Smoothness of an equation for the glottal flow rate versus the glottal area (L)

Jorge C. Lucero[a)]
Department of Computer Science, University of Brasilia, Brasilia, Federal District, 70910-900, Brazil

Jean Schoentgen
Laboratories of Image, Signal Processing and Acoustics, Faculty of Applied Sciences, Université Libre de Bruxelles, 50, Avenue Franklin D. Roosevelt, B-1050, Brussels, Belgium

(Received 24 January 2015; revised 24 March 2015; accepted 27 March 2015)

This letter proposes a correction to an equation by Titze [J. Acoust. Soc. Am. **75**, 570–580 (1984)] for the volume velocity of the glottal air flow given the glottal area and other laryngeal parameters. It shows that the equation produces non-differentiable waveforms at the instants of glottal closure and opening, if the glottal area is also not differentiable at those instants. By adding an air viscosity term to the equation, twice-differentiability is obtained. Also, the letter corrects a sign error in the original formulation. © *2015 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4919297]

[DAB] Pages: 2970–2973

I. INTRODUCTION

Numerical simulations of voice production routinely involve computing the volume velocity of the glottal air flow given the cross-sectional glottal area and other laryngeal parameters. A simple and widely used equation has been given by Titze (1984),

$$u_g = \frac{a_g c}{k_t}\left\{-\frac{a_g}{A^*} \pm \left[\left(\frac{a_g}{A^*}\right)^2 + \frac{4k_t}{c^2\rho}\left(p_s^+ - p_e^-\right)\right]^{1/2}\right\}, \quad (1)$$

where u_g is the flow volume velocity, a_g is the glottal area, c is the speed of sound, k_t is a transglottal pressure coefficient, ρ is the air density, p_s^+ and p_e^- are the incident pressure waves to the glottis coming from the subglottal and supraglottal vocal tracts, respectively, and A^* is an effective vocal tract area given by $1/A^* = 1/A_s + 1/A_e$, where A_s and A_e are the entry areas to the subglottal and supraglottal vocal tracts, respectively. According to Titze (1984), the plus sign for the term in brackets corresponds to $p_s^+ - p_e^- \geq 0$, and the negative sign corresponds to $p_s^+ - p_e^- \leq 0$.

From the computed glottal flow rate, reflected components of the subglottal and supraglottal pressure waves are obtained, and these are used as inputs to wave reflection analogs of the subglottal and supraglottal vocal tracts, respectively.

A drawback of Titze's equation, the remediation of which is the object of this letter, is that the derivative du_g/da_g does not tend to zero when $a_g \to 0$. This letter also shows that there is a sign error in the equation for the case $p_s^+ - p_e^- \leq 0$, the correct form of which is

$$u_g = \pm\frac{a_g c}{k_t}\left\{-\frac{a_g}{A^*} + \left[\left(\frac{a_g}{A^*}\right)^2 + \frac{4k_t}{c^2\rho}\left|p_s^+ - p_e^-\right|\right]^{1/2}\right\} \quad (2)$$

(note the absolute value for the difference $|p_s^+ - p_e^-|$), with the same sign convention as above.

The case of $p_s^+ - p_e^- \leq 0$ implies a negative glottal flow and should be rare during normal phonation. However, simulations of voice production performed as reported by Lucero *et al.* (2013) show brief periods of negative flow at voice offset, when the lung pressure is reduced to zero. Therefore, the sign correction is required in order to obtain a valid formula of general application.

Regarding the time derivative, a quick calculation shows that

$$\lim_{a_g \to 0} \frac{du_g}{da_g} = \pm 2\sqrt{\frac{|p_s^+ - p_e^-|}{k_t\rho}}, \quad (3)$$

which, in general, is not zero (except in the particular case of $p_s^+ = p_e^-$). Therefore, if the glottal area waveform is not differentiable (smooth) at the time of glottal closure (i.e., when the opposite vocal folds collide and close the glottis with the consequent interruption of the airflow), then neither will be the computed glottal flow. This result might pose a problem for the application of Eq. (2) for voice synthesis. It is well known that non-smoothness increases the energy content of higher harmonics, which results in artificial timbres of the synthesized sound.

Suppose that $t_c \leq t \leq t_o$ is the time interval in which the glottis is closed. Then, $a_g(t) = 0$ and $u_g(t) = 0$ for $t_c \leq t \leq t_o$, and $da_g/dt(t) = 0$ and $du_g/dt(t) = 0$ for $t_c < t < t_o$. Differentiability of $u_g(t)$ at t_c demands

$$\lim_{t \to t_c^-} \frac{du_g}{dt} = \lim_{t \to t_c^+} \frac{du_g}{dt} = 0. \quad (4)$$

Assume, for the sake of simplicity, that the glottal area is the only time-varying parameter in Eq. (2). Then

$$\frac{du_g}{dt} = \left(\frac{du_g}{da_g}\right)\left(\frac{da_g}{dt}\right). \quad (5)$$

[a)]Electronic mail: lucero@unb.br

Therefore, if $\lim_{a_g \to 0^+}(du_g/da_g) \neq 0$, differentiability of the glottal flow demands $\lim_{t \to t_c^-} da_g/dt = 0$, and so the glottal area must be differentiable at $t = t_c$. A similar conclusion is obtained for the instant of glottal opening $t = t_o$.

In fact, popular models of the vocal fold oscillation produce non-differentiable glottal area waveforms. For example, in the Phase Delayed Overlapping Sinusoid (PDOS) and Truncated and Exponentially Raised Sinusoid (TERS) models (Titze and Alipour, 2006) of the vocal fold kinematics and also in some lumped representations in terms of mass-damper-spring oscillators (e.g., Avanzini, 2008; Steinecke and Herzel, 1995; Story and Titze, 1995), the glottal area is first computed from the vocal fold displacement and next truncated at $a_g = 0$ or at some small value $\varepsilon > 0$. At the instants of contact onset and release, the glottal area is non-differentiable.

Smooth results may be obtained by producing a gradual area transition between an open and a closed glottis, at the expense of a complexity increase of the models. For example, a zip-like closure and opening may be simulated by considering a triangular geometry for the glottal channel (Birkholz *et al.*, 2011; Titze, 1984). In the two-mass model of the vocal folds, a gradual closure has been simulated by assuming that mechanical contact between the folds begins when the glottal area decreases below a positive threshold value, i.e., before a full glottal closure (Pelorson *et al.*, 1994). A simpler phenomenological approach is to "round" the corners of the glottal area waveform at the times of glottal closure and opening by applying shape transformations (Schoentgen and Lucero, 2013). However, an equally simple but physically based solution might be more desirable.

This letter proposes a third alternative: inclusion of an air viscosity term in Eq. (2). The rationale is that, when the glottis is narrow and almost closed, the viscosity of the air becomes a dominant factor (Pelorson *et al.*, 1994). In such a situation, the pressure drop ΔP along the glottal channel is commonly approximated by the Poiseuille formula for flow between parallel plates,

$$\Delta P = \frac{12\mu L^2 T u_g}{a_g^3}, \tag{6}$$

where μ is the air viscosity coefficient, L is the glottal length in the anterior-posterior direction, and T is the glottal depth in the direction of the airflow (van den Berg *et al.*, 1957). Equation (6) produces the sought $\lim_{a_g \to 0^+}(du_g/da_g) = 0$, and Sec. II shows how to incorporate it into the glottal flow computation.

We note that more complex models of flow viscosity have been proposed, accounting for glottal channels of arbitrary shapes, unsteady flow and other effects (e.g., Deverge *et al.*, 2003; Howe and McGowan, 2010; Krane and Wei, 2006). In principle, any of those models would achieve the desired smoothing. However, experimental assessments have found that the Poiseuille model provides a good prediction of observed flow behavior (Deverge *et al.*, 2003). At the same time, the simplicity of the model allows for its inclusion as a small correction to Titze's equation.

II. CORRECTION TO THE GLOTTAL FLOW EQUATION

We start with a modified form of Bernoulli's equation for the glottis,

$$P_s - P_e = \frac{k_t \rho}{2} \frac{|u_g| u_g}{a_g^2} + \gamma \frac{u_g}{a_g^3}, \tag{7}$$

where P_s and P_e are the subglottal and supraglottal air pressures, respectively, and $\gamma = 12\mu L^2 T$. Equation (7) is the expression used by Titze (1984) to derive his flow equation, to which an air viscosity term has been added (following van den Berg *et al.*, 1957). The absolute value in the first term is required to obtain a negative volume velocity when $P_s < P_e$.

Next, the steps outlined by Titze (1984) are retraced. The subglottal and supraglottal air pressures are decomposed into forward and backward components,

$$P_s = p_s^+ + p_s^-, \quad P_e = p_e^+ + p_e^-. \tag{8}$$

The reflected components at the glottis are

$$p_s^- = -\frac{\rho c u_g}{A_s} + r_s p_s^+, \tag{9}$$

$$p_e^+ = \frac{\rho c u_g}{A_e} + r_e p_e^-, \tag{10}$$

where r_s and r_e are subglottal and supraglottal reflection coefficients, respectively. In his derivation, Titze (1984) considered $r_s = r_e = 1$; however, the reflections coefficients appear as explicit parameters in updated formulations of the theory (Titze and Worley, 2009).

Substituting Eqs. (8)–(10) to into Eq. (7) yields

$$\frac{k_t \rho}{2} \frac{|u_g| u_g}{a_g^2} + \left(\frac{\rho c}{A^*} + \frac{\gamma}{a_g^3} \right) u_g - \delta_p = 0, \tag{11}$$

where we have used $\delta_p = [(1 + r_s)p_s^+ - (1 + r_e)p_e^-]$ to simplify the notation.

Equation (11) has the form

$$c_2 |u_g| u_g + c_1 u_g - c_0 = 0, \tag{12}$$

with $c_1, c_2 > 0$. Assuming $u_g \geq 0$, the equation becomes $c_2 u_g^2 + c_1 u_g - c_0 = 0$ with the solutions

$$u_g = \frac{-c_1 \pm \sqrt{c_1^2 + 4c_0 c_2}}{2c_2}. \tag{13}$$

However, the condition $u_g \geq 0$ is satisfied only when $c_0 \geq 0$ and the square root is taken with the positive sign.

Next, consider the case $u_g < 0$. Equation (12) takes the form $-c_2 u_g^2 + c_1 u_g - c_0 = 0$, with the solutions

$$u_g = \frac{c_1 \pm \sqrt{c_1^2 - 4c_0 c_2}}{2c_2}. \tag{14}$$

The condition $u_g < 0$ is satisfied only when $c_0 < 0$ and the square root is taken with the negative sign.

Combining both cases, we conclude that Eq. (12) has the solutions

$$u_g = \pm\left(\frac{-c_1 + \sqrt{c_1^2 + 4|c_0|c_2}}{2c_2}\right), \tag{15}$$

where the positive sign corresponds to $c_0 \geq 0$ and the negative sign corresponds to $c_0 < 0$.

Applying the above formula to Eq. (11) and rearranging terms, we obtain

$$u_g = \pm\frac{a_g c}{k_t}\left\{-\frac{a_g}{A^*} + \frac{\gamma}{\rho c a_g^2} + \left[\left(\frac{a_g}{A^*} + \frac{\gamma}{\rho c a_g^2}\right)^2 + \frac{2k_t}{c^2\rho}|\delta_p|\right]^{1/2}\right\}. \tag{16}$$

Note that this equation, with $\gamma = 0$ and $\delta_p = 2(p_s^+ - p_e^-)$ (i.e., $r_s = r_e = 1$), reduces to Eq. (2), with a sign correction for Titze's original Eq. (1).

Nevertheless, Eq. (16) poses a new problem: division by a small number when $a_g \to 0^+$ (in the terms with coefficient γ), which causes numerical instability. To solve Eq. (15) numerically, it is multiplied and divided by $c_1 + \sqrt{c_1^2 + 4|c_0|c_2}$, which results in

$$u_g = \frac{\pm 2|c_o|}{c_1 + \sqrt{c_1^2 + 4|c_0|c_2}}. \tag{17}$$

Noting further that the positive sign applies when $c_0 \geq 0$ and the negative sign applies when $c_0 < 0$, then the above equation simplifies to

$$u_g = \frac{2c_o}{c_1 + \sqrt{c_1^2 + 4|c_0|c_2}}. \tag{18}$$

Again, applying the new formula to Eq. (11) and rearranging terms, we obtain

$$u_g = \frac{2a_g^3\delta_p}{\frac{\rho c a_g^3}{A^*} + \gamma + \left[\left(\frac{\rho c a_g^3}{A^*} + \gamma\right)^2 + 2k_t\rho a_g^4|\delta_p|\right]^{1/2}}, \tag{19}$$

where division by a small value of a_g is avoided. Note also that, since $\gamma > 0$, the denominator is never 0 (its minimum value is 2γ).

A quick calculation shows that, when $a_g \to 0^+$, $u_g \to a_g^3\delta_p/\gamma$, which is the Poiseuille formula for viscous flow, and $du_g/da_g \to 0$. Further, $d^2u_g/da_g^2 \to 0$.

III. EXAMPLE

Let us illustrate the previous results with an example. A glottal area waveform with amplitude of $0.15\,\text{cm}^2$, frequency of 100 Hz, open quotient of 0.6 and skewing to the right was simulated with the TERS model (Titze and Alipour, 2006),

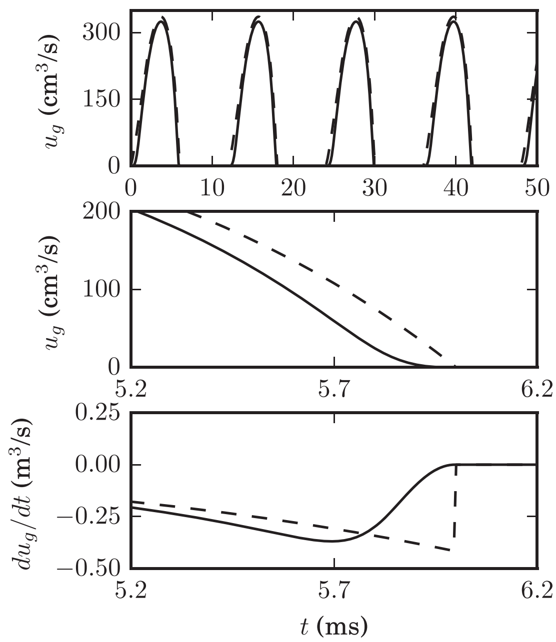

FIG. 1. (Top) Volume velocity of the glottal flow vs time. (middle) Expanded view of the volume velocity around the time of glottal closure. (Bottom) First derivative of the volume velocity. Full line shows the results from Eq. (19); dashed line shows the results from Eq. (2).

$$a_g(t) = 0.15\max\left[0, \left(1 + \frac{500t}{0.6}\right)\sin\left(\frac{100\pi t}{0.6}\right)\right]. \tag{20}$$

Next, the glottal flow rate was computed using Eqs. (2) and (19) with $L = 1.6$ cm, $T = 0.35$ mm, $k_t = 1$, $A_s = A_e = 3\ \text{cm}^2$, $c = 35\,000$ cm/s, $\rho = 0.00114\ \text{g/cm}^3$, $\nu = 0.000\,186$ g/(cm s) and a constant $\delta_p = 800$ Pa.

Results are shown in Fig. 1. Clearly, Eq. (19) produces a smooth waveform for the glottal flow rate, first derivative included, while Eq. (2) produces a non-differentiable waveform at the instant of glottal closure. At the same time, there is no relevant effect on the general shape of the glottal pulse, except for an amplitude reduction of 3.3%. The effect of the correction on the flow spectrum may be assessed by means of the spectral ratio $\text{SR} = 10\log(E_{\text{low}}/E_{\text{high}})$, where E_{low} and E_{high} are the energy contents below and above 1 kHz, respectively (Kitzing, 1986). Another useful measure is the spectral balance (SB), defined as the critical frequency which divides the spectrum into low and high frequency regions of equal energy. For the waveforms in the figure and a sampling frequency of 20 kHz, Eq. (2) produces SR = −3.5 dB and SB = 2295 Hz, whereas Eq. (19) produces SR = −3.2 dB and SB = 2138 Hz. The spectral ratio and balance decrease by approximately 7%, which indicates a clear reduction of the high frequency content in the new model and is of the same order of magnitude as the differences reported between strained and lax voices in long-term average spectra (Kitzing, 1986).

IV. CONCLUSION

In summary, this letter proposes Eq. (19) as a replacement for Eq. (1) from Titze (1984). By incorporating an air viscosity term, the new equation produces a smooth volume velocity of the glottal flow that is twice-differentiable at the instants of glottal closure and opening, even when the glottal area waveform is not smooth at those instants. The practical advantage of the proposed equation is that it allows for the use of simple vocal fold models which truncate the glottal area waveform at glottal closure.

In studies using the two-mass model of the vocal folds, increased smoothness of the glottal flow has been obtained by adding a flow component produced by deformation of the vocal folds during their collision (Pelorson *et al.*, 1994), and also by considering an incomplete glottal closure because of the presence of a posterior glottal opening (Zañartu *et al.*, 2014). If desired, both effects may be incorporated also to the above flow model. Note, however, that Eq. (19) allows for a simple adjustment of the degree of smoothing by using coefficient γ as a control parameter.

If air viscosity is neglected by letting $\gamma \approx 0$, then the denominator of Eq. (19) takes small values when $a_g \rightarrow 0$ causing numerical instability. In that case, Eq. (2) should be used instead, which incorporates a sign correction to Titze's original equation.

It must be noted also that our analysis has not considered the effect of a moving separation point of the air flow from the walls of the glottal surface. In fact, the glottal area a_g in Eqs. (2) and (19) must be computed at the point of airflow separation, and the location of that point moves during the oscillatory cycle (e.g., Pelorson *et al.*, 1994). Simplified representations of such movement may also cause discontinuities in the waveform of the glottal flow (e.g., Howe and McGowan, 2010).

ACKNOWLEDGMENTS

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior of Brazil.

Avanzini, F. (**2008**). "Simulation of vocal fold oscillation with a pseudo-one-mass physical model," Speech Commun. **50**, 95–108.

Birkholz, P., Kröger, B. J., and Neuscheafer-Rube, C. (**2011**). "Synthesis of breathy, normal, and pressed phonation using a two-mass model with a triangular glottis," in *Interspeech 2011*, Florence, Italy, pp. 2681–2684.

Deverge, M., Pelorson, P., Vilain, C., Lagrée, P.-Y., Chentouf, F., Willems, J., and Hirschberg, A. (**2003**). "Influence of collision on the flow through in-vitro rigid models of the vocal folds," J. Acoust. Soc. Am. **114**, 3354–3362.

Howe, M. S., and McGowan, R. S. (**2010**). "On the single-mass model of the vocal folds," Fluid Dyn. Res. **42**, 015001.

Kitzing, P. (**1986**). "LTAS criteria pertinent to the measurement of voice quality," J. Phonetics **14**, 477–482.

Krane, M. H., and Wei, T. (**2006**). "Theoretical assessment of unsteady effects in phonatory aerodynamics," J. Acoust. Soc. Am. **120**, 1578–1588.

Lucero, J. C., Schoentgen, J., and Behlau, M. (**2013**). "Physics-based synthesis of disordered voices," in *Interspeech 2013*, Lyon, France, pp. 587–591.

Pelorson, X., Hirschberg, A., van Hassel, R. R., Wijnands, A. P. J., and Auregan, Y. (**1994**). "Theoretical-study and experimental-study of quasi-steady-flow separation within the glottis during phonation—Application to a modified 2-mass model," J. Acoust. Soc. Am. **96**, 3416–3431.

Schoentgen, J., and Lucero, J. C. (**2013**). "Synthesis by rule of disordered voices," in *Advances in Nonlinear Speech Processing*, edited by T. Drugman and T. Dutoit (Springer-Verlag, Berlin), pp. 120–127.

Steinecke, I., and Herzel, H. (**1995**). "Bifurcations in an asymmetric vocal-fold model," J. Acoust. Soc. Am. **97**, 1874–1884.

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.

Titze, I. R. (**1984**). "Parametrization of the glottal area, glottal flow and vocal fold contact area," J. Acoust. Soc. Am. **75**, 570–580.

Titze, I. R., and Alipour, F. (**2006**). *The Myoelastic Aerodynamic Theory of Speech Production* (National Center for Voice and Speech, Iowa City, IA), Chap. 5, pp. 258–262.

Titze, I. R., and Worley, A. S. (**2009**). "Modeling source-filter interaction in belting and high-pitched operatic male singing," J. Acoust. Soc. Am. **126**, 1530–1540.

van den Berg, J. W., Zantema, J. T., and Doornenbal, P. (**1957**). "On the air resistance and the Bernoulli effect of the human larynx," J. Acoust. Soc. Am. **29**, 626–631.

Zañartu, M., Galindo, G. G., Erath, B. D., Peterson, S. D., Wodicka, G. R., and Hillman, R. E. (**2014**). "Modeling the effect of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction," J. Acoust. Soc. Am. **136**, 3262–3271.