Biomechanical modeling of voice register transitions

Isao T. Tokuda*
School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa 923-1292, Japan

Marco Zemke and Hanspeter Herzel
Institute for Theoretical Biology, Humboldt University of Berlin, Invalidenstr. 43, D-10115 Berlin, Germany

Biomechanical modeling is carried out to study register transitions of the human voice. Our model has a body-cover structure, which is composed of fourth masses. A smooth geometry is realized by introducing a polygon shape to the vocal fold model. The simulation study shows that the model can reproduce many complex phenomena such as register jumps, hysteresis, subharmonics, and chaos, observed in excised larynx experiment as well as in vocalization of untrained singers. Our two-dimensional bifurcation diagram can be regarded as a generalization of the voice range profile.

Keywords: 4-Mass Model, Body-Cover Structure, Polygon Geometry, Falsetto-Chest Transition, Hysteresis

I. INTRODUCTION

The exact definition of registers in the human voice is still under debate [1, 2]. Especially the quantitative analysis of transitions between the registers have not been investigated in much detail yet. Excised larynx experiments show different kinds of voice instabilities that appear close to the transition from chest to falsetto register [3, 4]. These instabilities include abrupt jumps between the two registers exhibiting hysteresis, aphonic episodes, subharmonics and chaos. To model these phenomena, we started with a three-mass cover model, which was constructed by adding one more mass on top of the two-mass model [5]. Because of the additional mass, the upper part of the vocal folds can produce a small amplitude waveform, which resembles the falsetto register. This falsetto-like register can easily coexist with the chest-like register, giving rise to hysteresis phenomena. Near the register transitions, subharmonics and chaos are observed, which reproduce even details of the excised larynx experiment.

For a deeper understanding of the register transition in human voice, several extensions are indispensable. Introduction of the body-cover structure is important for physiologically more realistic modeling of the larynx [6]. Recent studies also showed that a smooth geometry in vocal folds is important for a precise computation of the aerodynamic force, that can produce distinct registers [7–9]. Therefore, we extend our model to a four-mass body-cover polygon model. We make use of the bifurcation analysis to understand how small changes of parameter values can cause sudden qualitative changes in the dynamical behavior of the larynx.

II. 4-MASS MODEL

Figure 1 shows a schematic illustration of the four-mass polygon model. This model is composed of a body part \( m_b \) and a cover part, which is divided into three masses \( m_i \) (lower: \( i = 1 \), middle: \( i = 2 \), upper: \( i = 3 \)). Our basic modeling assumptions are the following:

1. Cubic nonlinearities of the oscillators are neglected.
2. Influence of vocal tract as well as subglottal resonances are not considered.
3. Additional pressure drop at inlet is neglected; the Bernoulli flow is considered only below the narrowest part of the glottis [10].
4. Symmetric motion between the left and the right vocal folds is assumed.

*isao@jaist.ac.jp
Our model equations read

\[ m_1 \ddot{y}_1 + r_1(y_1 - \dot{y}_b) + k_1(y_1 - y_b) + \Theta(-h_1)c_1(\frac{h_1}{2}) + k_{1,2}(y_1 - y_2) = F_1, \] (1)

\[ m_2 \ddot{y}_2 + r_3(y_2 - \dot{y}_b) + k_2(y_2 - y_2) + \Theta(-h_2)c_2(\frac{h_2}{2}) + k_{1,2}(y_2 - y_1) + k_{2,3}(y_2 - y_3) = F_2, \] (2)

\[ m_3 \ddot{y}_3 + r_3(y_3 - \dot{y}_b) + k_3(y_3 - y_2) + \Theta(-h_3)c_3(\frac{h_3}{2}) + k_{2,3}(y_3 - y_2) = F_3, \] (3)

\[ m_4 \ddot{y}_4 + r_4\dot{y}_b + k_4y_b + r_1(y_b - \dot{y}_1) + k_1(y_b - y_1) + r_2(y_b - \dot{y}_2) + k_2(y_b - y_2) + \]

\[ r_3(y_b - \dot{y}_3) + k_3(y_b - y_3) = 0. \] (4)

The dynamical variables \( y_i \) represent displacements of the masses \( m_i \), where the corresponding glottal opening length is given by \( h_i = h_{0i} + 2y_i \) (\( h_{0i} \): prephonatory length; \( i = 1, 2, 3 \)). The constant parameters \( r_i, k_i, c_i \) represent damping, stiffness, and collision stiffness of the masses \( m_i \), respectively, whereas \( k_{ij} \) represents coupling strength between two masses \( m_i \) and \( m_j \). The stiffness is determined as \( r_i = 2\zeta_i\sqrt{m_i k_i} \) using the damping ratio \( \zeta_i \). The collision function is approximated as \( \Theta(\xi) = 0 \) \((\xi \leq 0)\); \( \Theta(\xi) = 1 \) \((0 < \xi)\).

The aerodynamic force, \( F_1 \), acting on each mass is derived as follows. First, the vocal fold geometry is described by a pair of four mass-less plates as shown in Fig. 1. The flow channel height \( h(x, t) \) is a piecewise linear function, composed of \( h_{1,0} \) \((x_0 \leq x \leq x_1)\), \( h_{2,1} \) \((x_1 \leq x \leq x_2)\), \( h_{3,2} \) \((x_2 \leq x \leq x_3)\), and \( h_{4,3} \) \((x_3 \leq x \leq x_4)\), which are determined as

\[ h_{i,i-1}(x, t) = \frac{h_i(t) - h_{i-1}(t)}{x_i - x_{i-1}}(x - x_{i-1}) + h_{i-1}(t), \] (5)

where \( i = 1, 2, 3, 4 \) and \( h_0 \) and \( h_4 \) are constants. Assuming the Bernoulli flow, the pressure distribution \( P(x, t) \) below the narrowest part of the glottis, \( h_{\text{min}} = \min(h_{1,2}, h_{2,3}) \), is described as

\[ P_s = P(x, t) + \frac{\rho}{2} \left( \frac{U}{h(x,t)} \right)^2 = P_0 + \frac{\rho}{2} \left( \frac{U}{l} \right)^2, \] (6)

where \( \rho \) represents the air density, \( P_s \) is the subglottal pressure, and \( l \) is the length of the glottis. By setting \( P_0 = 0 \), the glottal volume flow velocity is computed as \( U = \sqrt{\frac{2P_s}{\rho}h_{\text{min}}}(\Theta(h_{\text{min}})) \). The aerodynamic forces on the plates are induced by the pressure \( P(x, t) \) along the flow channel. As \( m_1, m_2, m_3 \) support the plates, an
aerodynamic force on point  

\[ F_i(t) = \int_{x_{i-1}}^{x_i} \frac{x - x_{i-1}}{x_i - x_{i-1}} P(x, t) \, dx + \int_{x_i}^{x_{i+1}} \frac{x_{i+1} - x}{x_i - x_{i-1}} P(x, t) \, dx. \]  

This integral can be solved analytically [8] for the pressure distribution  \( P(x, t) \) described by Eq. (6).

As the default situation, the parameters are set as:  
\[ m_1 = 0.006 \text{ g}; \quad m_2 = 0.012 \text{ g}; \quad m_3 = 0.003 \text{ g}; \quad m_b = 0.05 \text{ g}; \quad k_1 = 0.004 \text{ g/\text{ms}^2}; \quad k_2 = 0.008 \text{ g/\text{ms}^2}; \quad k_3 = 0.002 \text{ g/\text{ms}^2}; \quad k_b = 0.03 \text{ g/\text{ms}^2}; \quad k_{1,2} = 0.001 \text{ g/\text{ms}^2}; \quad k_{2,3} = 0.0005 \text{ g/\text{ms}^2}; \]
\[ c_1 = 3k_1; \quad c_2 = 3k_2; \quad c_3 = 3k_3; \quad \zeta_1 = 0.1; \quad \zeta_2 = 0.1; \quad \zeta_3 = 0.1; \quad \zeta_b = 0.1; \quad h_{01} = 0.036 \text{ cm}^2; \quad h_{02} = 0.036 \text{ cm}^2; \quad h_{03} = 0.036 \text{ cm}^2; \quad x_0 = 0 \text{ cm}; \quad x_1 = 0.01 \text{ cm}; \quad x_2 = 0.017 \text{ cm}; \quad x_3 = 0.225 \text{ cm}; \quad x_4 = 0.265 \text{ cm}; \quad l = 1.4 \text{ cm}; \]
\[ g = 0.00113 \text{ g/\text{cm}^3}; \quad P_s = 0.008 \text{ g/cm\text{ms}^2}. \]

The tension parameter  \( Q \) is introduced to control the frequency of the four masses as  
\[ m_i' = m_i/Q; \quad k_i' = k_iQ \quad (i = 1, 2, 3, b). \]

### III. SIMULATIONS

With the default parameter setting with  \( Q = 1.285 \), we observed a chest-like phonation as visualized in Fig. 2 (a). We observe the known phase advance of the lower mass and complete glottal closure leading to the slightly skewed volume flow shown in Fig. 2 (b). This chest-like waveform has a frequency of 133 Hz. If we change the vocal fold geometry to  \( x_0 = 0 \text{ cm}; \quad x_1 = 0.075 \text{ cm}; \quad x_2 = 0.225 \text{ cm}; \quad x_3 = 0.265 \text{ cm}; \quad x_4 = 0.3 \text{ cm} \)
and set the tension parameter to  \( Q = 3.15 \), qualitatively distinct vibration pattern appears. Fig. 2 (c) shows phase-shifted vibrations of the upper two masses, whereas the lowest mass is wide open. The observed frequency of 330 Hz is much higher than the chest-like vibration. Due to the lack of the vocal fold collision, the volume flow waveform in Fig. 2 (d) shows an almost sinusoidal waveform. In this way, falsetto-like vibrations can be simulated.

![Fig. 2](image)

**FIG. 2:** (a), (c) Time series of glottal areas  \( h \) for chest and falsetto registers, respectively. (b), (d): Glottal volume flow  \( U(t) \) corresponding to (a) and (c).

Now we study the transition between the chest and the falsetto registers. By linearly changing all the parameters as  
\[ p = (1 - \lambda)p_{\text{chest}} + \lambda p_{\text{falsetto}} \quad (0 \leq \lambda \leq 1), \]
we have simulated the register transition. Figure 3 shows a 2-dimensional bifurcation diagram of the model in the range of  \( (\lambda, \theta) \in [0, 1] \times [0, 0.03] \). On the right side, the falsetto-like vibrations dominate the bifurcation diagram, whereas on the left side the chest registers are found. Aphonia is located at the bottom. Sudden jumps are found between the chest and falsetto registers accompanied by subharmonics and deterministic chaos. In the range of a relatively small subglottal pressure  \( P_s < 0.007 \), coexistence of chest and falsetto registers is observed. If we increase the transition parameter  \( \lambda \)
starting at zero the chest register persists in a relatively large parameter range, whereas for decreasing \( \lambda \) the falsetto register range is enlarged (the corresponding figure is not shown here). This hysteresis is a typical phenomenon of register changes observed in excised larynx experiment as well as in the voice range profile.

FIG. 3: (a) 2-dimensional bifurcation diagram in the range of \((\lambda, P_s) \in [0, 1] \times [0, 0.03]\). The blue and green regions correspond to the existence domain of chest and falsetto registers, respectively. The orange, red, and dark blue correspond to subharmonics, chaos (or torus), and aphonia. (b) Frequency distribution of the registers observed in (a).
IV. SUMMARY

Our simulations reveal that the proposed 4-mass polygon model can reproduce coexistence of chest and falsetto registers as well as complex transitions between them as observed in excised larynx experiment as well as in vocalization of untrained singers. The two-dimensional bifurcation analysis is a valuable tool to understand voice registers, which can be considered as a generalization of the voice range profile.

Our studies will take into account the effect of vocal tract as well as subglottal resonances on the voice registers. Rules shall be developed to associate muscle activities with the parameter configurations [11]. The bifurcation diagram will be compared with experimental measurements of register transitions in trained and untrained singers.

Acknowledgments

This work was supported by SCOPE (071705001) of Ministry of Internal Affairs and Communications (MIC), Japan.