ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 4aSCb: Voice and F0 Across Tasks (Poster Session)

4aSCb12. Modeling vocal fold asymmetries with coupled van der Pol oscillators
Jorge C. Lucero* and Jean Schoentgen

*Corresponding author's address: University of Brasilia, Brasilia, 70910-900, DF, Brazil, lucero@unb.br

Models of the glottal sound source are being developed to extend a recent synthesizer of disordered voices [Fraj et al., J. Acoust. Soc. Am. 132, 2603-2615 (2012)]. The synthesizer was based on a nonlinear wave-shaping algorithm which generates a glottal excitation to a concatenated-tube representation of the trachea and vocal tract. The purpose of the present work is to incorporate a physics-based model of the vibrating vocal folds in order to increase the anatomical fidelity of the synthesizer. Further, the model will permit to characterize left-right fold asymmetries and explore the effect of those asymmetries on the resultant vocal timbre. In this report, the vocal folds are represented as a system of two coupled Van der Pol oscillators with noise terms and a detuning factor between their natural frequencies. Regions of phase locked and unlocked oscillations are determined and illustrated with bifurcation diagrams. Also, the effect of frequency detuning on the resultant frequency jitter is analyzed. The results are discussed in terms of their implications for modeling abnormal vocal fold behavior. [Work supported by CNPq (Brazil) and FNRS (Belgium)].

Published by the Acoustical Society of America through the American Institute of Physics
INTRODUCTION

Recent studies have shown that the vocal fold oscillator during phonation may be modeled as van der Pol oscillator (Garrel et al., 2008; Laje et al., 2001; Lucero et al., 2011). Since its introduction almost a century ago (van der Pol, 1920), this well-known mathematical model has been applied to analyze nonlinear oscillations in a wide variety of systems. In case of human physiology, the model has been used to study from the heart beat (van der Pol and van der Mark, 1928) to Parkinsonian tremor (Beuter et al., 2003) and EEG dynamics (Kirschfeld, 2005). It has also been successfully applied to study the vocalization of songbirds (e.g., Laje and Mindlin, 2005), which shares similar principles than human vocalization.

The simplicity of the model helps the analysis of the basic mechanisms of phonation. Basically, it is a single mass-damper-spring oscillator; however, it also incorporates the transfer of energy from the airflow to the vocal folds due to the out-of-phase motion of the entry and exit glottal edges. In this way, it captures much of the oscillatory behavior of the popular and more complex two-mass model (Ishizaka and Flanagan, 1972).

Here, the model will be used to analyze the effect of vocal fold asymmetries on the oscillation. Several voice disorders are associated to desynchronization between the left and right folds, caused by a tissue tension or mass asymmetry, which may generate subharmonics, biphonation and irregular oscillations. Previous works have analyzed this phenomenon by using multi-mass models (e.g., Ishizaka and Isshiki, 1976; Steinecke and Herzl, 1995). The adoption of a simpler representation, as the van der Pol oscillator, might provide a clearer picture. Two aspects will be considered here: First, the effect of a tension imbalance will be analyzed, simulating, e.g., a unilateral laryngeal nerve paralysis. The purpose is to determine regions of synchronized (phase-locked) and desynchronized behavior as a function of the level of asymmetry. Next, the effect of the asymmetry on cycle-length jitter will be explored. Jitter denotes random cycle-to-cycle variations of the oscillation period caused by external perturbations to the vocal folds, and it appears in all voice signals. It has been shown that laryngeal pathologies increase the size of jitter. However, such increase may not necessarily be the result of a larger external perturbation, but also consequence of the asymmetry between the left and right oscillators, owing to the pathology (Schoentgen, 2001).

This work fits into the framework of the development of a synthesizer of disordered voices (Fraj et al., 2012). A previous version was based on a nonlinear wave-shaping algorithm which generates a glottal excitation to a concatenated-tube representation of the trachea and vocal tract. As a next step, the anatomical fidelity of the synthesizer will be increased by incorporating a physics-based model of the vibrating vocal folds. The present analysis will serve to test the suitability of the van der Pol oscillator for such a role.

VOCAL FOLD MODEL

The oscillatory motion of the vocal fold tissues is characterized as a surface wave that propagates in the direction of the airflow (Titze, 1988). Motion of the right vocal fold is described by the equation

$$M_r \ddot{\xi}_r + B_r (1 + \eta_r \dot{\xi}_r^2) \dot{\xi}_r + K_r \xi_r = P_g,$$

where $\xi_r$ is the tissue displacement at the midpoint of the glottis, $M_r$, $B_r$ and $K_r$ are the mass, damping and stiffness, respectively, per unit area of the vocal fold medial surface, $\eta_r$ is a nonlinear damping coefficient and $P_g$ is the mean air pressure at the glottis (Laje et al., 2001). A similar equation describes motion of the left fold.

Ignoring the effect of the sub- and supraglottal vocal tracts, the glottal mean air pressure is

$$P_g = \left( \frac{P_s}{k_1} \right) \frac{a_1 - a_2}{a_1}, \quad (a_1, a_2 > 0),$$

J. Lucero and J. Schoentgen
where $P_s$ is the subglottal pressure, $k_t$ is a transglottal pressure coefficient, and $a_1$ and $a_2$ are the cross-sectional glottal areas at the lower and upper edges of the vocal folds, respectively (Titze, 1988). The glottal areas are given by

$$a_{1,2}(t) = L\left[\xi_0 + \xi_{r,(l)}(t \pm \tau_{r,(l)})\right] + L\left[\xi_0 + \xi_{l,(r)}(t \pm \tau_{l,(r)})\right],$$

where $L$ is the vocal fold length, $\xi_0$ is half the glottal width when the the vocal folds are at rest, and $\tau_{r,l}$ is the time delay for the surface wave to travel half the glottal height $T$. The positive and negative signs correspond to $a_1$ and $a_2$, respectively.

In case of small displacements and small wave delay, Eq. (2) may be approximated by

$$P_g = \frac{P_s}{k_t\xi_0}(\tau_r\dot{\xi}_r + \tau_l\dot{\xi}_l).$$

Here, we will consider the effect of a vocal tension imbalance, represented as an asymmetry in the stiffness coefficient. In such a case, dropping the $r$ or $l$ sub-index of all parameters except the stiffness coefficient, letting $C = P_s\tau/(k_t\xi_0)$ and assuming $P_s$ large enough so that $C \geq B/2$, applying the change of variable $x_{1,2} = \sqrt{\eta B/(2C-B)}\xi_{1,2}$, and changing the time scale $t \rightarrow (\sqrt{M/K_r})t$, produces the system of two coupled van der Pol oscillators

$$\ddot{x}_r - \mu(1 - x_r^2)x_r + x_r = a(\dot{x}_l - \dot{x}_r),$$

$$\ddot{x}_l - \mu(1 - x_l^2)x_l + Qx_l = a(\dot{x}_r - \dot{x}_l),$$

where $\mu = (2C - B)/\sqrt{MK_r}$, $a = C/\sqrt{MK_r} \geq 0$ is a coupling parameter and $Q = K_l/K_r \geq 0$ is an asymmetry parameter.

When $Q = 1$, both oscillators are identical. For a stable limit-cycle oscillations, $\mu$ must be positive, which produces the condition $C > B/2$ required in the transformation above. The threshold condition $\mu = 0$ yields $P_s = k_t\xi_0B/(2\tau)$, which is the so-called phonation threshold pressure. Note also that $\mu \geq 0$ implies $a \geq \mu/2$.

To have an idea of parameter values, let us consider a normal adult male configuration (Titze, 1988) with $M = 0.5$ g/cm$^2$, $B = 50$ dyne s/cm$^3$, $K = 200000$ dyne/cm$^3$, $\tau = 1$ ms, $k_t = 1.1$, $\xi_0 = 0.1$ cm, $P_s = 800$ Pa, which produce $\mu = 0.30$ and $a = 0.23$. For the asymmetry parameter, we only need to consider the case $0 \leq Q \leq 1$. If $Q > 1$, then $K_l > K_r$. In that case, a simple change of time scale will rewrite the above system with an equivalent asymmetry parameter $0 \leq Q \leq 1$ applied to the right oscillator instead of the left one. This range of values for $Q$ (applied to the left fold) means that the left vocal fold is less stiff than the right, as in, e.g., a paralysis affecting the left fold. The smaller $Q$, the more severe the pathology.
Let us look for phase-locked (synchronized) oscillatory solutions, in which the right and left vocal folds oscillate with the same frequency and constant phase difference. We apply a harmonic balance method (Mickens, 1996), which consists of approximating the solution of the differential equations by a truncated Fourier series. Thus, we assume the first-order solution

\[ x_r(t) = R_r \cos(\omega t - \theta_r) \]  \hspace{1cm} (7)
\[ x_l(t) = R_l \cos(\omega t - \theta_l) \]  \hspace{1cm} (8)

Replacing into Eqs. (5) and (6), ignoring the higher harmonic terms and solving, we obtain, after some algebra,

\[ \omega^2 = \frac{(1 + Q)}{2} \]  \hspace{1cm} (9)
\[ \sin \phi = -\frac{(1 - Q)}{(2\omega\alpha)} \]  \hspace{1cm} (10)
\[ \mu(1 - R^2/4) = \alpha(1 - \cos \phi) \]  \hspace{1cm} (11)

where \( \phi = \theta_r - \theta_l \) and \( R = R_r = R_l \).

According to Eq. (9), the oscillation frequency decreases as the asymmetry increases, from its maximum value of \( \omega = 1 \) to \( \omega \to (\sqrt{2})/2 \) for \( Q \to 0 \). According to Eq. (10), \(-\pi \leq \phi \leq 0\); therefore, the right (healthy) side always precedes the left (flaccid) side. Both results agree with previous studies on asymmetric vocal folds (Steinecke and Herzel, 1995).

For given values of \( \alpha \) and \( Q \), the above equations may have none, one or two solutions.

Changes in the number of solutions (bifurcations) occur at

\[ Q = 1 - 2\omega\alpha \]  \hspace{1cm} (12)

and

\[ \alpha(1 - \cos \phi) = \mu \]  \hspace{1cm} (13)

The stability of the (unique) equilibrium position at \( x_r = x_l = 0 \) is given by the roots of the associated characteristic equation

\[ s^4 + 2(\alpha - \mu)s^3 + [1 + Q + \mu(\mu - 2\alpha)]s^2 + (\alpha - \mu)(1 + Q)s + Q = 0 \]  \hspace{1cm} (14)

Application of the standard Routh-Hurwitz criterion produces two conditions: one is Eq. (13), at which a pair of roots cross the imaginary axis (Hopf bifurcation). Another one is

\[ \alpha = \mu \]  \hspace{1cm} (15)

at which two pairs of roots cross the imaginary axis.

The above set of conditions lead to the bifurcation diagram in Fig. 2. There are three regions, associated to phase-locked oscillations (dark gray), unlocked oscillations (light gray), and no oscillations (white). When \( Q = 1 \), both folds oscillate with frequency \( \omega = 1 \), phase difference \( \phi = 0 \) and amplitude \( R \approx 2 \). As \( Q \) decreases, the phase difference increases in absolute value, and reaches the maximum value of \( |\phi| = \pi/2 \) in curve \( b \). At the same time, both the oscillation amplitude and frequency decrease. Fig. 3 shows an example of the variation of the oscillation parameters with \( Q \). In the Figure, point \( B \) corresponds to a value of \( Q \approx 1 - \mu = 0.7 \). For higher values of \( Q \), any value of \( \alpha \) above oscillation threshold produces a phase-locked oscillation. For smaller values of \( Q \), the coupling must be increased to obtain phase-locking.

Fig. 4 shows an example of unlocked oscillations. The frequencies of the right and left oscillators fluctuate in the range 0.90–1 and 0.60–0.85, respectively.
**FIGURE 2:** Bifurcation diagram for $\mu = 0.3$. Full lines $a$, $b$ and $c$ are associated to Eqs. (12), (13) and (15). A stable locked-phase oscillation exists in the blue region above curve $a$, at the left of point $A$, and above curve $b$, at the right of point $A$. In the dark blue region between curves $a$ and $b$, an additional locked-phase oscillation exists, but it is unstable. In the red region below line $c$, at the left of point $A$, and below curve $b$, at the right of point $A$, the oscillators are unlocked. Black line $d$ indicates the oscillation threshold $\alpha = \mu/2$, and marks a lower limit of validity of the van der Pol model for the vocal folds.

**FIGURE 3:** Oscillation amplitude (top panel), phase difference between oscillators (middle panel) and oscillation frequency (bottom panel) vs. $Q$, for $\mu = 0.3$ and $\alpha = 0.4$. The stars indicate results obtained by direct numerical solution of Eqs. (5) and (6).

**EFFECT OF THE ASYMMETRY ON JITTER**

Random perturbations to the stiffness are introduced in the form

$$\ddot{x}_r - \mu(1 - x_r^2)\dot{x}_r + (1 + \alpha\epsilon)x_r = \alpha(\dot{x}_l - \dot{x}_r),$$

(16)
FIGURE 4: Displacement of the right (blue) and left (red) oscillators for $Q = 0.5$, $\mu = 0.3$ and $a = 0.22$.

$$\ddot{x}_l - \mu(1-x_l^2)x_l + Q(1+\alpha)\dot{x}_l = a(\dot{x}_r - \dot{x}_l),$$ (17)

where $a$ is a small scaling coefficient and $\varepsilon$ is a white noise term.

The above stochastic differential equations were solved on the time interval $0 \leq t \leq 1000$ by applying a standard Euler-Maruyama algorithm (Higham, 2001), with a time step $h = 0.001$. White noise was simulated as

$$\varepsilon = \sqrt{h} \times \begin{cases} +1, & p = 0.5 \\ -1, & p = 0.5 \end{cases}$$ (Schoentgen, 2001). The scaling parameter was set at $a = 0.05$, which produces jitter values around 1%. From the solutions, cycle lengths of the glottal area, represented by the sum $x_r(t) + x_l(t)$, were computed by using a high-precision zero-crossing algorithm (Titze and Liang, 1993). Finally, jitter was computed as the mean of the absolute cycle-to-cycle length difference over the last 100 cycles, relative to their mean.

Fig. 5 show results when varying $Q$ and $\alpha$, while keeping a constant level of the stiffness perturbation. The fast increase at the left end of the curves occurs when the system reaches the boundaries of the region of phase-locked behavior. Outside that region, the oscillations desynchronize and become toroidal and therefore, jitter measures are no longer valid. Within the phase-locked region, jitter increases when $Q$ decreases (larger left-right asymmetry). Also, it increases when the coupling parameter $\alpha$ is decreased. This variation of jitter with $Q$ and $\alpha$ matches previous results from a kinematic model of the vocal fold oscillation (Schoentgen, 2001).

The dependance of jitter on the coupling $\alpha$ is very small and only appears towards the boundaries of the phase-locked region. In fact, Eq. (9) shows that, up to the first order approximation, the phase-locked frequency (and consequently its period) is a function of $Q$ only.

FIGURE 5: Jitter vs. stiffness asymmetry $Q$, for $\mu = 0.3$, $a = 0.05$ and various values of the coupling $\alpha$. 
CONCLUSIONS

The proposed van der Pol representation of the vocal folds is simple, yet it is useful for a theoretical investigation of the vocal fold dynamics. Previous studies have used it to analyze normal phonation; here, it has been extended to explore the effect of asymmetries in the vocal folds.

The model predicts phase-locked oscillations of the right and left vocal folds for small tension imbalances. When the imbalance is larger, phase-locked oscillations may be obtained by increasing the right-left coupling. The model also shows that phase locking has the effect of attenuating the level of jitter on the glottal area oscillation, even if the amplitude of the perturbations that causes the jitter is kept constant.

For the range of parameters analyzed, only phase-locked or unlocked oscillations were detected. More complex behavior with the presence of subharmonics and irregular oscillations, which may be observed in pathological voices, were not produced. Possibly, the model may be too simple for those complex patterns. In fact, the model is not able to simulate collision between the opposite vocal folds. Collision act as an addictional coupling factor and must have a significant effect on the synchronization of the oscillations. All these aspects will be considered in extensions of the model for further analyses.

ACKNOWLEDGMENTS

This work has been supported by a Scientific Cooperation Program between CNPq (Brazil) and FNRS (Belgium).

REFERENCES


van der Pol, B. and van der Mark, J. (1928). “The heartbeat considered as a relaxation oscillation, and an electrical model of the heart”, Phil. Mag. 6, 763–775.