IOP Publishing — Fluid Dynamics Research

Fluid Dyn. Res. **42** (2010) 015001 (16pp) doi:10.1088/0169-5983/42/1/015001

INVITED PAPER

On the single-mass model of the vocal folds

M S Howe[1,3] **and R S McGowan**[2]

[1] Boston University, College of Engineering, 110 Cummington Street, Boston, MA 02215, USA
[2] CReSS LLC, 1 Seaborn Place, Lexington, MA 02420, USA

E-mail: mshowe@bu.edu

Received 9 June 2008, in final form 13 April 2009
Published 18 January 2010
Online at stacks.iop.org/FDR/42/015001

Communicated by S Akishita

Abstract
An analysis is made of the fluid–structure interactions necessary to support self-sustained oscillations of a single-mass mechanical model of the vocal folds subject to a nominally steady subglottal overpressure. The single-mass model of Fant and Flanagan is re-examined and an analytical representation of vortex shedding during 'voiced speech' is proposed that promotes cooperative, periodic excitation of the folds by the glottal flow. Positive feedback that sustains glottal oscillations is shown to occur during glottal contraction, when the flow separates from the 'trailing edge' of the glottis producing a low-pressure 'suction' force that tends to pull the folds together. Details are worked out for flow that can be regarded as locally two-dimensional in the glottal region. Predictions of free-streamline theory are used to model the effects of quasi-static variations in the separation point on the glottal wall. Numerical predictions are presented to illustrate the waveform of the sound radiated towards the mouth from the glottis. The theory is easily modified to include feedback on the glottal flow of standing acoustic waves, both in the vocal tract beyond the glottis and in the subglottal region.

1. Introduction

'Voiced speech' refers to that component of speech that involves sound generation by the vocal folds of the larynx (Fant 1960, Flanagan 1972, Stevens 1998). In its simplest manifestation, the vocal tract consists of a uniform waveguide divided into two sections at the glottis, which is the opening of variable width between the two vocal folds. 'Voicing' occurs when high pressure in the lungs opens the glottis producing a mean flow that induces quasi-periodic vibrations of the folds and the pulsing of air through the glottis,

[3] Author to whom any correspondence should be addressed.

0169-5983/10/015001+16$30.00

at frequencies f_0 that rarely exceed 400 Hz for adult speakers. The lung–glottis system is therefore equivalent to a monopole source of sound, but the overall monopole strength is determined by the details of the hydrodynamic and structural interactions occurring at the vocal folds. Vorticity is shed from the vibrating folds, and feedback of pressure fluctuations produced by the convecting vorticity drive the folds against elastic restoring forces into self-sustaining oscillations—in consequence the sound produced by the application of a nominally steady subglottal overpressure can alternatively be attributed to a combination of aeroacoustic *dipoles* determined by the vortex drag on the folds, together with a relatively weak monopole produced by small volumetric variations of fold structure.

The lung–glottis monopole interpretation is emphasized in early theoretical studies of voicing (Fant 1960, Flanagan 1972, Ishizaka and Flanagan 1972). More recent work has emphasized the aeroacoustic problem associated with the mechanism of sound production by vortex–surface interactions, both numerically (Zhao *et al* 2002, Hofmans *et al* 2003, Duncan *et al* 2006) and analytically (Howe and McGowan 2007).

Voicing often starts with a closed glottis subject to a subglottal overpressure of about 10 cm of water (∼1 kPa); muscular adjustment ensures that this is just sufficient to cause the folds to separate. Figure 1 (adapted from Zhao *et al* 2002) shows various stages of glottal motion during a complete cycle (of period $1/f_0$) for an idealized, axi-symmetric model of the folds that was used in numerical simulations. A subglottal pressure rise (on the left in the figure) causes fold separation from the 'rear' and the opening of the glottis, initially forming a converging channel as indicated in figure 1(b). Separation is subsequently accompanied by a 'rocking' motion of the folds (figures 1(c)–(f)) usually attributed to wavelike undulations of the exposed surfaces ('epithelia') of the folds (Titze 1994, Stevens 1998). A steady overpressure from the lungs produces a periodic repetition of the sequence of states illustrated in figure 1, except that numerical constraints required Zhao *et al* (2002) to adopt an 'open' rest state (figure 1(a)), for which the glottal open area was small but finite. The actual motion of the folds is governed by the fluid–structure interaction that occurs during voicing. The numerical work of Zhao *et al* (2002) and the subsequent analytical treatment of Howe and McGowan (2007) took no explicit account of these interactions—the fold motion during each cycle was prescribed to be similar to that illustrated in figure 1; this permitted attention to be focussed on the fluid mechanical aspects of sound generation.

Ishizaka and Flanagan (1972) developed the first fully interactive theory of voicing, in which the self-sustaining coupled oscillations of the vocal folds and the air are calculated simultaneously. According to this theory the motion of the folds was driven by surface pressure fluctuations determined by a model of the unsteady flow that was entirely irrotational *within* the glottis. Self-sustained oscillations could be realized provided each fold was modelled by two-adjacent, axially aligned and elastically coupled masses (respectively providing crude representations of the fore and aft sections of a rocking fold of figure 1). The changes in the potential flow produced by the relative motions of these masses produced the required phase relation between surface pressures and surface motions necessary to achieve self-sustaining oscillations. This treatment evolved from an earlier 'single-mass' model of Fant (1960) and Flanagan (1972) (figure 2) in which the glottal flow was approximated by flow through a uniform orifice whose cross-section varied in a prescribed manner with time, because potential flow fluctuations within the glottis cannot supply the necessary positive feedback from the flow required to maintain the system in oscillatory motion. More recent studies of the single-mass model have included coupling of the dynamics of the fold structure and the unsteady air flow through the glottis. Fulcher *et al* (2006) proposed a one-dimensional (1D) equation for fold motion involving an heuristic negative Coulomb damping term to represent aerodynamic excitation. This was extended by Zanartu *et al* (2007),

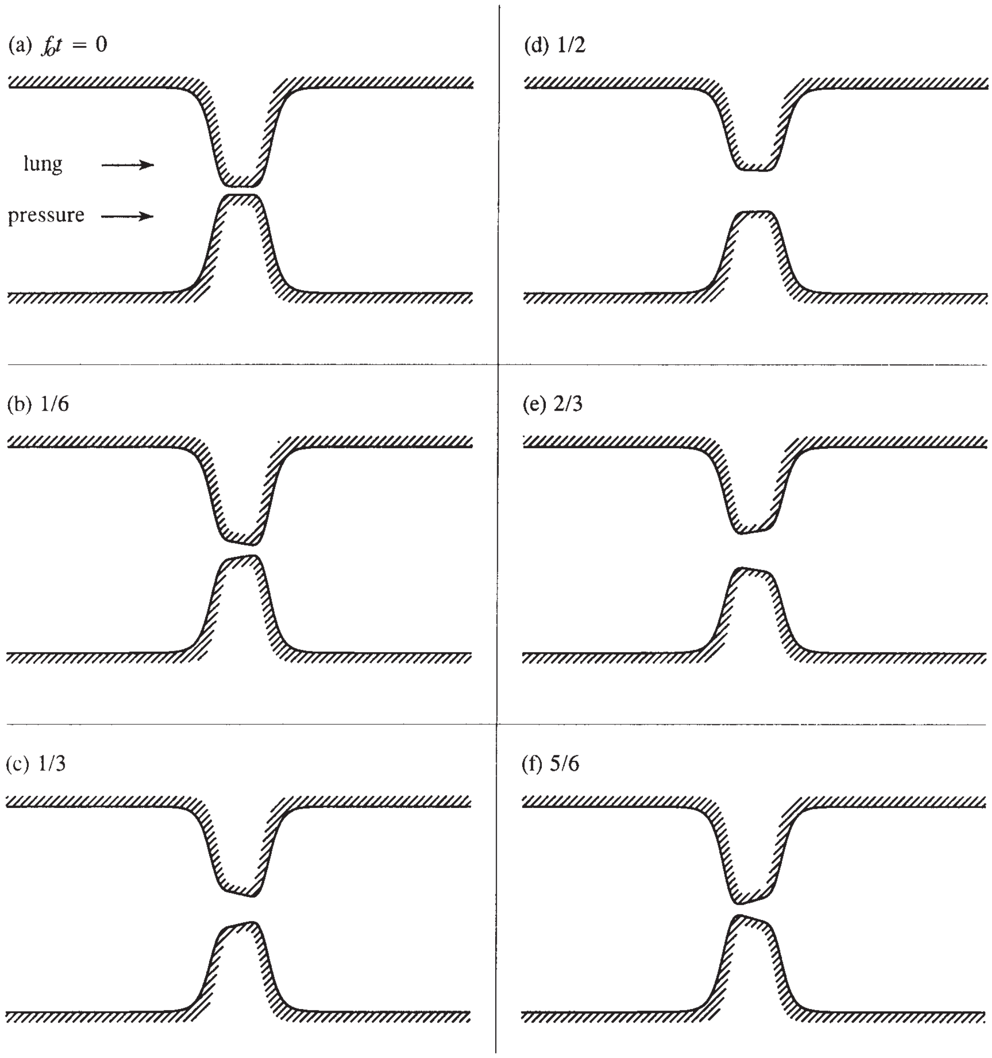

Figure 1. Variations in the geometry of the vocal folds over one cycle at intervals of $\frac{1}{6}$ of a period according to the model of Zhao *et al* (2002).

who achieved self-sustaining oscillations by calculating the negative damping force by assuming irrotational flow within the glottis modulated by a prescribed unsteady discharge coefficient (or *contraction ratio*) of the glottal jet and by acoustic feedback from the vocal tract.

However, these single-mass approximations take no account of the possibility of flow separation within the glottis, which might be expected to be the dominant 'negative damping' mechanism maintaining the motion of the folds. The point of separation from the glottal wall changes during their inward and outward oscillations producing a much stronger variation in surface forces, much larger than those attributable merely to changes in the jet discharge coefficient. The manner in which the separation point moves for the simple single-mass model of figure 2 is illustrated in figure 3. This shows the contemplated picture of the flow from

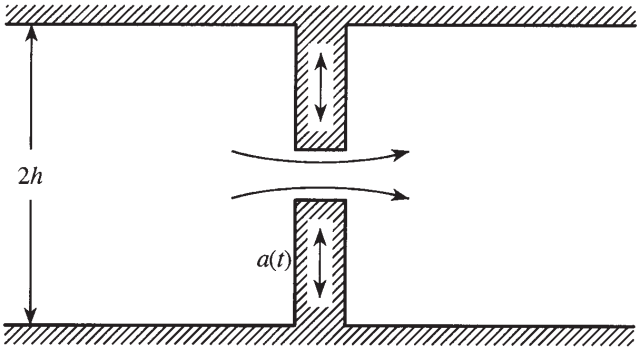

Figure 2. Simple orifice approximation to the glottis with prescribed area variations.

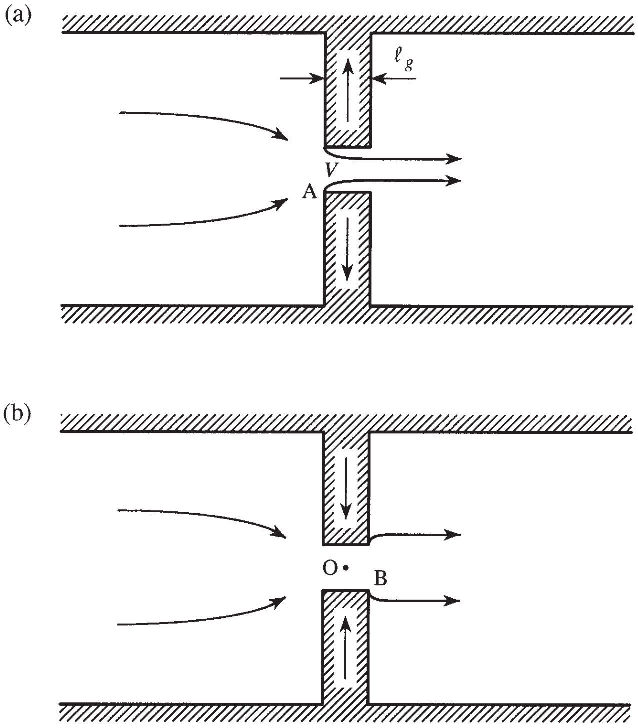

Figure 3. 'One mass' model for self-sustained oscillations of the glottis: (a) vortex shedding occurs from the leading edge of an expanding glottis; (b) shedding occurs from the aft edge during contraction.

left to right through the glottis in two conditions (a) when the glottis is expanding, and (b) when its area decreases. In both cases vorticity is shed from a 'trailing edge'. In case (a) the incoming flow separates at the front corners A of the orifice, so that the pressure applied over the glottal walls is uniform and approximately the same as the uniform pressure in the vocal tract just downstream of the glottis. In voiced speech the effective Strouhal number of the unsteady motion within the glottis $f_0\ell_g/V \sim 0.1$ (where the mean flow speed across the glottis of axial extent ℓ_g is typically of order $V \sim 20\,\mathrm{m\,s^{-1}}$). This is sufficiently small that in a first approximation it may be assumed that, over a cycle of oscillation at frequency f_0, the fluid passes through a continuous sequence of quasi-static states of motion.

Therefore, during expansion vorticity is shed from the fore edge A of the orifice in the manner illustrated in figure 3(a). After attaining its maximum size the glottis begins to contract and initially continues to shed vorticity from A. However, continued contraction eventually causes re-attachment of the separated flow to the advancing glottal wall, so that the forward (or subglottal) section of the wall is wetted by the incoming airflow. The initial corner motion around A then has the essential high-speed, low-pressure characteristics of potential flow, which tends to promote a rapid collapse of separation and complete wetting of the glottal walls, such that free shear layer formation now jumps to the aft corners B of the orifice (figure 3(b)), beyond which the flow expands into a jet whose cross-section is marginally larger than that of the orifice. The resulting accelerated flow within the glottis and a corresponding large decrease in the glottal wall pressure (below that in the downstream region), produce 'suction' forces over the glottal walls that promote the rate of decrease in glottal area. This extended picture of the motion within and near the glottis evidently includes the effects on surface force produced by fluctuations in both the discharge coefficient and the separation point. It should be noted, however, that the phase of the separated flow and vortex shedding from the ends of the glottis is effectively the *reverse* of that implied by the deformable model of Zhao *et al* (2002).

This paper is concerned with an analytical investigation of the implications of the model of figure 3 for voicing. An equation of motion for the single-mass model that includes nonlinear feedback from the unsteady glottal flow is derived in section 2. The method of Howe and McGowan (2007) is then used in section 3 to determine the coupled nonlinear system of equations governing the production of sound by a steady subglottal pressure rise applied to the folds. Sample numerical predictions of the waveform transmitted into the upper vocal tract are discussed in section 4.

2. Fluid-coupled motion of the vocal folds

We consider a simplified, two-dimensional model of the vocal tract. It is treated as an infinite, rigid duct of rectangular cross-section and area $A = 2h \times \ell_3$, where $2h$ is the duct width shown in figure 2 and ℓ_3 is the spanwise dimension, out of the plane of the paper in figure 3. Take the origin of coordinates $\mathbf{x} = (x, y, z)$ at the centre-point O of the glottis, with the x-axis parallel to the duct axis and in the direction of mean flow from the lungs; the y- and z-axes are then taken to form a right-handed system, respectively, parallel to the h and ℓ_3 directions. The two-dimensional folds of thickness $\ell_g \ll h$ extend over the full span ($|z| < \ell_3/2$) of the duct and have length $a(t) \leqslant h$ varying with the time t.

2.1. Separation during glottal closure: the suction force

It is assumed that for the major part of a voicing cycle the motion within the glottis is at sufficiently high Reynolds number that the direct action of viscosity within the flow can be neglected. Pressure feedback of the flow on the fold is then confined to the walls of the glottis, i.e. to the end-faces of the rectangular fold models. The end forces are negligible when the glottis expands and separation occurs at the entrance A of the mean flow through the glottis. During glottis contraction the end faces experience suction forces tending to promote glottal closure, and when the aspect ratio ℓ_g/h is small the force may be determined by neglecting the finite thickness ℓ_g of the glottis. The suction effect is then calculated by

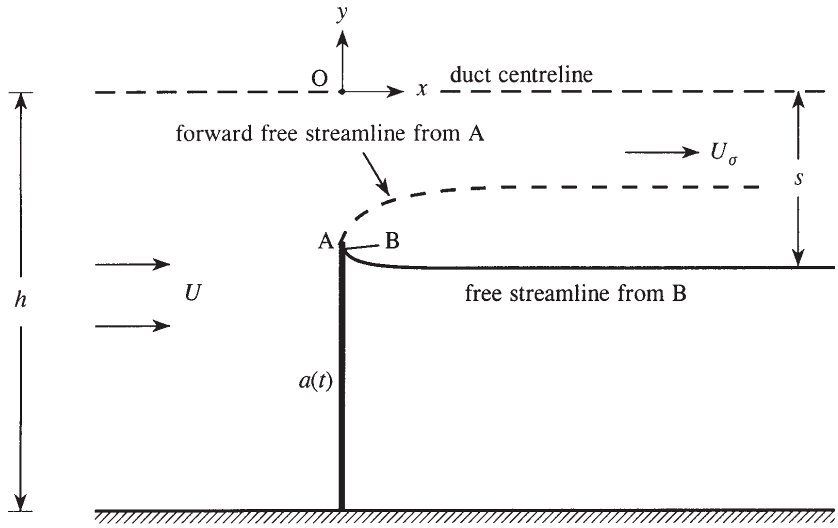

Figure 4. Free streamline model used to calculate the suction force on the glottis wall and the jet contraction ratio σ. The separation point B on the supraglottal face $x = +0$ of the blade is a small distance ℓ_B below the tip.

replacing the rectangular fold by a rigid *blade* of infinitesimal thickness (lying along the y-axis) and using the theory of potential flow around the blade edge (Batchelor 1967, Birkhoff and Zarantonello 1957, Gurevich 1965, Howe 2007) according to the simplified free streamline model depicted in figure 4.

In the contraction phase of the glottal motion the free streamline is shed from point B, which is now displaced from the tip to the right-hand (or supraglottal) side of the thin-blade vocal fold model of figure 4; point B is a small distance ℓ_B, say, from the tip on the surface $x = +0$ of the blade (too small to be shown clearly in the figure). The free streamline forms the quasi-static boundary of the jet flow emerging from the glottis. In the downstream region the uniform *half-width* of the jet is denoted by $s(t)$ and the velocity within the jet is uniform and equal to U_σ. This velocity is related to the mean axial velocity V within the glottis by

$$V = \sigma U_\sigma, \quad \sigma = \frac{s(t)}{h - a(t)}, \tag{1}$$

where the contraction ratio $\sigma > 1$ during contraction of the glottis.

The value of σ and the separation distance ℓ_B can be found for a prescribed value of a/h by transforming the region of the $\mathcal{Z} = x + iy$ plane occupied by the flow in figure 4 (at uniform speed $U = U_\sigma s/h$ in the duct from $x = -\infty$, bounded by the centreline of the duct and the free streamline from B in $x > 0$) onto the upper half of the ζ plane, by means of the mapping (Birkhoff and Zarantonello 1957, Gurevich 1965, Howe 2007)

$$\mathcal{Z} + i(h - a) = \frac{3^{1/2}s}{\pi} \int_0^\zeta \frac{\xi(2i + \sqrt{3\xi - 1})\, d\xi}{(\xi + \beta)\sqrt{\xi + 1}(i + \sqrt{3\xi - 1})^2}, \quad \operatorname{Im} \zeta > 0, \tag{2}$$

where the path of integration passes above all singularities of the integrand on the real axis and the square roots in the integrand take their principal values. The coefficient $\beta > 1$ is real

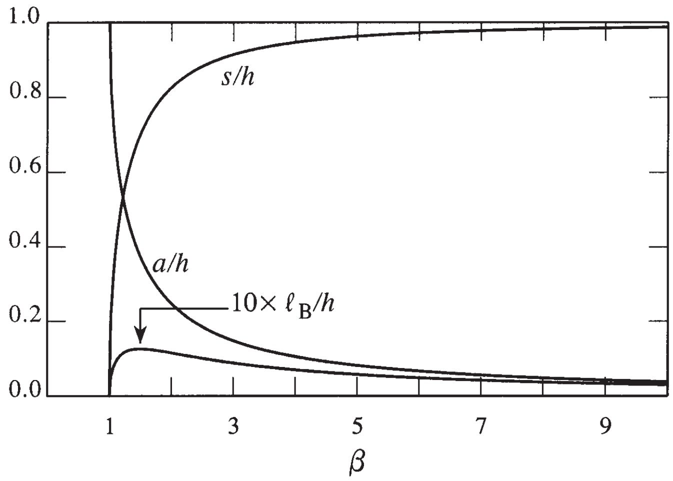

Figure 5. Variation of the ratios s/h, a/h, ℓ_B/h defined in terms of the parameter β by equations (3).

and determines parametrically the ratios

$$\left.\begin{aligned} \frac{s}{h} &= \frac{\sqrt{\beta-1}(1+\sqrt{1+3\beta})^2}{3^{1/2}\beta(2+\sqrt{1+3\beta})} \\ \frac{a}{h} &= \frac{3^{1/2}}{\pi}\frac{s}{h}\int_0^1 \frac{\xi(2+\sqrt{1+3\xi})\,\mathrm{d}\xi}{(\xi-\beta)\sqrt{1-\xi}(1+\sqrt{1+3\xi})^2} \\ \frac{\ell_B}{h} &= \frac{3^{1/2}}{\pi}\frac{s}{h}\int_0^{1/3} \frac{\xi(2+\sqrt{1-3\xi})\,\mathrm{d}\xi}{(\xi+\beta)\sqrt{1+\xi}(1+\sqrt{1-3\xi})^2} \end{aligned}\right\}. \tag{3}$$

The variations of these quantities with β are displayed in figure 5.

The velocity potential w (of the motion within the jet and in the glottis) in the ζ plane corresponds to the flow generated by a point source of strength $U_\sigma s$ on the real axis at $\zeta = -\beta$, so that

$$w = \frac{U_\sigma s}{\pi}\ln(\zeta+\beta), \quad \mathrm{Im}\,\zeta > 0. \tag{4}$$

The blade tip ($\mathcal{Z} = -\mathrm{i}(h-a)$) maps into the origin in the ζ plane. It therefore follows from equations (2) and (4) that near the blade tip

$$w \sim \frac{U_\sigma s}{\pi}\ln\beta + \frac{U_\sigma s}{\pi\beta}\left(\frac{8\pi\mathrm{i}\beta}{3^{3/2}s}\right)^{1/2}\sqrt{\mathcal{Z}+\mathrm{i}(h-a)}. \tag{5}$$

This characterizes the singular behaviour of the velocity near the edge, in terms of which the suction force $\mathcal{F}$ per unit span can be calculated (Batchelor 1967, equation 6.5.4, Howe 2007, equation 3.3.7) in the form

$$\mathcal{F} = \frac{2}{3^{3/2}\pi\beta}\rho_o U_\sigma^2 s, \tag{6}$$

where ρ_o is the mean air density.

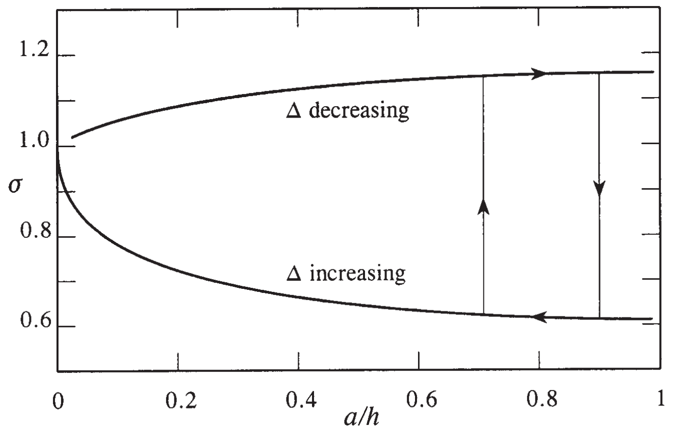

Figure 6. Dependence of the jet contraction ratio σ on the nondimensional blade length a/h for closing and expanding glottal width $\Delta = 2(h-a)$.

2.2. *Separation during glottal expansion*

During the expansion phase of the glottal motion separation occurs at the 'leading edge' A of the blade. The free streamline is tangential to the blade at separation and results in the formation of a jet (the broken line free streamline from A in figure 4) with a contraction ratio σ typically of order 0.62 for the relevant motions of the vocal fold model. In this case we obtain the parametric representations

$$\left.\begin{aligned}\frac{s}{h} &= \frac{\beta'}{1+\sqrt{1+\beta'^2}}\\ \frac{a}{h} &= \frac{s}{h}\left[\frac{2}{\pi\beta'}\tan^{-1}\left(\frac{1}{\beta'}\right)+\frac{\sqrt{1+\beta'^2}}{\beta'}-1\right]\end{aligned}\right\}, \quad \beta'>0. \tag{7}$$

The flow velocity at the blade tip is finite and there is no suction force.

2.3. *The jet contraction ratio*

The parametric representations (3) and (7), respectively, for the closing and expanding glottis can be used to calculate the dependence of the contraction ratio σ on the nondimensional length a/h of the blade during the opening and closing phases of the glottal motion (figure 6).

Evidently, our simple thin-blade model of the glottis implies that during periodic opening and closing $\sigma \equiv \sigma(t)$ will exhibit jump discontinuities at the minimal and maximal glottal areas, and may be considered to traverse a closed loop of the type indicated in the figure for each cycle of the motion. This conclusion is consistent with the experimental findings of Park and Mongeau (2007) and has been used in the Zanartu *et al* (2007) model of vocal fold mechanics. The discontinuities correspond to rapid variations in the separation point on the surface of a real glottal fold, and would in practice occur over a time interval $\sim C\ell_g/V$, where C is a coefficient of order 2 or 3 and V is the current mean velocity within the glottis.

2.4. Equation of motion of the glottal folds

Symmetric motions of the model folds of figure 3 are assumed to be governed by the following oscillator equation for the fold length $a(t)$:

$$\frac{\mathrm{d}^2 a}{\mathrm{d}t^2} + \gamma \frac{\mathrm{d}a}{\mathrm{d}t} + \Omega^2 (a - a_0) = \frac{F(t)}{m}, \tag{8}$$

where m is the effective mass of each fold, a_0 is the equilibrium fold length, γ is a structural damping coefficient and Ω is the fold resonance frequency determined by the elastic properties of the fold. The aerodynamic force $F = \ell_3 \mathcal{F}$ when $\mathrm{d}a/\mathrm{d}t \geqslant 0$ (case (b) of figure 3) and otherwise is zero, i.e.

$$F = \begin{cases} \dfrac{2}{3^{3/2}\pi\beta} \rho_o U_\sigma^2 \sigma(t) \ell_3 [h - a(t)], & \dfrac{\mathrm{d}a}{\mathrm{d}t} \geqslant 0, \\ 0, & \dfrac{\mathrm{d}a}{\mathrm{d}t} < 0. \end{cases} \tag{9}$$

Equations (8) and (9) involve the two unknown functions $a(t)$, $U_\sigma(t)$, and must be solved simultaneously with the equation describing the production of aerodynamic sound. Observe, that provided $a = a_0 < h$ in the rest state ($\mathrm{d}a/\mathrm{d}t = 0$), motion of the folds is initiated *spontaneously* by the arrival of a pressure load from the lungs, because any 'starting flow' through the glottis is necessarily irrotational, and therefore produces a nonzero suction force (see, e.g., Howe 2007).

The mutual impact of the vocal folds occurs when $a \to h$. In solving equation (8) it is assumed that the impact is inelastic, at which time $\mathrm{d}a/\mathrm{d}t$ is reduced to zero.

3. The aerodynamic sound

3.1. The aerodynamic sound equation

To calculate the sound produced it is assumed that the vocal tract is acoustically rigid and that the characteristic frequencies are small enough that only plane waves can propagate. The fluid motion is taken to be *homentropic* and of low Mach number, such that the convection of sound by the flow can be ignored. The equation governing sound production then reduces to (Lighthill 1952, Howe 1998)

$$\left(\frac{1}{c_o^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) B = \mathrm{div}(\boldsymbol{\omega} \wedge \mathbf{v}), \tag{10}$$

where c_o is the mean speed of sound and B is the *total enthalpy*, defined for homentropic flow by

$$B = \int \frac{\mathrm{d}p}{\rho} + \frac{1}{2} v^2, \tag{11}$$

where $\mathbf{v}$ is the velocity, $\boldsymbol{\omega} = \mathrm{curl}\, \mathbf{v}$ the vorticity and $\rho \equiv \rho(p)$ is the density.

The total enthalpy B is constant throughout the flow in the absence of vorticity and moving boundaries (Howe 1998). Fluctuations in B propagate as sound away from regions of unsteady vorticity where $\mathrm{div}(\boldsymbol{\omega} \wedge \mathbf{v}) \neq 0$ and from regions of boundary motion. In the propagation zone (several duct diameters from the glottis) the motion is irrotational with velocity potential $\varphi(x, t)$, say, where Bernoulli's equation indicates that $B = -\partial\varphi/\partial t$, and the acoustic pressure is $p = -\rho_o \partial\varphi/\partial t = \rho_o B$.

The acoustic equation (10) is solved by use of Green's function $G(\mathbf{x}, \mathbf{y}, t, \tau)$, which is the field generated by a point source at $\mathbf{y} \equiv (x', y', z')$ at time τ, and satisfies

$$\left(\frac{1}{c_o^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) G = \delta(\mathbf{x}-\mathbf{y})\delta(t-\tau), \quad G = 0 \text{ for } t < \tau. \tag{12}$$

We also impose the condition that the normal derivative $\partial G/\partial x_n = 0$ on the instantaneous moving surface of the glottis and on the walls of the vocal tract. Further details of the calculation of $G(\mathbf{x}, \mathbf{y}, t, \tau)$ in the compact limit in which the acoustic source region is small compared to the acoustic wavelengths are given by Howe and McGowan (2007). It will be sufficient here to quote the following representation of Green's function for the present problem in the compact limit:

$$\begin{aligned} G(\mathbf{x}, \mathbf{y}, t, \tau) &\approx \frac{c_o}{2A} \mathrm{H}\left(t - \tau - \frac{|x - Y|}{c_o}\right) \\ &\approx \frac{c_o}{2A}\left\{\mathrm{H}\left(t - \tau - \frac{|x|}{c_o}\right) + \frac{\operatorname{sgn}(x)}{c_o} Y(\mathbf{y}, \tau)\delta\left(t - \tau - \frac{|x|}{c_o}\right)\right\}, \quad |x| \to \infty. \end{aligned} \tag{13}$$

In this formula H denotes the Heaviside unit step function, and Y is the (1D) Kirchhoff vector defined such that $\partial Y/\partial x' \to 1$ as $x' \to \pm\infty$. Y is a solution of Laplace's equation satisfying $\partial Y/\partial x'_n = 0$ on the surface of the vocal tract and on the vocal folds at time τ, and can be normalized such that

$$Y \sim x' \pm \frac{\ell'(\tau)}{2}, \quad x' \to \pm\infty, \tag{14}$$

where ℓ' is the time-dependent Rayleigh 'end correction' of the glottis (Rayleigh 1945), given by

$$\ell'(\tau) = \int_{-\infty}^{\infty} \left(\frac{\partial Y}{\partial x'}(\mathbf{y}, \tau) - 1\right) \mathrm{d}x', \tag{15}$$

where the integration is along the axis of the duct. The function $Y(\mathbf{y}, \tau)$ can be interpreted as the velocity potential of incompressible flow through the glottis having unit speed at large distances from the vocal folds. Approximation (13) is applicable provided the observation point $\mathbf{x}$ lies in the acoustic far field of the glottis (see Howe and McGowan (2007) for further discussion).

3.2. Sound produced by a steady subglottal pressure

Let a subglottal pressure rise $p_I \mathrm{H}(t - x/c_o)$ of amplitude p_I impinge on the glottis at $x = 0$ at time $t = 0$. This disturbance triggers motion of the glottal folds, induces jet flow through the glottis and generates sound that propagates away from the glottis in both directions in the vocal tract. By following the procedure described in detail by Howe and McGowan (2007), the resulting net pressure within the vocal tract can be expressed in the form

$$\begin{aligned} p(x, t) = p_I &\left\{\mathrm{H}\left(t - \frac{x}{c_o}\right) - \frac{\ell' \operatorname{sgn}(x)}{2c_o}\,\delta\left(t - \frac{|x|}{c_o}\right)\right\} \\ &- \rho_o \int (\boldsymbol{\omega} \wedge \mathbf{v})(\mathbf{y}, \tau) \cdot \frac{\partial G}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}, t, \tau)\, \mathrm{d}^3\mathbf{y}\, \mathrm{d}\tau \\ &+ \rho_o \oint_{S(\mathbf{y},\tau)} G(\mathbf{x}, \mathbf{y}, t, \tau) \frac{\partial \mathbf{v}}{\partial \tau}(\mathbf{y}, \tau) \cdot \mathrm{d}\mathbf{S}(\mathbf{y})\, \mathrm{d}\tau, \end{aligned} \tag{16}$$

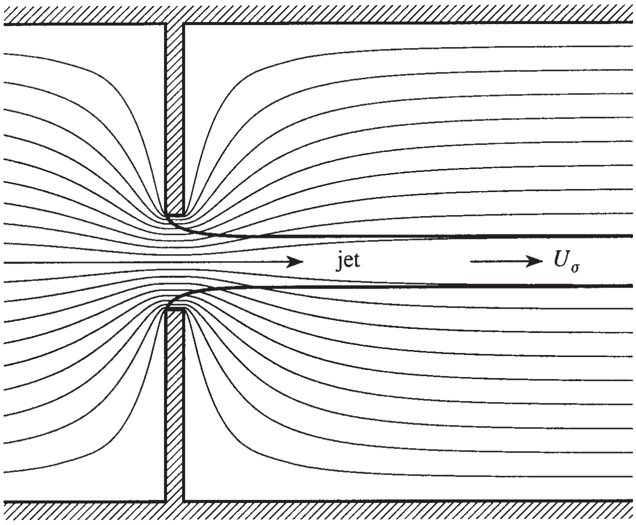

Figure 7. Quasi-static jet modelled by free streamline flow with asymptotic jet speed U_σ during the expansion phase of the glottal motion. Sound production is confined to the region close to the glottis where the 'streamlines' of the Green's function axial Kirchhoff vector $Y(\mathbf{x}, t)$ cut through the jet boundary.

where we have put $\ell' = \ell'(0)$, the end correction at the arrival time of the pressure rise from the lungs. The time integrations are over all values of the source time $-\infty < \tau < \infty$; the first integration is over the region of the fluid containing vorticity, here confined to the shear layers of the jet, and the surface integral is taken over the surfaces $S(\mathbf{y}, \tau)$ of the vocal folds.

The terms in the second line of equation (13) are, respectively, the leading-order monopole and dipole terms in a formal multipole expansion of Green's function. Only the dipole term contributes to the volume source integral in (16), so that

$$-\int (\boldsymbol{\omega} \wedge \mathbf{v})(\mathbf{y}, \tau) \cdot \frac{\partial G}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}, t, \tau)\, \mathrm{d}^3\mathbf{y}\, \mathrm{d}\tau \approx -\frac{\operatorname{sgn}(x)}{2A} \left[\int \frac{\partial Y}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v}\, \mathrm{d}^3\mathbf{y} \right]_{t-|x|/c_o}. \tag{17}$$

The vorticity is confined to the free streamlines at the edges ('free shear layers') of the jet emerging from the glottis. Figure 7 illustrates the general form of these free streamlines during glottal opening. Also shown are the 'streamlines' of the hypothetical flow defined by the velocity potential $Y(\mathbf{y}, \tau)$. On the upper/lower free streamlines of the jet the vorticity $\boldsymbol{\omega} = \pm U_\sigma \delta(s_\perp)\mathbf{k}$, where $s_\perp$ is the distance measured in the direction of the outward normal of the jet and $\mathbf{k}$ is a unit vector out of the plane of the paper; the vorticity convection velocity $\mathbf{v} = \frac{1}{2} U_\sigma \mathbf{t}$, where $\mathbf{t}$ is a unit vector tangential to the free streamline. Therefore, $\boldsymbol{\omega} \wedge \mathbf{v}$ is parallel to $\pm \mathbf{k} \wedge \mathbf{t}$, the outward unit normal from the jet, so that the main contribution to integral (17) is from that section of the jet *very close to the glottis* where the jet shear layers are cut by the Y-streamlines. The velocity U_σ does not vary with position over this short length of the jet, and this shows that

$$\int \left(\frac{\partial Y}{\partial \mathbf{y}} \cdot \boldsymbol{\omega} \wedge \mathbf{v} \right) (\mathbf{y}, \tau)\, \mathrm{d}^3\mathbf{y} = \ell_3 U_\sigma^2(\tau) \int_0^\infty \left(\frac{\partial Y}{\partial s_\perp} \right)_{s_\perp = 0} \mathrm{d}s,$$

where the final integration is with respect to the distance s along either free streamline from the glottis. Continuity and the limiting behaviours (14) of the function Y imply that

$2\ell_3 \int_0^\infty (\partial Y/\partial s_\perp)_{s_\perp=0}\, ds = A$, provided that the asymptotic jet width $2s$ is small compared to the width $2h$ of the vocal tract. Therefore the integrals in (17) reduce to

$$-\int (\boldsymbol{\omega} \wedge \mathbf{v})(\mathbf{y}, \tau) \cdot \frac{\partial G}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}, t, \tau)\, \mathrm{d}^3\mathbf{y}\, \mathrm{d}\tau \approx -\frac{\mathrm{sgn}(x)}{4} U_\sigma^2 \left(t - \frac{|x|}{c_o}\right). \tag{18}$$

The final integral on the right of (16) determines the *monopole* sound produced by the small periodic variations in vocal fold volume, which occur at a rate $2\ell_3\ell_g\, \mathrm{d}a/\mathrm{d}t$. The monopole sound radiates a symmetric waveform in both directions from the glottis. Only the first term in the multipole expansion (13) of G contributes to this term, yielding

$$\oint_{S(\mathbf{y},\tau)} G(\mathbf{x}, \mathbf{y}, t, \tau) \frac{\partial \mathbf{v}}{\partial \tau}(\mathbf{y}, \tau) \cdot \mathrm{d}\mathbf{S}(\mathbf{y})\, \mathrm{d}\tau = \frac{\ell_3\ell_g c_o}{A} \frac{\mathrm{d}a}{\mathrm{d}t}\left(t - \frac{|x|}{c_o}\right). \tag{19}$$

Using results (18) and (19) representation (16) of the sound in the vocal tract becomes

$$p(x,t) \approx p_I \left\{ \mathrm{H}\left(t - \frac{x}{c_o}\right) - \frac{\ell' \,\mathrm{sgn}(x)}{2c_o} \delta\left(t - \frac{|x|}{c_o}\right) \right\} - \frac{\mathrm{sgn}(x)}{4} \rho_o U_\sigma^2 \left(t - \frac{|x|}{c_o}\right) + \frac{\ell_3\ell_g}{A} \rho_o c_o \frac{\mathrm{d}a}{\mathrm{d}t}\left(t - \frac{|x|}{c_o}\right). \tag{20}$$

The second term in the brace brackets is a transient that represents the 'passive' scattering of the incident pressure rise at the glottis occurring over the short time $\sim O(\ell'/c_o)$.

3.3. Self-consistent equation for U_σ

The final, monopole term on the right of equation (20) is defined in terms of the motion $a(t)$ of the vocal folds. To obtain a similar representation of the U_σ term we consider

$$p'(x,t) = p_I \left\{ \mathrm{H}\left(t - \frac{x}{c_o}\right) - \frac{\ell' \,\mathrm{sgn}(x)}{2c_o} \delta\left(t - \frac{|x|}{c_o}\right) \right\} - \frac{\mathrm{sgn}(x)}{4} \rho_o U_\sigma^2 \left(t - \frac{|x|}{c_o}\right). \tag{21}$$

The corresponding plane wave acoustic particle velocity in $x > 0$ is $p'(x,t)/\rho_o c_o$. If

$$U(t) = \lim_{x \to +0} \frac{p'(x,t)}{\rho_o c_o} \tag{22}$$

is the limiting value of this velocity just downstream of the glottis, then equation (21) shows that

$$2c_o U + \frac{1}{2} U_\sigma^2 = \frac{2p_I}{\rho_o}\left(\mathrm{H}(t) - \frac{\ell'}{2c_o}\delta(t)\right) \approx \frac{2p_I}{\rho_o} \mathrm{H}\left(t - \frac{\ell'}{2c_o}\right), \tag{23}$$

provided $f_0\ell'/c_o \ll 1$. But continuity through the glottis requires from (1) that $U_\sigma = V/\sigma$, where $V[h - a(t)] = Uh$, and therefore that

$$U_\sigma = \frac{h}{[h - a(t)]} \frac{U}{\sigma}, \tag{24}$$

so that equation (23) becomes

$$2c_o U + \frac{1}{2}\left(\frac{h}{\sigma[h - a(t)]}\right)^2 U^2 = \frac{2p_I}{\rho_o} \mathrm{H}\left(t - \frac{\ell'}{2c_o}\right), \tag{25}$$

with solution representing positive pressure pulsing towards $x > 0$

$$U(t) = 2c_o\sigma\left(\frac{h-a(t)}{h}\right)\left(\sqrt{\left(\frac{\sigma[h-a(t)]}{h}\right)^2 + \frac{p_I}{\rho_o c_o^2}} - \frac{\sigma[h-a(t)]}{h}\right), \quad t > \frac{\ell'}{2c_o}. \tag{26}$$

This result determines $U(t)$ when $a(t)$ is known, after which the radiation in the vocal tract can be taken in the form

$$p(x,t) = \rho_o c_o U\left(t - \frac{x}{c_o}\right) + \frac{\ell_3 \ell_g}{A}\rho_o c_o \frac{da}{dt}\left(t - \frac{x}{c_o}\right), \quad x \to +\infty \tag{27a}$$

$$= p_I\left[\mathrm{H}\left(t - \frac{x}{c_o}\right) + \mathrm{H}\left(t + \frac{x}{c_o}\right)\right] - \rho_o c_o U\left(t + \frac{x}{c_o}\right)$$
$$+ \frac{\ell_3 \ell_g}{A}\rho_o c_o \frac{\mathrm{d}a}{\mathrm{d}t}\left(t + \frac{x}{c_o}\right), \quad x \to -\infty. \tag{27b}$$

4. Numerical results

The computations of Zhao *et al* (2002) and the corresponding analytical predictions of Howe and McGowan (2007) for the model vocal folds with the prescribed motions illustrated in figure 1 indicated that predicted waveforms for steady excitation of the glottis by a constant over pressure p_I rapidly attains a quasi-steady form. Indeed equation (23) is the precise quasi-static analogue of the nonlinear, lumped-parameter equation of Fant (1960) and Flanagan (1972) and of the corresponding equation (6.5) in Howe and McGowan (2007).

Predictions of the acoustic waveform (27*a*) for radiation downstream of the glottis (towards the mouth, $x > 0$) are shown in figures 8 and 9. These results are derived by numerical integration of the glottal equation (8), wherein the force $F(t)$ is given in terms of $a(t)$ by equations (9), (24) and (26). The solid-line curve (——) in figure 8(a) depicts three cycles of the pressure normalized by the incident wave amplitude p_I when the frequency $f_0 \equiv \Omega/2\pi = 125$ Hz (typical of an adult male). In all of our predictions it is assumed that the incident step pressure rise $p_I = 10$ cm of water (~1 kPa) and also that $c_o = 340\,\mathrm{m\,s^{-1}}$, $\rho_o = 1.23\,\mathrm{kg\,m^{-3}}$. The following typical values have also been assumed for other parameters entering the formulae:

$$h = 10\,\mathrm{mm}, \quad \ell_3/h = 1, \quad \ell_g/h = 0.3, \quad m = 0.5 \times 10^{-4}\,\mathrm{kg}, \quad \gamma/\Omega = 0.01. \tag{28}$$

Motion of the vocal folds starts with the arrival of the step pressure rise at the glottis at $t = 0$. At this time it is assumed that $a = a_0$, $\mathrm{d}a/\mathrm{d}t = 0$, where in figure 8 $a_0/h = 0.95$. The initial flow through the undisturbed glottis must be irrotational and necessarily produces a suction force that causes the first movement of the folds. Integration of equation (8) may therefore be started from a condition of rest at $t = 0$. Figure 8(a) also displays the separate contributions from the terms on the right of (27*a*): the first ($\rho_o c_o U$) is the 'vortex sound' contribution from the jet superimposed on the mean transmitted pressure of amplitude p_I (• • •); the second is the much weaker monopole (- - - -) associated with volume changes of the vocal folds.

The corresponding variations in $a(t)/h$ are shown in figure 8(b). The maximum acoustic pressure occurs at the retarded time at which $a(t)$ is a minimum, when the glottis opening is a maximum. At this instant vortex shedding jumps from the rear to the forward end of the glottal wall, corresponding to the upward jump in the contraction ratio in figure 6. In practice,

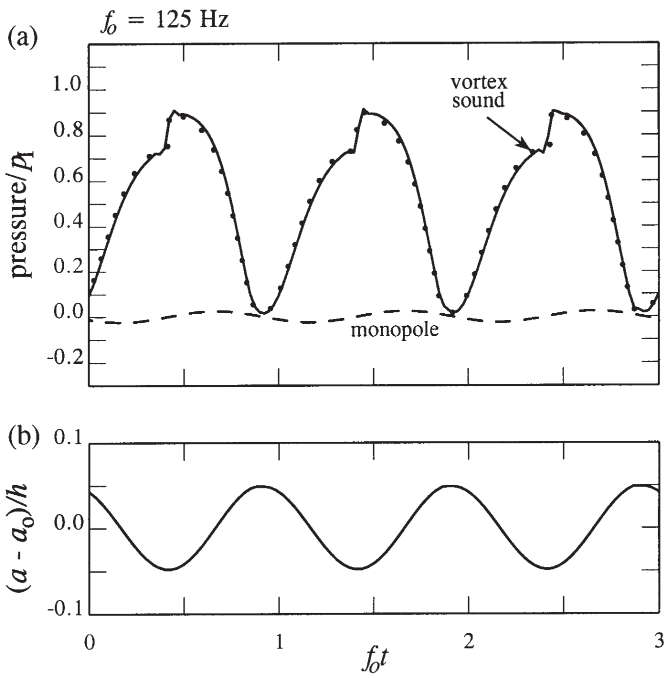

Figure 8. (a) Three cycles of the pressure ([27*a*]) (——) radiated from the glottis towards the mouth ($x > 0$) for conditions ([28]) when $f_0 = 125$ Hz, $a_0/h = 0.95$ and $p_I = 10$ cm of water. Also shown are the separate contributions from the vortex sound superimposed on the mean transmitted pressure p_I (• • •) and the monopole sound ([19]) (- - - -). (b) The corresponding variation of vocal fold amplitude $a(t)/h$.

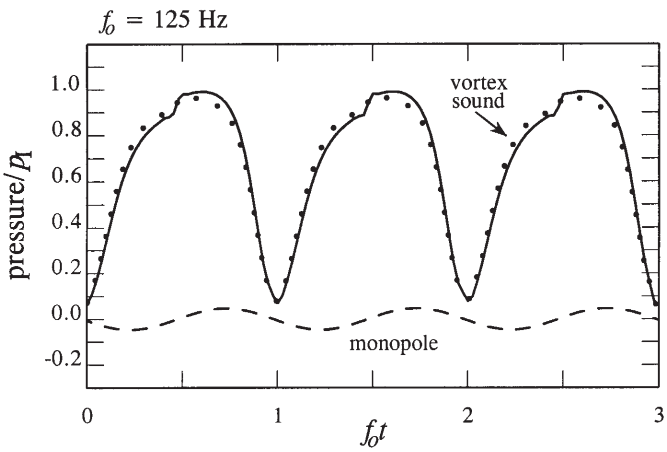

Figure 9. As for figure 8(a) with $a_0/h = 0.9$.

the variation in shedding position would occur in a finite time—in our case the discontinuity produces a corresponding unrealistic jump in the acoustic pressure. This phenomenon was also noted by Zanartu *et al* (2007); the acoustic jump was avoided by adopting an artificially smoothed variation of $\sigma(t)$ across the discontinuity.

Figure 9 illustrates the corresponding acoustic waveforms when $a_0 = 0.9h$, when other parametric values are unchanged. In this case the rapid increase in the contraction ratio has a smaller effect because both before and after the jump the glottis is sufficiently 'open' to allow for almost complete transmission of the incident pressure p_I into $x > 0$. The fundamental oscillations in both figures 8 and 9 are at the natural frequency f_0 of the folds. The fold motion is very regular, and practically sinusoidal; the variation in the suction force with the rapid changes in the flow separation point is responsible for the rapid variation in the transmitted acoustic pressure near its peak.

5. Conclusion

Voiced speech is a consequence of self-sustained oscillations of the vocal folds when subjected to slowly varying subglottal pressure loading. This is achieved by cooperative, periodic excitation of the folds by the unsteady glottal flow. For the single-mass mechanical model discussed in this paper, positive feedback from the flow occurs when the glottis is contracting causing separation from the glottal trailing edge and the consequent appearance of a low-pressure 'suction' that tends to pull the folds together.

The detailed calculations reported in this paper are based on the assumption that the flow in the glottal region is 2D. This is not an essential limitation. An axisymmetric model of the same kind can be worked out as a minor extension of the analysis (in this case by making use, for example, of experimental data on the unsteady jet contraction ratio σ reported by Park and Mongeau (2007)). Similar modifications can be made to account for the feedback on the glottal flow of standing acoustic waves, both in the vocal tract beyond the glottis and in the subglottal region. In the earlier single-mass models of Fulcher *et al* (2006) and Zanartu *et al* (2007), the surface forces driving the glottal motion are modulated by the unsteady variations in the discharge coefficient. In this paper it has been argued that the variations in the discharge coefficient are merely a manifestation of the periodic motion of the separation point of the jet from the glottal wall, and that it is this variation in the separation position that actually controls the time dependence of the surface force.

Acknowledgments

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.

References

Batchelor G K 1967 *An Introduction to Fluid Dynamics* (Cambridge: Cambridge University Press)
Birkhoff G and Zarantonello E H 1957 *Jets, Wakes and Cavities* (New York: Academic Press)
Duncan C, Zhai G and Scherer R 2006 Modeling coupled aerodynamics and vocal fold dynamics using immersed boundary methods *J. Acoust. Soc. Am.* **120** 2859–71
Fant G 1960 *Acoustic Theory of Speech Production* (The Hague: Mouton)
Flanagan J L 1972 *Speech Analysis Synthesis and Perception* 2nd edn (New York: Springer)
Fulcher L P, Scherer R C, Melnykov A, Gateva V and Limes M E 2006 Negative Coulomb damping, limit cycles, and self-oscillation of the vocal folds *Am. J. Phys.* **74** 386–93
Gurevich M I 1965 *Theory of Jets in Ideal Fluids* (New York: Academic)
Hofmans G C J, Groot G, Ranucci M, Graziani G and Hirschberg A 2003 Unsteady flow through in-vitro models of the glottis *J. Acoust. Soc. Am.* **113** 1658–75
Howe M S 1998 *Acoustics of Fluid–Structure Interactions* (Cambridge: Cambridge University Press)

Howe M S 2007 *Hydrodynamics and Sound* (Cambridge: Cambridge University Press)

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** 367–92

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** 1233–67

Lighthill M J 1952 On sound generated aerodynamically part I: general theory *Proc. R. Soc.* A **211** 564–87

Park J B and Mongeau L 2007 Instantaneous orifice discharge coefficient of a physical, driven model of the human larynx *J. Acoust. Soc. Am.* **121** 442–55

Lord Rayleigh 1945 *Theory of Sound* vol 2 (New York: Dover))

Stevens K N 1998 *Acoustic Phonetics* (Cambridge, MA: MIT Press)

Titze I R 1994 *Principles of Voice Production* (Upper Saddle River, NJ: Prentice-Hall)

Zanartu M, Mongeau L and Wodicka G R 2007 Influence of acoustic loading on an effective single mass model of the vocal folds *J. Acoust. Soc. Am.* **121** 1119–29

Zhao W, Zhang C, Frankel S H and Mongeau L 2002 Computational aeroacoustics of phonation, part I: computational methods and sound generation mechanisms *J. Acoust. Soc. Am.* **112** 2134–46