equation (35) is shown true below in a discussion of the aeroacoustic sources in Sec. II B 5. We assert, along with the auxiliary condition $\lim_{\tau \rightarrow -\infty} \bar{Q}(\tau) = 0$, that the equation

$$\begin{aligned} \bar{\ell}(\tau) \frac{\partial \bar{Q}(\tau)}{\partial \tau} - \frac{A_{\text{sub}}}{2\pi\rho_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{Q}(\tau') (\hat{Z}_{\text{sub}}(\omega) \\ + \hat{Z}_{\text{sub}}(\omega)) e^{-i\omega(\tau-\tau')} d\omega d\tau' = \mathcal{F}(\tau) \end{aligned} \quad (36)$$

is satisfied by the glottal volume velocity $Q(\tau)$, i.e., that $\bar{Q}(\tau) = Q(\tau)$, provided that Eq. (35) holds. Note that the operator on the left-hand side of Eq. (36) is the adjoint to the operator on the left-hand side of Eq. (34).

To prove that $\bar{Q} = Q$, multiply Eq. (34) by $\bar{Q}(\tau)$, Eq. (36) by $\beta(\tau)$, add the results and integrate with respect to τ from $-\infty$ to ∞ . With $\bar{Q}(\tau) \rightarrow 0$ as $\tau \rightarrow -\infty$ and $\beta(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$, the symmetry properties of $\hat{Z}_{\text{sup}}(\omega)$ and $\hat{Z}_{\text{sub}}(\omega)$, and noting that $G_0(\mathbf{x}, t; -\mathbf{0}, \tau) = 0$ for $\mathbf{x}$ in the supra-glottal vocal tract results in

$$\int_{-\infty}^{\infty} \frac{\partial G_0(\mathbf{x}, t; +\mathbf{0}, \tau)}{\partial \tau} \bar{Q}(\tau) d\tau = - \int_{-\infty}^{\infty} \beta(\tau) \mathcal{F}(\tau) d\tau, \quad (37)$$

where the symmetry properties of $\hat{Z}_{\text{sub}}(\omega)$ and $\hat{Z}_{\text{sup}}(\omega)$ noted after Eq. (32) have been used. Multiplying both sides of

Eq. (37) by A_{sup} , taking the derivative with respect to x_1 , and the limit as $x_1 \rightarrow +0$ gives

$$\begin{aligned} \lim_{x_1 \rightarrow +0} A_{\text{sup}} \int_{-\infty}^{\infty} \frac{\partial G_0(\mathbf{x}, t; +\mathbf{0}, \tau)}{\partial x_1} \frac{\partial \bar{Q}(\tau)}{\partial \tau} d\tau \\ = \lim_{x_1 \rightarrow +0} A_{\text{sup}} \frac{\partial B(\mathbf{x}, t)}{\partial x_1}, \end{aligned} \quad (38)$$

where integration by parts has been performed and Eq. (35) has been invoked. As before, let ϕ be the velocity potential in a neighborhood of S_{sup} and use the definition of stagnation enthalpy, B (namely, $B = -\partial\phi/\partial t$ in a region without vorticity). The right-hand side of Eq. (38) is

$$\lim_{x_1 \rightarrow +0} A_{\text{sup}} \frac{\partial B(\mathbf{x}, t)}{\partial x_1} = \lim_{x_1 \rightarrow +0} A_{\text{sup}} \frac{\partial^2 \phi}{\partial x_1 \partial t} = -\frac{\partial Q(t)}{\partial t}, \quad (39)$$

because G_0 satisfies Eq. (4) and is symmetric in the spatial variables (Morse and Feshbach, 1953)

$$\lim_{x_1 \rightarrow +0} A_{\text{sup}} \frac{\partial G_0(\mathbf{x}, t; +\mathbf{0}, \tau)}{\partial x_1} = -\delta(t - \tau). \quad (40)$$

It follows that the left-hand side of Eq. (38) is $-\frac{\partial \bar{Q}(t)}{\partial t}$. Thus, with the condition $\lim_{\tau \rightarrow -\infty} Q(\tau) = 0$, $\bar{Q} = Q$, if Eq. (35) is true. Equation (36) is the desired Fant equation under the same condition.

5. Aeroacoustic sources

In order to show that Eq. (35) is true, it is necessary to examine Eq. (7). Initially the glottis is closed and the lungs are presumed to contract slowly, which means that lungs are a source of total enthalpy $B_{\text{sub}}(\mathbf{y}, \tau)$ in the sub-glottal tract. As discussed previously in Sec. II A, the active lung contraction occurs outside of the control volume $V(\tau)$, so that the effect of this source is taken into account by the flow of air passing through the surface S_L inside the bronchial tubes, and the first term in Eq. (7) provides the mathematical expression for this source.

From Eq. (7) the enthalpy just below the glottis due to the lung source is

$$\begin{aligned} B_{\text{sub}}(-\mathbf{0}, t) &= \int_{-\infty}^{\infty} \left[\oint_{S_L} G(-\mathbf{0}, t; \mathbf{y}, \tau) \frac{\partial \mathbf{v}(\mathbf{y}, \tau)}{\partial \tau} \cdot \mathbf{S}(\mathbf{y}) \right. \\ &\quad \left. - \nu \oint_{S_{\text{subwall}}} \frac{\partial G(-\mathbf{0}, t; \mathbf{y}, \tau)}{\partial \mathbf{y}} \wedge \boldsymbol{\omega} \cdot d\mathbf{S}(\mathbf{y}) \right] d\tau \\ &\quad - \int_{-\infty}^{\infty} \int_{V_{\text{sub}}} \frac{\partial G(-\mathbf{0}, t; \mathbf{y}, \tau)}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v} d^3 \mathbf{y} d\tau \\ &= - \int_{-\infty}^{\infty} \left[\oint_{S_L} \frac{\partial G(-\mathbf{0}, t; \mathbf{y}, \tau)}{\partial \tau} \cdot \mathbf{v}(\mathbf{y}, \tau) \mathbf{S}(\mathbf{y}) \right. \\ &\quad \left. + \nu \oint_{S_{\text{subwall}}} \frac{\partial G(-\mathbf{0}, t; \mathbf{y}, \tau)}{\partial \mathbf{y}} \wedge \boldsymbol{\omega} \cdot \mathbf{S}(\mathbf{y}) \right] d\tau \\ &\quad - \int_{-\infty}^{\infty} \int_{V_{\text{sub}}} \frac{\partial G(-\mathbf{0}, t; \mathbf{y}, \tau)}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v} d^3 \mathbf{y} d\tau, \end{aligned} \quad (41)$$

where V_{sub} is the portion of the control volume in the sub-glottal tract. Thus the total enthalpy just below the glottis increases due to the first term in the integral, but it is reduced from what it otherwise would be at the glottis by viscous loss in the bronchial tubes and by vortex shedding at any sharp edges in the bronchial tree. The latter effect should be

small compared to the viscous losses when the flow is slow enough so that the Reynolds numbers are small. Define $p_{\text{sub}}(\tau) = \rho_0 B_{\text{sub}}(-\mathbf{0}, \tau)$, which is approximately the slowly increasing pressure just below the glottis before it opens. When $p_{\text{sub}}(\tau = 0)$ is attained the vocal folds begin to vibrate, and it is assumed that the lungs continue to supply the same level of total enthalpy below the glottis.

When the glottis opens, the B at a field point $\mathbf{x}$ in the supra-glottal tract due to the sub-glottal enthalpy is

$$\begin{aligned} B_{\text{sub}}(\mathbf{x}, t) &= \frac{1}{\rho_0} \int_{-\infty}^{\infty} \oint_{S_{\text{sup}}} p_{\text{sub}}(\tau = 0) \frac{\partial G(\mathbf{x}, t; \mathbf{0}, \tau)}{\partial \mathbf{y}} \cdot \mathbf{S}(\mathbf{y}) d\tau \\ &= \frac{A_{\text{sub}}}{\rho_0} \int_{-\infty}^{\infty} p_{\text{sub}}(\tau = 0) \beta(\tau) d\tau, \end{aligned} \quad (42)$$

where continuity of $B_{\text{sub}}(-\mathbf{0}, \tau)$ from just below the glottis to just above the acoustically compact glottis is assumed and Eq. (15) has been used for the second equality. Define $\mathcal{F}_{\text{sub}}(\tau) = \frac{A_{\text{sub}}}{\rho_0} p_{\text{sub}}(\tau) H(\tau)$, and then Eq. (42) can be written

$$B_{\text{sub}}(\mathbf{x}, t) = \int_{-\infty}^{\infty} \mathcal{F}_{\text{sub}}(\tau) \beta(\tau) d\tau, \quad (43)$$

where $\mathbf{x}$ is in the supra-glottal tract and $t > \tau$.

In this development, the volume changes in the vocal folds, as represented by the first term on the right-hand side in Eq. (7), are unknown and not considered here. They can be added in a later development when they become known. The viscous term, the third term on the right-hand side of Eq. (7), is small compared to the volume integral of the vorticity interaction term in the glottis, which is considered next, except when the glottis is nearly closed.

The volume integral on the right-hand side of Eq. (7) is a dipole source, and its main contribution is from the jet within distance $\sim \sqrt{A_g(\tau)}$ downstream of the glottis. Its strength is modulated by the variations in the glottis area $A_g(\tau)$ produced by fold vibration (Howe and McGowan, 2007, 2011a). The field, B_{ω} , resulting from the vortex source is written,

$$\begin{aligned} B_{\omega}(\mathbf{x}, t) &= - \int_{-\infty}^{\infty} \int_{V(\tau)} \left(\frac{\partial G(\mathbf{x}, \mathbf{y}, t - \tau)}{\partial \mathbf{y}} \right) \cdot \boldsymbol{\omega} \wedge \mathbf{v}(\mathbf{y}, \tau) d^3 \mathbf{y} d\tau \\ &\approx - \int_{-\infty}^{\infty} \int_{V(\tau)} \beta(\tau) \frac{\partial Y(\mathbf{y}, \tau)}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v}(\mathbf{y}, \tau) d^3 \mathbf{y} d\tau, \end{aligned} \quad (44)$$

where Eq. (8) has been used. Define

$$\mathcal{F}(\tau) = \mathcal{F}_{\text{sub}}(\tau) - \int_{V(\tau)} \frac{\partial Y(\mathbf{y}, \tau)}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v}(\mathbf{y}, \tau) d^3 \mathbf{y}. \quad (45)$$

With Eqs. (42)–(45), Eq. (7) reduces to the simple form of Eq. (35) for $\mathbf{x}$ in the supra-glottal vocal tract and $t > \tau > 0$. Thus the Fant equation, Eq. (36), is true.

6. Final form of the Fant equation

The dependence on Q in Eq. (36) is made more explicit when the vortex sound source term, the second term on the

right-hand side of Eq. (45) is evaluated in terms of $Q(\tau)$. This has been done previously in [Howe and McGowan \(2007\)](#) under the assumptions that glottal jet velocity, $U_\sigma(\tau)$, does not vary substantially over a distance of glottal length scale $\sim \sqrt{A_g(\tau)}$, the glottal area is small compared to the supra-glottal tract area, $A_g(\tau) \ll A_{\text{sup}}$, and that the Strouhal number $\sim f_0 \sqrt{A_g(\tau)} / U_\sigma(\tau)$ is small, where f_0 is the fundamental frequency of voicing.

$$\int_{V(\tau)} \frac{\partial Y(\mathbf{y}, \tau)}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v}(\mathbf{y}, \tau) d^3 \mathbf{y} \approx \frac{A_{\text{sup}} U_\sigma^2(\tau)}{2} = \frac{A_{\text{sup}}}{2 A_g^2(\tau)} \frac{Q^2(\tau)}{\sigma^2(\tau)}, \quad (46)$$

where $\sigma(t)$ is the jet contraction ratio. With Eqs. (45) and (46), Eq. (36), can be written

$$\begin{aligned} \bar{\ell}(\tau) \frac{\partial Q(\tau)}{\partial \tau} + \frac{A_{\text{sup}}}{2 A_g^2(\tau)} \frac{Q^2(\tau)}{\sigma^2(\tau)} - \frac{A_{\text{sub}}}{2 \pi \rho_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(\tau') \\ \times (\hat{Z}_{\text{sub}}(\omega) + \hat{Z}_{\text{sup}}(\omega)) e^{-i\omega(\tau-\tau')} d\omega d\tau' \\ = A_{\text{sub}} \frac{P_{\text{sub}}}{\rho_0}, \end{aligned} \quad (47)$$

for $\tau > 0$. Equation (47) is the desired Fant equation.

C. Low frequency impedance models

In order to examine the effects of the sub-glottal and supra-glottal tracts, the input impedances need to be specified. The functional forms of the input impedances $\hat{Z}_{\text{sub}}(\omega)$ and $\hat{Z}_{\text{sup}}(\omega)$ are now examined for a range of frequencies low enough that only plane wave propagation can occur in the sub- and supra-glottal vocal tracts. The case where loss is neglected is examined first, and then the effects of losses at solid surfaces and due to radiation into the atmosphere are accounted for as a perturbation to these results. Note that while loss caused by changes in entropy have not been included as part of the derivation thus far, their effect on propagation are small so that the Fant equation is a valid approximation, even when heat conduction losses at vocal tract walls are included as an empirical adjustment to the bandwidths of the input impedances, as occurs below.

The input impedances for the sub- and supra-glottal vocal tracts are examined in the case that the equation of motion in the frequency domain is described by the Fourier transform of Eq. (2), but with $\frac{\partial \hat{B}(\mathbf{x}, \omega)}{\partial x_n} = 0$ on solid surfaces. Because only plane wave propagation is assumed to occur, we can write $\hat{B} = \hat{B}(x_1, \omega)$. The boundary conditions at the ends of the sub- and supra-glottal tracts are now considered. For the supra-glottal vocal tract of length $\bar{L}_{\text{sup}} = \bar{L}_{\text{sup}}(\omega)$ including a frequency-dependent end correction ([Howe and McGowan, 2011b](#)), these conditions are

$$\frac{d\hat{B}(+0, \omega)}{dx_1} = 0 \quad \text{and} \quad \hat{B}(\bar{L}_{\text{sup}}, \omega) = 0. \quad (48)$$

For the sub-glottal tract the simplified model of [Ishizaka et al. \(1976\)](#) is used for low frequency sub-glottal propagation (see Fig. 6 of [Ishizaka et al., 1976](#)). The distance from

the glottis to a large expansion in bronchial area is about 20 cm for male adults and is denoted $\bar{L}_{\text{sub}}$, which also includes a frequency dependent end correction. The boundary conditions in the sub-glottal tract are similar to those of the supra-glottal tract,

$$\frac{d\hat{B}(-0, \omega)}{dx_1} = 0 \quad \text{and} \quad \hat{B}(-\bar{L}_{\text{sub}}, \omega) = 0. \quad (49)$$

In the Appendix two models are considered for low-frequency propagation: (1) Sturm-Liouville problems, and (2) concatenated cylindrical tubes, which is commonly used to compute speech acoustics when cross-sectional vocal tract areas are specified. In both cases it is shown that the infinity of poles and zeros of the input impedance are simple and real, and they interleave one another. From these facts, it is possible to write them explicitly as products of factors ([Whittaker and Watson, 1927](#)). For the supra-glottal input impedance

$$\hat{Z}_{\text{sup}}(\omega) = -\frac{i\omega\rho_0\Lambda(\omega)}{2\pi A_{\text{sup}}} \prod_{n=1}^{\infty} \left(\frac{(1-\omega/\psi_n)(1+\omega/\psi_n)}{(1-\omega/\omega_n)(1+\omega/\omega_n)} \right), \quad (50)$$

where $\omega_n, \psi_n > 0$, and $\omega_n < \psi_n < \omega_{n+1}$. $\Lambda(\omega) = \bar{L}_{\text{sup}}$ is the length scale that is necessary in order that this input impedance agrees with the case of a tube of constant cross-sectional area of length $\bar{L}_{\text{sup}}$ in the Appendix of [Howe and McGowan \(2011b\)](#). The symmetry in the products about $\omega = 0$ is because the inverse Fourier transforms of the purely imaginary $\hat{Z}_{\text{sup}}(\omega)$ is real for real ω . A similar expression defines the input impedance for the sub-glottal tract. In this form, the input impedances are defined on the complex ω plane with poles and zeros on the real axis.

Now that the behavior of the input impedances is understood for the lossless case, loss is included as a perturbation. This has the effect of moving the zeros and poles into the lower half of the complex ω plane in such a way that the ordering of the real parts of these points is preserved from the lossless case.¹ Generalizing to the case where losses are present provides the following functional forms for the input impedances,

$$\hat{Z}_{\text{sup}}(\omega) = -\frac{i\omega\rho_0\bar{L}_{\text{sup}}}{2\pi A_{\text{sup}}} \prod_{n=1}^{\infty} \left(\frac{(1-\omega/r_n)(1+\omega/r_n^*)}{(1-\omega/s_n)(1+\omega/s_n^*)} \right), \quad (51)$$

$$\hat{Z}_{\text{sub}}(\omega) = -\frac{i\omega\rho_0\bar{L}_{\text{sub}}}{2\pi A_{\text{sub}}} \prod_{n=1}^{\infty} \left(\frac{(1-\omega/R_n)(1+\omega/R_n^*)}{(1-\omega/S_n)(1+\omega/S_n^*)} \right), \quad (52)$$

$r_n = \psi_n - i\zeta_n, s_n = \omega_n - i\gamma_n, R_n = \Psi_n - i\Xi_n, S_n = \Omega_n - i\Gamma_n$. Here $\Psi_n, \Omega_n, \gamma_n, \zeta_n, \Gamma_n, \Xi_n > 0$, and $\Omega_n < \Psi_n < \Omega_{n+1}$.

Because this development is for frequencies for which there is little variation of $\hat{B}(x_1, \omega)$ across the tracts' cross-sections, this representation is useful only for a finite number

of zeros and poles. Further, because we are interested in the interaction with the volume velocity at the glottis, which generally has most of its energy in a frequency range well below the third resonant frequencies of either the sub- or supra-glottal vocal tracts, the above representations will be restricted to the first two pole pairs and the first zero pair. (A pair of points are where the real parts are non-zero, have opposite sign, and equal magnitude.)

The Fant equation, Eq. (47), can now be rewritten in the low frequency approximation. First, the integral with respect to ω in the third term on the left-hand side of Eq. (47) is evaluated. With all the poles in the lower half plane, this integral is zero for $\tau - \tau' < 0$ because the contour can be closed in the upper-half of the complex frequency plane. For $\tau - \tau' > 0$ the contour can be closed in the lower half-plane and evaluated by the method of residues. Let

$$\begin{aligned}\mathcal{J}_1 &= \frac{s_1(s_1 - r_1)(s_1 + r_1^*)\mathcal{C}(s_1)}{(s_1 + s_1^*)(s_1 - s_2)(s_1 + s_2^*)}, \\ \mathcal{J}_2 &= \frac{s_2(s_2 - r_1)(s_2 + r_1^*)\mathcal{C}(s_2)}{(s_2 + s_2^*)(s_2 - s_1)(s_2 + s_1^*)}, \\ \mathcal{R}_n &= \text{Re}(\mathcal{J}_n), \quad \mathcal{I}_n = \text{Im}(\mathcal{J}_n),\end{aligned}\quad (53)$$

where $\mathcal{C}(\omega)$ is correction factor for higher order zeros and poles. To a first approximation,

$$\mathcal{C}(\omega) = 1 + i2\omega \left(\frac{\gamma_3}{|r_2|^2} - \frac{\xi_2}{|s_3|^2} \right). \quad (54)$$

With these definitions and fixing the frequency dependent length, $\bar{L}_{\text{sup}}$,

$$\begin{aligned}\int_{-\infty}^{\infty} \hat{Z}_{\text{sup}}(\omega) e^{-i\omega(\tau-\tau')} d\omega \\ = -\frac{2\rho_0 \bar{L}_{\text{sup}}}{A_{\text{sup}}} H(\tau - \tau') \sum_{n=1}^2 \mathcal{A}_n e^{-\gamma_n(\tau-\tau')} \\ \times \cos[\omega_n(\tau - \tau') - \phi_n],\end{aligned}\quad (55)$$

where

$$\mathcal{A}_n = \left(\frac{|s_1||s_2|}{|r_1|} \right)^2 \sqrt{\mathcal{R}_n^2 + \mathcal{I}_n^2} \quad \text{and} \quad \phi_n = \arctan(\mathcal{I}_n/\mathcal{R}_n). \quad (56)$$

In the lossless case this reduces to,

$$\begin{aligned}\int_{-\infty}^{\infty} \hat{Z}_{\text{sup}}(\omega) e^{-i\omega(\tau-\tau')} d\omega = -\frac{2\rho_0 \bar{L}_{\text{sup}}}{A_{\text{sup}}} \left(\frac{\omega_1 \omega_2}{\psi_1} \right)^2 H(\tau - \tau') \\ \times \frac{1}{(\omega_2^2 - \omega_1^2)} \{ (\psi_1^2 - \omega_1^2) \cos \omega_1(\tau - \tau') \\ + (\omega_2^2 - \psi_1^2) \cos \omega_2(\tau - \tau') \}.\end{aligned}\quad (57)$$

Amplitudes $\mathcal{B}_n$ and phases Φ_n can be defined for the sub-glottal tract in an analogous way to those defined in the supra-glottal tract by Eq. (56). In this low frequency approximation the Fant equation [Eq. (47)] is written,

$$\begin{aligned}\bar{\ell}(\tau) \frac{\partial Q(\tau)}{\partial \tau} + \frac{A_{\text{sup}}}{2A_g^2(\tau)} \frac{Q^2(\tau)}{\sigma^2(\tau)} - A_{\text{sub}} \int_{-\infty}^{\infty} H(\tau - \tau') Q(\tau') \\ \times \left\{ \frac{2\bar{L}_{\text{sup}}}{A_{\text{sup}}} \sum_{n=1}^2 \mathcal{A}_n e^{-\gamma_n(\tau-\tau')} \cos[\omega_n(\tau - \tau') - \phi_n] \right. \\ \left. + \frac{2\bar{L}_{\text{sub}}}{A_{\text{sub}}} \sum_{n=1}^2 \mathcal{B}_n e^{-\Gamma_n(\tau-\tau')} \cos[\Omega_n(\tau - \tau') - \Phi_n] \right\} d\tau' \\ = A_{\text{sub}} \frac{p_{\text{sub}}}{\rho_0}\end{aligned}\quad (58)$$

for $\tau > 0$.

III. NUMERICAL SIMULATIONS

A. Definitions

The Fant equation in the low frequency approximation of Eq. (58) is solved numerically, following the procedure introduced in [Howe and McGowan \(2011b\)](#). Define,

$$\begin{aligned}K_{2n-1}(\tau) &= \int_{-\infty}^{\tau} Q(\tau') e^{-\gamma_n(\tau-\tau')} \cos[\omega_n(\tau - \tau') - \phi_n] d\tau', \\ K_{2n}(\tau) &= \int_{-\infty}^{\tau} Q(\tau') e^{-\gamma_n(\tau-\tau')} \sin[\omega_n(\tau - \tau') - \phi_n] d\tau'\end{aligned}\quad (59)$$

and

$$\begin{aligned}J_{2n-1}(\tau) &= \int_{-\infty}^{\tau} Q(\tau') e^{-\Gamma_n(\tau-\tau')} \cos[\Omega_n(\tau - \tau') - \Phi_n] d\tau', \\ J_{2n}(\tau) &= \int_{-\infty}^{\tau} Q(\tau') e^{-\Gamma_n(\tau-\tau')} \sin[\Omega_n(\tau - \tau') - \Phi_n] d\tau'\end{aligned}\quad (60)$$

for $n = 1, 2$. With these definitions, the integro-differential equation is replaced by a system of first-order differential equations,

$$\begin{aligned}\frac{dQ}{d\tau} &= \left[A_{\text{sub}} \frac{p_{\text{sub}}}{\rho_0} - \frac{A_{\text{sup}}}{2A_g^2(\tau)} \frac{Q^2(\tau)}{\sigma^2(\tau)} \right. \\ &\quad \left. - A_{\text{sub}} \sum_{n=1}^2 \left(\frac{2\bar{L}_{\text{sup}}}{A_{\text{sup}}} \mathcal{A}_n K_{2n-1}(\tau) \right. \right. \\ &\quad \left. \left. + \frac{2\bar{L}_{\text{sub}}}{A_{\text{sub}}} \mathcal{B}_n J_{2n-1}(\tau) \right) \right] / \bar{\ell}(\tau), \\ \frac{dK_{2n-1}}{d\tau} &= Q - \omega_n K_{2n} - \gamma_n K_{2n-1}, \\ \frac{dK_{2n}}{d\tau} &= \omega_n K_{2n-1} - \gamma_n K_{2n}, \\ \frac{dJ_{2n-1}}{d\tau} &= Q - \Omega_n J_{2n} - \Gamma_n J_{2n-1}, \\ \frac{dJ_{2n}}{d\tau} &= \Omega_n J_{2n-1} - \Gamma_n J_{2n},\end{aligned}\quad (61)$$

with initial conditions $Q = K_{2n-1} = K_{2n} = J_{2n-1} = J_{2n} = 0$ for $\tau = 0$ and $n = 1, 2$.

The glottis is assumed to be of length L in the x_1 direction, to have a rectangular cross-sections orthogonal to the x_1 direction of constant spanwise dimension d (into the paper in Fig. 1). The width, $w = w(x_1, \tau)$, (the vertical direction in Fig. 1) is time dependent so that the minimum glottal area $A_g(\tau)$ varies in time,

$$\frac{A_g(\tau)}{A_{\text{sub}}} = a_0 + a_1[1 - \cos(\omega_0\tau)], \quad (62)$$

$f_0 = \omega_0/2\pi$ is the voice fundamental frequency. The potential $Y(y, \tau)$ within the glottis and Rayleigh length $\bar{\ell}(\tau)$ introduced in Eq. (10) can be approximated as,

$$Y(y, \tau) \approx \frac{A_{\text{sub}} y_1}{A_g(\tau)} - \frac{\bar{\ell}(\tau)}{2}, \quad -L/2 < y_1 < L/2, \\ \text{where } \bar{\ell}(\tau) \approx \frac{A_{\text{sub}} L}{A_g(\tau)}. \quad (63)$$

B. Numerical results

The numerical simulations were performed with the following fixed parameter values,

$$A_{\text{sup}} = A_{\text{sub}} = 2 \times 10^{-4} \text{ m}^2, \quad \bar{L}_{\text{sup}} = 0.17 \text{ m}, \\ \bar{L}_{\text{sub}} = 0.2 \text{ m}, \quad L = 0.005 \text{ m}, \\ c_0 = 345 \text{ m/s}, \quad \rho_0 = 1.2 \text{ kg/m}^3, \quad p_{\text{sub}} = 1600 \text{ Pa}, \\ a_0 = 0.001, \quad a_1 = 0.05.$$

The sub-glottal properties are set according to [Ishizaka et al. \(1976\)](#). The first three sub-glottal resonances, or formants, are at 640 Hz, 1400 Hz, and 2100 Hz, with approximate bandwidths of 256 Hz, 156 Hz, and 175 Hz, respectively.² The zeros were assumed to be midway between the surrounding poles, with bandwidths derived from the same plot of Q-factors as used to derive the bandwidths of the resonances (Fig. 4 in [Ishizaka et al., 1976](#)). (Bandwidth is the width of the resonance peak 3 dB down from its power maximum.) This gives the first two zeros as 1020 Hz and 1750 Hz with bandwidths of 206 Hz and 165 Hz, respectively. The first

simulation is for the case where there is no supra-glottal resonance, so that $\hat{Z}_{\text{sup}} = \rho_0 c_0 / A_{\text{sup}}$ for a semi-infinite tube. The results of these simulations are shown in Fig. 2 for $f_0 = 100, 280, 400$, and 800 Hz in terms of normalized glottal volume velocity, $Q_{\text{norm}}(t) = (Q(t)\rho_0 c_0) / (A_{\text{sub}} p_{\text{sub}})$. The horizontal axis is time normalized by f_0 , which gives the number of glottal cycles. The dashed lines are the volume velocity that results when the sub-glottal tract is replaced by a semi-infinite tube, with $\hat{Z}_{\text{sub}} = \rho_0 c_0 / A_{\text{sub}}$.

The amplitude of the glottal volume velocity is slightly enhanced by the presence of the sub-glottal resonances for all the fundamental frequencies f_0 . The enhancement is largest for the lowest fundamental frequency, $f_0 = 100$ Hz. There is evidence of rightward skewing of the glottal pulse for $f_0 = 280$ Hz, which increases at $f_0 = 400$ Hz. As predicted by [Rothenberg \(1981\)](#), and confirmed by [Ananthapadmanabha and Fant \(1982\)](#) and [Titze \(2008\)](#), as the sub-glottal input impedance becomes more “inductive” (i.e., with a larger positive imaginary part) for the higher fundamental frequency rightward, skewing of the glottal pulse becomes greater. Note that the relatively large bandwidth of the first sub-glottal resonance means that this effect is seen well-below its first resonance frequency. At $f_0 = 800$ Hz the effect of being above the first resonance, yet still well below the first zero and higher poles of the input impedance, is seen to skew the volume velocity pulse to the left. This is what occurs with a “capacitive” load on the glottis ([Titze, 2008](#)).

The effects of supra-glottal resonances are now considered. Possible resonance frequencies, formants, appropriate for an adult male American English production of /i/ are taken to be 280 Hz, 2250 Hz, and 2750 Hz ([Olive et al., 1993](#)). Bandwidths are estimated by interpolation of the data presented in [Stevens \(1998\)](#) from [Fant \(1962\)](#) and [Fujimura and Lindqvist \(1970\)](#) to be 65 Hz, 98 Hz, and 113 Hz, respectively. The first two zeros of the supra-glottal input impedance are set to be located at the mean frequency of the surrounding resonance at 1265 Hz and 2500 Hz with bandwidths 69 Hz and 105 Hz, respectively. The resulting $Q_{\text{norm}}(t)$ is shown in Fig. 3 for $f_0 = 100, 260, 275$, and 290 Hz. The dashed curve for comparison is the condition (base condition) discussed above with the sub-glottal resonances alone. The

FIG. 2. Normalized glottal volume velocity, $Q_{\text{norm}} = (Q(t)\rho_0 c_0) / (p_{\text{sub}} A_{\text{sub}})$, versus normalized time, $f_0 t$, at four different fundamental frequencies. The supra-glottal impedance is that for a semi-infinite tube, $(\rho_0 c_0) / A_{\text{sup}}$. The dark curves have sub-glottal input impedance specified in the text according to [Ishizaka et al. \(1976\)](#). Sub-glottal formant frequencies, with bandwidths in brackets are 640 Hz (256 Hz), 1400 Hz (156 Hz), and 2100 Hz (175 Hz). The dashed curve for comparison is when the sub-glottal tract is a semi-infinite tube with input impedance $(\rho_0 c_0) / A_{\text{sub}}$. (a) $f_0 = 100$ Hz, (b) $f_0 = 280$ Hz, (c) $f_0 = 400$ Hz, and (d) $f_0 = 800$ Hz.

FIG. 3. Normalized glottal volume velocity, $Q_{norm} = (Q(t)\rho_0 c_0)/(p_{sub} A_{sub})$, versus normalized time, $f_0 t$ at four different fundamental frequencies. Sub-glottal resonances are as specified for the dark curve in Fig. 2. The dark curves are for supra-glottal formants appropriate for /i/, which are, with bandwidths in brackets, 280 Hz (65 Hz), 2250 Hz (98 Hz), and 2750 Hz (113 Hz). The dashed curves are when supra-glottal impedance is $(\rho_0 c_0)/A_{sup}$ for a semi-infinite tube. (a) $f_0 = 100$ Hz, (b) $f_0 = 260$ Hz, (c) $f_0 = 275$ Hz, and (d) $f_0 = 290$ Hz.

amplitude of the volume velocity pulse is enhanced substantially above the base condition for $f_0 = 100$ Hz, but there is no change in the skewing of the pulse. The enhancement of the glottal pulsing is diminished when f_0 changes from 100 Hz to frequencies in the neighborhood of the first supra-glottal formant. Note also that steady-state takes two or three pulses to become established with the fundamental frequency close to the first formant frequency. There is a slight increase in rightward skewing of the glottal volume velocity pulse at $f_0 = 260$ Hz and 275 Hz compared to 100 Hz. Also there is leftward skewing at $f_0 = 290$ Hz. These show that the supra-glottal input impedances supply “inductive” and “capacitive” perturbations to the sub-glottal input impedances for this low first supra-glottal first formant.

Howe and McGowan (2011b) find a strong effect of source-tract interaction when the fundamental frequency is very close to a Helmholtz resonance frequency (see Fig. 5 in Howe and McGowan, 2011b). This interaction has the effect of reducing the peak flow to the degree that there are two local maxima and a local minimum in a single area function pulse. Such behavior could not be induced here for /i/, but there is evidence of a change in curvature near the peak of the glottal pulse for fundamental frequencies close to the

formant frequency. We suggest that the propagation losses in the present system are larger than in Howe and McGowan (2011b), so that this type of interaction does not appear for low formant frequencies.

While the first supra-glottal formant frequency of /i/ is low, and, thus, strong source-tract interaction might be expected with this vowel, its second supra-glottal formant frequency is high. What happens when both the first and second formant frequencies are relatively low, as for American English /r/? The first three formants for this phoneme in word initial position can be as low as 250 Hz, 700 Hz, and 1400 Hz for an adult male American English speaker according to Olive *et al.* (1993). The bandwidths are 69 Hz, 54 Hz, and 73 Hz, respectively, as computed from interpolated data. The interspersed zeros of the supra-glottal input impedance are 475 Hz and 1050 Hz with bandwidths of 51 Hz and 62 Hz, respectively. Figure 4 shows the results for $f_0 = 100, 240, 245$, and 260 Hz. The results are not qualitatively different from the results for the vowel /i/ shown in Figs. 3, except that the pulses may have higher amplitude for /r/. The reason for the similarity between the results appears to be the existence of the zero of the input impedance between the formants, which tends to cancel the effect of the second formant

FIG. 4. Normalized glottal volume velocity, $Q_{norm} = (Q(t)\rho_0 c_0)/(p_{sub} A_{sub})$, versus normalized time, $f_0 t$ at four different fundamental frequencies. Sub-glottal resonances are as specified for the dark curve in Fig. 2. The dark curves are for supra-glottal formants appropriate for /r/, which are, with bandwidths in brackets, 250 Hz (69 Hz), 700 Hz (54 Hz), and 1400 Hz (73 Hz). The dashed curves are when supra-glottal impedance is $(\rho_0 c_0)/A_{sup}$ for a semi-infinite tube. (a) $f_0 = 100$ Hz, (b) $f_0 = 240$ Hz, (c) $f_0 = 245$ Hz, and (d) $f_0 = 260$ Hz.

FIG. 5. Peak amplitudes for /i/ versus fundamental frequency, f_0 at a distance $r = 400$ cm from the lips. (a) Normalized far field pressure, $10\log_{10}(p_{f_{\max}}/p_{\text{sub}})$ versus f_0 , and (b) normalized volume velocity $10\log_{10}(Q_{\max}\rho_0 c_0/p_{\text{sub}}A_{\text{sub}})$ versus f_0 .

moving closer to the first formant. This cancellation can be easily be seen in the lossless case with reference to Eq. (57). The denominator of the interaction term increases when the first and second formants move closer together, but the magnitudes of the coefficients of the two terms in brackets decrease, because the zero frequency is between the two formant frequencies.

The cases of /i/ and /r/ would apparently confirm the observation of Ananthapadmanabha and Fant (1982) that only one formant from each the sub-glottal tract and the supra-glottal tract is necessary to capture the main effects of source-tract interaction, at least for conversational fundamental frequencies. However, we have assumed that the zeros of the input impedances are placed equidistant between neighboring poles. This is not necessarily the case, and in fact important phenomena can be missed if this assumption is taken to always be the case. Titze and Story (1997) have shown that a narrow epilarynx tube has the effect of attracting the second, third, and fourth formants toward one another. Also, an examination of their Fig. 4 indicates that the zeros of the input impedance are adjusted in such away that the magnitude of the input impedance is greatly enhanced in the frequency region of these formants. This would have the effect of making the glottal volume flow's interaction with the higher supra-laryngeal formants much more important than shown here. In particular one could expect ripples to appear in the glottal volume velocity waveform when appropriate modifications are made to input impedance (Titze, 2008).

The far field acoustic pressure can be considered within the present context. The transfer function between the volume velocity at the glottis, $\hat{Q}_\ell(\omega)$, and the volume velocity at the lips, $\hat{Q}_\ell(\omega)$, is (Fant, 1960)

$$\frac{\hat{Q}_\ell(\omega)}{\hat{Q}(\omega)} \approx \prod_{n=1}^3 \left(\frac{1}{(1 - \omega/s_n)(1 + \omega/s_n^*)} \right), \quad (64)$$

where higher-order poles and radiation through the vocal tract walls are neglected. In the far field at a distance r from the lips, the acoustic pressure fluctuations, p_f , is given by

$$p_f(r, t) = \frac{\rho_0}{4\pi r} \frac{\partial Q_\ell(t - r/c_0)}{\partial t}, \quad (65)$$

where scattering by the body has been neglected and the area surrounded by the lips is considered a point source for the frequencies under consideration (< 1 kHz). Far field pressure can be computed as a function of time from the solution of the Fant equation, $Q(t)$, by employing Eqs. (64) and (65).

Figures 5 and 6 show normalized normalized peak far field pressure, $p_{f_{\max}}/p_{\text{sub}}$ at a distance $r = 4$ m from the lips and peak glottal volume velocity $(Q_{\max}\rho_0 c_0)/(p_{\text{sub}}A_{\text{sub}})$ as functions of fundamental frequency, f_0 , for the vowel /i/ discussed in Fig. 3 and the /r/ discussed in Fig. 4. Both normalized peak quantities are plotted on decibel scales. In both cases there are relatively small decreases in peak glottal volume velocity at either sub-glottal or supra-glottal formants. In particular, two local minima in the peak glottal volume velocity can be noted in the frequency neighborhood of the first sub-glottal formant at 640 Hz and the second supra-glottal formant nearby at 700 Hz for /r/. On the other hand the filtering properties of the vocal tract makes peak far field pressures increase at the supra-glottal formants, as seen by Titze (2008) and in Howe and McGowan (2011b). The normalized peak far field pressure can be considered a measure of voice efficiency (Howe and McGowan, 2011b).

A final example is provided by the low back vowel /a/ with its relatively high frequency first formant and low second formant frequency. The formant frequencies with their bandwidths are, 750 Hz (bandwidth = 54 Hz), 1100 Hz (bandwidth = 64 Hz), and 2600 Hz (bandwidth = 108 Hz) (Olive et al., 1993). The first two zeros are 925 Hz (bandwidth = 59 Hz) and 1850 (bandwidth = 86 Hz). The glottal volume velocity waveforms are shown in Fig. 7 for $f_0 = 100, 650, 745$, and 800 Hz. While the final three fundamental frequencies are very high, they show a great amount of source tract interaction in the form of two local maxima per glottal pulse for at least a 150 Hz range in fundamental frequency. Part of the reason for this could be the proximity of the first supra-glottal formant (750 Hz) to the first sub-glottal formant (640 Hz). Despite the complexity of the

FIG. 6. Peak amplitudes for /t/ versus fundamental frequency, f_0 at a distance $r = 400$ cm from the lips. (a) Normalized far field pressure, $10\log_{10}(p_{fmax}/p_{sub})$ versus f_0 , and (b) normalized volume velocity $10\log_{10}(Q_{max}\rho_0 c_0/p_{sub}A_{sub})$ versus f_0 .

waveforms there is evidence of extra rightward pulse skewing for $f_0 = 650$ Hz and some leftward skewing from the base condition for $f_0 = 745$ and 800 Hz, after a couple of glottal pulses. $f_0 = 745$ Hz is apparently above the actual first supra-glottal formant frequency when losses are taken into account. That is damping decreases the actual formant frequency from the nominal specification of 750 Hz to one less than 745 Hz (see footnote 2). The complexity of the waveform seems to counteract any maxima in the far field peak pressure as a function of f_0 that could appear around the first supra-glottal formant frequency (Fig. 8).

IV. DISCUSSION AND CONCLUSION

A Fant equation has been derived for the vocal tract when the specific form of the Green's function remains unknown. This is made possible by the mathematical observation in [Howe and McGowan \(2011b\)](#) that the Fant equation is the adjoint of an equation that results from satisfying matching conditions between segments of Green's functions in the glottal region on one hand and Green's functions

segments in the sub-glottal and supra-glottal regions on the other hand. The derivation requires that the glottal region be acoustically compact and that the governing equations of motion for the air be ones appropriate for low Mach number, homentropic flow, which is the case for speech production. The only other assumption is that the fluid motion be essentially potential and without cross-modes in the vocal tract regions immediately outside the glottal region. The resulting Fant equation, Eq. (47), is written in terms of measurable quantities, namely the sub-glottal and supra-glottal input impedances.

One of the assumptions in the derivation of the Fant equation is that the fluid flow just upstream and downstream of the glottal region is governed by the one-dimensional wave equation, and there are no acoustic cross modes present there. This is in order that the input impedances be well-defined, but it does not preclude the existence of cross-modal propagation beyond the immediate neighborhood of the glottis. It is when specific propagation models are used to find the functional form of the input impedances is it convenient to impose one-dimensional propagation throughout

FIG. 7. Normalized glottal volume velocity, $Q_{norm} = (Q(t)\rho_0 c_0)/(p_{sub}A_{sub})$, versus normalized time, $f_0 t$ at four different fundamental frequencies. Sub-glottal resonances are as specified for the dark curve in Fig. 1. The dark curves are for supra-glottal formants appropriate for /a/, which are, with bandwidths in brackets, 750 Hz (54 Hz), 1100 Hz (64 Hz), and 2600 Hz (108 Hz). The dashed curves are when supra-glottal impedance is $(\rho_0 c_0)/A_{sup}$ for a semi-infinite tube. (a) $f_0 = 100$ Hz, (b) $f_0 = 650$ Hz, (c) $f_0 = 745$ Hz, and (d) $f_0 = 800$ Hz.

FIG. 8. Peak amplitudes for /a/ versus fundamental frequency, f_0 at a distance $r = 400$ cm from the lips. (a) Normalized far field pressure, $10\log_{10}(p_{fmax}/p_{sub})$ versus f_0 , and (b) normalized volume velocity $10\log_{10}(Q_{max}\rho c_0/p_{sub}A_{sub})$ versus f_0 .

the vocal tract. This is a common approximation in speech science, where there has been no, or very little, attention to the possibility of cross-mode propagation in the vocal tract.

Assuming that low frequency propagation is governed by a Sturm-Liouville problem or by a concatenated cylindrical tube model with one-dimensional propagation, the functional forms of the sub- and supra-glottal input impedances can be derived. These functions of complex frequency are completely specified by their simple poles and zeros. Simulations based on the first two poles, or formants, and estimated zeros show that the lossy sub-glottal system affects the skewing of the glottal pulse over a wide range of frequencies. Adding the formants of the supra-glottal system provide perturbations to the glottal pulse's shape caused by the sub-glottal system. If zeros of the input-impedance are presumed to fall equidistant in frequency between neighboring poles it appears that the first sub-glottal and supra-glottal formants dominate the behavior of source-tract interaction at conversational fundamental frequencies. However, this assumption does not appear to be realistic when a relatively narrow epilarynx tube is taken into account (Titze and Story, 1997). This tube causes the higher formants to interact with the glottal volume velocity much more strongly and to cause ripples in the pulse. It would be best to use a concatenated tube model, as described in the second part of the Appendix, to derive the input impedance in most cases. These observations are made for source-tract interactions for which the glottal area is specified (level I), and, thus do not represent the full source-tract problem in speech.

The more practical and realistic source-tract interaction results will obtain when the Fant equation, Eq. (47), derived in this work is applied to the full problem that includes the fluid-structure interaction between the air in the glottis and the vocal folds. This is known as a level II model (Titze, 2008). The past few years has seen theoretical, numerical simulation, and experimental work in the area of level II source-tract interaction (e.g., Švec *et al.*, 2008; Titze, 2008; Titze *et al.*, 2008; Titze and Morely Worely, 2009; Tokuda *et al.*, 2007; Zañartu *et al.*, 2011). These studies show that source-tract interaction is an important factor when considering voice

register changes in response to changes in laryngeal settings. These changes are not just gradual changes in volume velocity pulse skewing with the level I interactions examined here, but bifurcations with discontinuous jumps in fundamental frequency and vocal fold vibration mode.

To proceed to the full level II situation, a model of fluid-structure interaction in the glottal region would be included, such as the body-cover model of Story and Titze (1995). The Fant equation, Eq. (47), would be solved simultaneously for the volume flow $Q(t)$ with the fluid-structure interaction model, because the Fant equation requires knowledge of the glottal area $A_g(t)$ and the fluid-structure interaction model requires knowledge of $Q(t)$. (Note that the time domain pressures upstream and downstream of the glottis can be obtained by convolving the volume velocity with the inverse Fourier transforms of the sub-glottal and supra-glottal impedances, respectively.)

The present work has concentrated on obtaining a Fant equation that can be used for both level I and level II source-tract interaction problems. Before applying this equation to level II problems, more research directed towards understanding the details of glottal fluid flow and the resulting fluid-structure interaction needs to be done. So far, recent research has indicated a role for suction, or Coanda, forces created with flow that changes direction (Howe and McGowan, 2009; McGowan and Howe, 2010). The behavior of the separation point during the glottal cycle is also an important consideration (e.g., Pelorson *et al.*, 1994; Howe and McGowan, 2009). An effect that brings asymmetry to the vocal fold oscillation problem involves both Coanda forces and flow separation, the Coanda effect, also needs to be included eventually (Erath *et al.*, 2011). Progress in that research will permit a consideration of full source-tract interaction in the future.

ACKNOWLEDGMENT

This work was supported by a subaward of Grant No. 1R01 DC009229 from the National Institute on Deafness and other Communication Disorders to the University of

California, Los Angeles. We thank two anonymous reviewers for helping to improve this paper.

APPENDIX

The functional forms of $\hat{Z}_{\text{sup}}(\omega)$ and $\hat{Z}_{\text{sub}}(\omega)$ are examined in models of lossless plane wave propagation. One model is provided by Sturm-Liouville problems, and the other is provided by concatenated cylindrical tubes. The characterizations of the functional form of the input impedance will be derived for the supra-glottal vocal tract, as the derivations for the sub-glottal tract are completely analogous.

A. Sturm-Liouville problems

First the case when the plane wave propagation in the sub- or supra-glottal vocal tract can be described by a lossless Sturm-Liouville equation with homogeneous boundary conditions (a Sturm-Liouville problem) is considered. An example of such a lossless equation is the Webster horn equation, which is valid when cross-sectional areas do not change with axial position y_1 too rapidly in relation to the ratio of the speed of sound, c_0 , to characteristic frequency (Pierce, 1989). However, this equation does not exhaust the possible ways that the low-frequency propagation problem could be described by a Sturm-Liouville equation. It is assumed that, in the absence of losses and sources in the sub- and supra-glottal tracts, that the Fourier transform of Eq. (2) can be written in the form

$$\frac{d}{dx_1} \left(\mathcal{K}(x_1) \frac{d\hat{B}(x_1, \omega)}{dx_1} \right) - (\mathcal{H}(x_1) - \lambda) \hat{B}(x_1, \omega) = 0, \quad (\text{A1})$$

where $\lambda = \omega^2/c_0^2$, $\mathcal{K}(x_1) > 0$ and $\mathcal{H}(x_1)$ are real-valued functions (Ince, 1956). Coupling Eq. (A1) with Eq. (48) constitutes a Sturm-Liouville problem. The minimal conditions that permit a second-order differential equation to be written in this self-adjoint operator form in the sub- and supra-glottal regions, such as the continuity of $\frac{\mathcal{K}(x_1)}{dx_1}$ and $\mathcal{H}(x_1)$ have been presumed (Ince, 1956). Because of the strictly monotonic relation between λ and ω for $\omega > 0$ a function of one of these variables can be considered a function of the other.

We draw on the classical theory of Sturm-Liouville problems and their Green's functions to characterize the functional form of $\hat{Z}_{\text{sup}}(\omega)$ (Courant and Hilbert, 1953; Ince, 1956; Morse and Feshbach, 1953). The Green's function, which is self-adjoint, $\mathcal{G}(x_1, y_1, \lambda)$ satisfies

$$\begin{aligned} & \left[\frac{d}{dx_1} \left(\mathcal{K}(x_1) \frac{d}{dx_1} \right) - (\mathcal{H}(x_1) - \lambda) \right] \mathcal{G}(x_1, y_1, \lambda) \\ & = -\delta(x_1 - y_1), \\ & \frac{d\mathcal{G}(+0, y_1, \lambda)}{dx_1} = \frac{d\mathcal{G}(x_1, +0, \lambda)}{dy_1} = 0, \quad \text{and} \\ & \mathcal{G}(\bar{L}_{\text{sup}}, y_1, \lambda) = \mathcal{G}(x_1, \bar{L}_{\text{sup}}, \lambda) = 0. \end{aligned} \quad (\text{A2})$$

It can be shown that $\mathcal{G}(x_1, y_1, \lambda)$ with $y_1 > x_1 \geq +0$ can be written (Courant and Hilbert, 1953),

$$\begin{aligned} \mathcal{G}(x, y, \lambda) &= -H(y_1 - x_1) \frac{\hat{u}(x_1, \lambda) \hat{v}(y_1, \lambda)}{W(\lambda)}, \quad \text{with} \\ W(\lambda) &= \mathcal{K}(x_1) \left[\hat{u}(x_1, \lambda) \frac{\partial \hat{v}(x_1, \lambda)}{\partial x_1} - \frac{\partial \hat{u}(x_1, \lambda)}{\partial x_1} \hat{v}(x_1, \lambda) \right]. \end{aligned} \quad (\text{A3})$$

Both $\hat{u}$ and $\hat{v}$ are solutions to Eq. (A1), and each satisfies one of the boundary conditions in Eq. (48): $\frac{\partial \hat{u}(+0, \lambda)}{\partial x_1} = \hat{v}(\bar{L}_{\text{sup}}, \lambda) = 0$. The Wronskian, $W(\lambda)$, is not a function of the spatial variable x_1 , and it will be zero at certain values of λ , which is when functions $\hat{u}$ and $\hat{v}$ are linearly dependent and satisfy both boundary conditions. In this case, the particular value of λ is an eigenvalue and these functions are the same eigenfunction (to within a constant multiple).

If x_1 is set to $+0$ in Eq. (A3), then in the limit as $y_1 \rightarrow +0$,

$$\lim_{y_1 \rightarrow +0} \mathcal{G}(+0, y_1, \lambda) = \left(\frac{-1}{\mathcal{K}(+0)} \right) \frac{\hat{v}(+0, \lambda)}{\frac{\partial \hat{v}(+0, \lambda)}{\partial x_1}}. \quad (\text{A4})$$

With Eqs. (A2) and (A4), the supra-glottal input impedance defined completely analogous with Eq. (29) becomes

$$\hat{Z}_{\text{sup}}(\omega) i \omega \rho_0 \lim_{y_1 \rightarrow +0} \mathcal{G}(+0, y_1, \lambda) = \frac{-i \omega \rho_0}{\mathcal{K}(+0)} \frac{\hat{v}(+0, \lambda)}{\frac{\partial \hat{v}(+0, \lambda)}{\partial x_1}}. \quad (\text{A5})$$

The zeros of $\hat{Z}_{\text{sup}}(\omega)$ correspond to the zeros of $\hat{v}(+0, \lambda)$ and the poles of $\hat{Z}_{\text{sup}}(\omega)$ to the zeros of $\frac{\partial \hat{v}(+0, \lambda)}{\partial x_1}$.

The facts known about the $\hat{v}(x_1, \lambda)$ are usually stated in terms of $\hat{v}(x_1, \lambda)$ as a function of x_1 at λ eigenvalues. First, for Sturm-Liouville problems there is an infinite sequence of positive eigenvalues $\lambda_0, \lambda_1, \dots$ (without a finite limit point) *i.e.* those values λ_n for which $\frac{\partial \hat{v}(+0, \lambda_n)}{\partial x_1} = 0$, so that the $\hat{v}(x_1, \lambda_n)$ are eigenfunctions. For each eigenvalue, λ_n all the zeros of $\hat{v}(x_1, \lambda_n)$ and $\frac{\partial \hat{v}(+0, \lambda_n)}{\partial x_1}$ are simple and real. Further the zeros of $\hat{v}(x_1, \lambda_n)$ interleave those of $\frac{\partial \hat{v}(+0, \lambda_n)}{\partial x_1}$, and there are exactly n zeros of $\hat{v}(x_1, \lambda_n)$ for $0 < x_1 < \bar{L}_{\text{sup}}$.

These properties are also true of $\hat{v}(+0, \lambda)$ and $\frac{\partial \hat{v}(+0, \lambda)}{\partial x_1}$, considered as functions of λ . The fact that this is true follows from the following relationship between x_1 and λ for Sturm-Liouville problems. Let $\kappa(\lambda') = \sqrt{\frac{\lambda' - \mathcal{H}(+0)}{\mathcal{K}(+0)}}$, with $\lambda' > 0$ and $\lambda' - \mathcal{H}(+0) > 0$. Let $\theta = \kappa(x_1 - \bar{L}_{\text{sup}})$, so that $\theta = \theta(x_1, \kappa)$. It can be shown that solutions to $\frac{d^2 V(\theta)}{d\theta^2} + V(\theta) = 0$ and its first and second order derivatives can be made arbitrarily close to $\hat{v}(x_1, \lambda)$ and its corresponding derivatives for x_1 sufficiently close to $+0$ and λ sufficiently close to λ' . $d\theta = \frac{\partial \theta(+0, \lambda)}{\partial x_1} dx_1 + \frac{\partial \theta(+0, \lambda)}{\partial \lambda} d\lambda$, $d\theta = 0$ requires that

$$\frac{\partial \lambda(\theta, x_1)}{\partial x_1} \bigg|_{x_1=+0} = - \frac{\frac{\partial \theta(+0, \lambda)}{\partial x_1}}{\frac{\partial \theta(+0, \lambda)}{\partial \lambda}} > 0, \quad (\text{A6})$$

with the inequality following from the definition of θ . Therefore, λ is a strictly monotonic function of x_1 for constant phase θ . So the pattern of maxima, minima, and zeros of

$\hat{v}(x_1, \lambda)$ as a function of x_1 for $0 < x_1 < \bar{L}_{\text{sup}}$ discovered in the classical theory of Sturm-Liouville problems is the result of the same pattern found passing from negative x_1 to positive x_1 through $x_1 = +0$ as λ increases: Namely, interleaving simple, real zeros of $\hat{v}(+0, \lambda)$ and $\frac{\partial \hat{v}(+0, \lambda)}{\partial x_1}$, as λ increases from 0 to ∞ . Therefore, $\hat{Z}_{\text{sup}}(\omega)$ possesses an infinity of interleaving zeros and poles as ω increase from 0 to ∞ .

B. Concatenated cylindrical tubes

The low frequency approximation to the supra-glottal input impedance in the case that the supra-glottal vocal tract is approximated by N sections of cylindrical tubes of equal length, l , so that $\bar{L}_{\text{sup}} = Nl$ is now considered. The tubes are numbered with the tube that ends at the lips the first tube. The n th tube possesses cross-sectional area A_n , and let $Z_n = \frac{\rho_0 c_0}{A_n}$. Continuity of pressure and volume velocity at the junctions between the tubes is assumed. With these continuity assumptions, the input impedance looking toward the lips from the end of the n th tube away from the lips is (Lighthill, 1978)

$$Z_n^{\text{in}} = Z_n \frac{Z_{n-1}^{\text{in}} + iZ_n \tan(\theta)}{Z_n + iZ_{n-1}^{\text{in}} \tan(\theta)}, \quad (\text{A7})$$

where $\theta = \omega l / c_0$ and the impedance at the lips can be denoted Z_0^{in} , with $Z_0^{\text{in}} = 0$. Thus,

$$Z_1^{\text{in}} = iZ_1 \tan \theta. \quad (\text{A8})$$

Here it is claimed that for all $n, N \geq n \geq 1$,

$$Z_n^{\text{in}} = iZ_n F_n(\theta), \quad (\text{A9})$$

where the poles and zeros of real valued $F_n(\theta)$ are real and simple. Also, $\frac{dF_n(\theta)}{d\theta} > 0$, except at poles of $F_n(\theta)$, which is equivalent to requiring that neighboring poles be separated by a zero, and that neighboring zeros be separated by poles, when the poles and zeros are real and simple. Further, $F_n(m\pi) = 0$ for all integers m .

That such $F_n(\theta)$ exist for all $n, N \geq n \geq 1$ can be shown by induction. $F_1(\theta) = \tan(\theta)$ satisfies Eq. (A9) with all the required properties following that equation. Suppose that $F_{n-1}(\theta)$ exists that satisfies Eq. (A9) and all the properties following that equation. From the induction hypothesis and Eq. (A7)

$$Z_n^{\text{in}} = iZ_n \frac{Z_{n-1} F_{n-1}(\theta) + Z_n \tan(\theta)}{Z_n - Z_{n-1} F_{n-1}(\theta) \tan(\theta)}. \quad (\text{A10})$$

If $F_n(\theta)$ is identified with the real-valued fraction in Eq. (A10), then Eq. (A9) is satisfied and $F_n(m\pi) = 0$ for all integer m . The poles for $F_n(\theta)$ are the θ for which,

$$\cot(\theta) = \frac{Z_{n-1}}{Z_n} F_{n-1}(\theta), \quad (\text{A11})$$

and the zeros of $F_n(\theta)$ are the θ for which,

$$-\tan(\theta) = \frac{Z_{n-1}}{Z_n} F_{n-1}(\theta). \quad (\text{A12})$$

Both $\cot(\theta)$ and $-\tan(\theta)$ are strictly monotone decreasing functions of θ that do not intersect for any θ . Under the induction hypothesis, $F_{n-1}(\theta)$ is a strictly monotone increasing function. Thus the poles and zeros of $F_n(\theta)$ are real and simple. [Note that there is the possibility that either Eq. (A11) or (A12) could be satisfied for values of θ for which the functions attain either $\pm\infty$.] Calculus shows that $\frac{dF_n(\theta)}{d\theta} > 0$ for values of θ that are not poles of $F_n(\theta)$. Thus the countable infinity of poles and zeros of Z_n^{in} are real and simple, and they separate one another for all $n, N \geq n \geq 1$. In particular this is true for $\hat{Z}_{\text{sup}}(\omega) = Z_N^{\text{in}}$.

¹When wall distensibility is accounted for in realistic boundary conditions, as they would be in Eqs. (3) and (5), the speed of sound becomes a function of frequency $c(\omega)$. It can be assumed that the real part of $c(\omega)$, is a monotonically increasing function of ω for the frequency range of interest, and that it approaches c_0 as frequency increases (Lighthill, 1978; Davies *et al.*, 1993). Further, the increase of the real part of $c(\omega)$ is presumed to be such that the ordering of the real parts of the poles and zeros of the input impedance functions remains unchanged from the hard-wall case.

²The formant frequencies specified in this section will refer to the real part of the impedance poles. Measured formant frequencies will be perturbed to a value less than these nominal specifications because of losses.

- Ananthapadmanabha, T. V., and Fant, G. (1982). "Calculation of true glottal flow and its components," *Speech Commun.* **1**, 167–184.
- Courant, R., and Hilbert, D. (1953). *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York), Vol. 1, pp. 291–295, 351–358.
- Davies, P. O. A. L., McGowan, R. S., and Shadle, C. H. (1993). "Practical flow-duct acoustics applied to the vocal tract," in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. R. Titze (Singular Publishing Group, San Diego, CA), pp. 93–142.
- Erath, B. D., Peterson, S. D., Zañartu, M., Wodicka, G. R., and Plesniak, M. W. (2011). "A theoretical model of the pressure field arising from asymmetric intraglottal flows applied to a two-mass model of the vocal folds," *J. Acoust. Soc. Am.* **130**, 389–403.
- Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, Hague, Netherlands), pp. 42, 265–272.
- Fant, G. (1962). "Formant bandwidth data," *Speech Transmission Laboratory Quarterly Progress and Status Report 1* (Royal Institute of Technology, Stockholm), pp. 1–3.
- Fant, G. (1986). "Glottal flow: models and interaction," *J. Phonetics* **14**, 393–399.
- Flanagan, J. L., and Cherry, L. (1968). "Excitation of vocal-tract synthesizers," *J. Acoust. Soc. Am.* **45**, 764–769.
- Flanagan, J. L., and Landgraf, L. L. (1968). "Self-oscillating sources for vocal-tract synthesizers," *IEEE Trans. Audio Electroacoust.* **AU-16**, 57–64.
- Fujimura, O., and Lindqvist, J. (1970). "Sweep-tone measurements of vocal-tract characteristics," *J. Acoust. Soc. Am.* **49**, 541–558.
- Habib, R. H., Chalker, R. B., Suki, B., and Jackson, A. C. (1994). "Airway geometry and wall mechanical properties estimated from subglottal input impedances in humans," *J. Appl. Physiol.* **77**, 441–451.
- Howe, M. S. (1998). *Acoustics of Fluid-Structure Interactions* (Cambridge University Press, Cambridge, U.K.), pp. 1–560.
- Howe, M. S. (2002). *Theory of Vortex Sound* (Cambridge University Press, Cambridge, U.K.), pp. 1–216.
- Howe, M. S. (2008). "Rayleigh Lecture 2007: Flow-surface interaction noise," *J. Sound Vib.* **314**, 113–146.
- Howe, M. S., and McGowan, R. S. (2007). "Sound generated by aerodynamic sources near a deformable body, with application to voiced speech," *J. Fluid Mech.* **592**, 367–392.
- Howe, M. S., and McGowan, R. S. (2009). "Nonlinear flow-structure coupling in a mechanical model of the vocal folds and subglottal system," *J. Fluids Struct.* **25**, 1299–1317.
- Howe, M. S., and McGowan, R. S. (2011a). "On the generalized Fant equation," *J. Sound Vib.* **330**, 3123–3140.
- Howe, M. S., and McGowan, R. S. (2011b). "Production of sound by unsteady throttling of flow into a resonant cavity, with application to voiced speech," *J. Fluid Mech.* **672**, 428–450.

- Ince, E. L. (1956). *Ordinary Differential Equations* (Dover Publications, New York), pp. 210, 215, 223–235.
- Ishizaka, K., and Flanagan, J. L. (1972). “Synthesis of voiced sounds from a two-mass model of the vocal cords,” *Bell Syst. Tech. J.* **51**, 1233–1267.
- Ishizaka, K., Matsudaira, M., and Kaneko, T. (1976). “Input acoustic-impedance measurement of the subglottal system,” *J. Acoust. Soc. Am.* **60**, 190–197.
- Lighthill, J. (1978). *Waves in Fluids* (Cambridge University Press, Cambridge, U. K.), pp. 89–120.
- McGowan, R. S., and Howe, M. S. (2010). “Comments on single-mass models of vocal fold vibration,” *J. Acoust. Soc. Am.* **127**, EL215–EL222.
- Morse, P. M., and Feshbach, H. (1953). *Methods of Theoretical Physics* (McGraw-Hill, New York), pp. 719–729, 804–807, 869–874.
- Olive, J. P., Greenwood, A., and Coleman, J. (1993). *Acoustics of American English Speech: A Dynamic Approach* (Springer-Verlag, New York), pp. 104, 208.
- Pelorsen, X., Hirschberg, A., van Hassel, R. R., and Wijnands, A. P. J. (1994). “Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model,” *J. Acoust. Soc. Am.* **96**, 3416–3431.
- Pierce, A. D. (1989). *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, Woodbury, NY), pp. 360–361.
- Rothenberg, M. (1981). “Acoustic interaction between the glottal source and the vocal tract,” in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo), pp. 305–328.
- Stevens, K. N. (1998). *Acoustic Phonetics* (MIT Press, Cambridge, MA), pp. 258–260.
- Story, B. H., and Titze, I. R. (1995). “Voice simulation with a body-cover model of the vocal folds,” *J. Acoust. Soc. Am.* **97**, 1249–1260.
- Svec, J. G., Sundberg, J., and Hertegard, S. (2008). “Three registers in an untrained female singer analyzed by videokymography, strobolar-yngoscopy, and sound spectrography,” *J. Acoust. Soc. Am.* **123**, 347–353.
- Titze, I. R. (1984). “Parameterization of the glottal area, glottal flow, and vocal contact area,” *J. Acoust. Soc. Am.* **75**, 570–580.
- Titze, I. R. (2008). “Nonlinear source-filter coupling in phonation: Theory,” *J. Acoust. Soc. Am.* **123**, 2733–2749.
- Titze, I. R., Riede, T., and Popolo, P. (2008). “Nonlinear source-filter coupling in phonation: Vocal exercises,” *J. Acoust. Soc. Am.* **123**, 1902–1915.
- Titze, I. R., and Story, B. H. (1997). “Acoustic interactions of the voice source with the lower vocal tract,” *J. Acoust. Soc. Am.* **101**, 2234–2243.
- Titze, I. R., and Worley, A. S. (2009). “Modeling source-filter interaction in belting and high-pitched operatic singing,” *J. Acoust. Soc. Am.* **126**, 1530–1540.
- Tokuda, I. T., Zemke, M., Kob, M., and Herzel, H. (2007). “Biomechanical modeling of register transitions and the role of vocal tract resonators,” *J. Acoust. Soc. Am.* **122**, 519–531.
- Whittaker, E. T., and Watson G.N. (1927). *A Course in Modern Analysis* (Cambridge University Press, Cambridge, U. K.), pp. 136–137.
- Zañartu, M., Mehta, D. D., Ho, J. C., Wodicka, G. R., and Hillman, R. E. (2011). “Observation and analysis of in vivo vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study,” *J. Acoust. Soc. Am.* **129**, 326–339.