Modelling Bird Songs: Voice Onset, Overtones and Registers

Riccardo Zaccarelli¹, Coen P.H. Elemans², W. Tecumseh Fitch³, Hanspeter Herzel¹

¹: Institute for Theoretical Biology (ITB), Humboldt University, Invalidenstr. 43, D-10115 Berlin, Germany
²: Experimental Zoology Group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, The Netherlands
³: School of Psychology, University of St. Andrews, St. Andrews, Fife, KY16 9AJ, Scotland

Summary
We analyze two symmetric two-mass models of the avian syrinx. Our first model applies to songbirds and is a rescaled version of the well-known human two-mass model. Our second model (trapezoidal model) introduces a smoother geometry and is used to simulate the ring dove (Streptopelia risoria) syrinx. Simulations show that both models exhibit self-sustained vibrations. We show that the occurrence of collisions and the intensity of harmonics depend strongly on the configuration of the syrinx. The songbird model does not present instabilities. The trapezoidal model, however, displays coexisting limit-cycles that represent vibrations with, and without collisions at the same pressure. Register-like transitions are accompanied by subharmonics and deterministic chaos.

PACS no. 05.45.-a, 43.60.+d, 43.64.+r, 43.70.+i

1. Introduction

Two-mass models of mammalian vocal fold vibration have been used successfully to describe the normal voice [1, 2], vocal fold paralysis [3, 4, 5, 6], phonation onset [7], voice instabilities at high pressures [8] and source-tract coupling [9, 10].

In contrast to most mammals, birds do not generate vocalizations with their larynx but with their uniquely vocal organ, the syrinx [11]. Instead of vocal folds, thickened membranes called “labia” serve as vibrating tissue [12, 13, 14, 15]. However, sound production in birds is thought to be based on aerodynamical principles similar to that of human phonation [12, 13, 16]. Consequently, similar modelling approaches might be applicable and several types of biomechanical models have been developed [12, 16, 17, 18, 19]. It is not obvious that rescaling of the original two-mass model will lead to appropriate oscillations at realistic driving pressures and damping ratios. The syrinx is generally much smaller than the human larynx, which leads to smaller areas for the interactions between airflow and vibrating structures and to higher fundamental frequencies of the produced sound.

In this paper, we develop two biomechanical models of the syrinx to study the onset of sound generation and control of higher harmonics (overtones) in the absence of source-tract coupling.

Low-order models, as discussed in this paper, are oversimplifications of the physiology. More realistic geometries [2], additional degrees of freedom [20], coupling to the resonators [21] or a more detailed description of the jet separation [2] can improve the simulated sound signal significantly. In this paper we consider rather simple models with somewhat less realistic output. These core models allow comprehensive bifurcation analyses and we can address basic questions: Do we find reasonable vibrations of the masses even for much smaller geometries? How do symmetric upper and lower masses and a trapezoidal shape influence sound onset, spectral slope and voice instabilities?

In a first model version (rescaled two-mass model), we adapt the well-known simplified two-mass model to the dimensions of a songbird syrinx. Songbirds have two pairs of bilateral labia, which can operate as two independent sound sources [22]. Our model considers one unilateral pair of labia.

In the second model (trapezoidal model), a more realistic geometry of the so-called LTM (Lateral Tympanic Membranes) is adopted to describe the syrinx of a non-songbird: the ring dove (Streptopelia risoria). The upper and lower masses are connected with massless plates on which pressure can act. Such a configuration with massless plates has been applied to reproduce experimental data from the human voice and to design laryngeal prostheses [21].

Both models exhibit self-sustained oscillations at physiologically realistic parameter values. In the classical two-mass model, as well as in the model of the songbird syrinx, collisions occur at medium pressures leading to strong harmonics. In the model of the dove syrinx, however, collision is partially avoided, leading to more pure
tones. We relate these observations to the widely discussed topic of how birds control the intensity of their harmonics [23, 24, 25, 26, 27, 28, 29, 30]. The trapezoidal model exhibits coexisting vibratory regimes that resemble vocal registers. At the same subsyringeal pressure, vibrations with and without collisions are possible. Slow variation of subsyringeal pressure can induce subharmonics, deterministic chaos and a sudden jump to the other “registers”.

2. The Models: Overview

Both models are governed by the same equations of motion:

\[
\frac{dx_1}{dt} = v_1, \\
\frac{dv_1}{dt} = \frac{1}{m_1} \left( F_1 - r_1 v_1 - k_1 x_1 + I_1 - k_c (x_1 - x_2) \right), \\
\frac{dx_2}{dt} = v_2, \\
\frac{dv_2}{dt} = \frac{1}{m_2} \left( F_2 - r_2 v_2 - k_2 x_2 + I_2 - k_c (x_2 - x_1) \right),
\]

where \( F_i \) are the pressure forces derived from the Bernoulli equation and the jet assumption and \( I_i \) represent the collision forces. We will give a detailed description of both models in the next sections.

3. Rescaled Two-mass Model

3.1. Derivation and voice onset

The classical two-mass model directly describes the well-known phase shift between upper and lower edge of aerodynamically driven vibrating tissue [31]. Originally, this model was derived in order to reproduce human vocal fold vibrations [1]-[10]. Therefore, it is not obvious that a rescaling of the original two-mass model will lead to appropriate oscillations at realistic driving pressures and damping ratios.

The fundamental frequency \( F_0 \) of many bird songs is on the order of 1 kHz. If we assume similar tissue elasticities \((k_1, k_2)\) and density \(\rho\) as in human vocal fold modelling [1], we can derive the appropriate parameters of the rescaled model shown in Table I. We maintain the 5:1 ratio of lower to upper mass thickness (compare Figure 1).

The pressures are derived in earlier studies [5] from the Bernoulli equation using a jet assumption:

\[
P_1 = P_2 \left( 1 - \Theta(a_{\min}) \left( \frac{a_{\min}}{a_{1}} \right)^2 \right) \Theta(a_1),
\]

\[
P_2 = 0,
\]

where \( \Theta(x) \) is the Heaviside function

\[
\Theta(x) = \begin{cases} 
1 & \text{if } x > 0, \\
0 & \text{if } x \leq 0.
\end{cases}
\]

The forces \( F_i, I_i \), read as in previous studies [5]:

\[
F_1 = l d_1 P_1, \\
F_2 = 0, \\
I_i = -\Theta(-a_i) c_i a_i, \quad i \in \{1, 2\}.
\]

Viscous resistance of the vibrating tissues can be expressed in terms of damping ratio \( \zeta = r/2\sqrt{km} \) [1]. Because of the decreasing mass \( m \), \( r \) was rescaled to keep approximately the same \( \zeta \) values as in [1].

There is almost no information available on the prephonatory shape of vibrating tissues in avian phonation studies. Therefore, we have chosen rest areas \( a_{01} \) and \( a_{02} \) that allow easy vibrations, i.e. a low onset pressure. Systematic variations of the configuration revealed that a rectangular or a slightly convergent shape allow realistic onset pressures, close to the values reported in other models, e.g. [32].

Figure 2 shows the onset of self-sustained oscillations (solid Hopf bifurcation line) for increasing subsyringeal pressure \( P_s \) and varying stiffness \( k_1 \). The onset pressure of the rescaled model can be below 0.004 g/(cm ms²) (400 Pascal ≈ 4 cm H₂O¹) around our default parameters. Thus, the pressures are measured in g/(cm ms²)

\[\text{All units are given in centimeters, grams and milliseconds and their corresponding combinations: hence pressure is measured in g/(cm ms²)}\]
rescaled dimensions with proper damping ratios and rest areas lead to an onset of oscillations at realistic pressure values. Therefore, our rescaled two-mass model can serve as a first step to model vibrating tissues in the syrinx.

3.2. Intensity of overtones

The intensity of higher harmonics (overtones) is a widely discussed topic in bird song studies [23, 24, 25, 26, 27, 28, 29, 30]. A pure tone (e.g., a sine wave) has no overtones. Collisions of the vibrating tissues, however, lead to pronounced harmonics [31]. To study the intensity of overtones in the rescaled two-mass model, we calculate the power spectrum of the flow derivative \( \frac{dU}{dt} = U \) (we recall that \( U \) is a reasonable approximation of the radiated sound pressure [1]). We introduce a simple measure of the intensity of overtones - the Harmonics Ratio (HR):

\[
HR = 10 \log \left( \frac{H_1}{H_0} \right)
\]

which is calculated from the spectral power \( H_0 \) at the fundamental frequency and the power at the first harmonic, \( H_1 \). This quantity is closely related to the widely used spectral slope [31, 33]. Values below \(-20\) dB indicate that the signal has weak overtones.

Figure 2 shows the values of HR (in grey scale) for varying sub-syringeal pressure \( P_s \) and stiffness \( k_1 \). Only in the immediate neighborhood of phonation onset do nearly sinusoidal oscillations occur (e.g., at point A in Figure 2). Point B in Figure 2 represents a more typical situation: collisions of the masses lead to rather strong harmonics (Figure 3). In summary, the rescaled two-mass model can be used to model sound production with pronounced overtones. Extensive exploration of the effects of parameter variations revealed that the observed periodic vibrations with collision are quite robust. No register transitions or nonlinear phenomena such as subharmonics were found.

4. Trapezoidal Model

4.1. Geometrical aspects

In this section, we will relax some over-simplifications of the original two-mass model along the lines of Refs. [2, 21]. The trapezoidal model aims to describe the syrinx of the ring dove (Streptopelia risoria). In contrast with songbirds, the vocal organ of ring doves is located at the bronchotracheal junction [34], i.e., above the bifurcation of the trachea into the bronchi. The anatomy of the ring dove syrinx (see Figure 4) suggests that a smoother model is more appropriate (Streptopelia decaocto, [35]; Streptopelia risoria, C.P.H. Elemans unpublished results). Therefore, each side of the LTM is modelled as a system of two masses linked together by three massless plates (see Figure 5), as presented in [21]. The parameters listed in Table II were obtained from anatomical studies (C.P.H. Elemans, unpublished results). Instead of the estimated total mass of approximately 9 mg, we assume a vibrating mass of 2 mg in order to achieve a reasonable fundamental frequency. Two important modifications of the two-mass model are introduced:

- symmetry between upper and lower masses (i.e., \( m_1 = m_2, k_1 = k_2, r_1 = r_2 \)) as in [21], because there is no anatomical reason to introduce a small upper mass as in the classical two-mass model
- smoothed geometry via lower and upper plates which are characterized by the height parameters \( d_1 \) and \( d_3 \), respectively.

These modifications affect the calculation of pressure and collision forces significantly. For example, even for a closed syrinx there is a pressure force acting on the lower mass via the lower plate.
the corresponding area. The collision forces $I_1, I_2$ can be written as a linear function of the areas $a_1, a_2$.

In our trapezoidal model, most of the quantities have to be generalized to adapt to the new model geometry. First, we assume for simplicity that all forces above the imaginary horizontal midline at $d_M = d_1 + d_2$ act on the upper mass whereas all forces below that midline drive the lower mass (midline assumption). Note that in Lous et al. [21], $F_1$ and $F_2$ are derived from a balance between forces and torques leading to a more detailed derivation of the force terms.

Second, we define the syringeal area at height $z$ as $a(z) = 2 lx(z)$, where $x(z)$ is obtained by means of plate equations, i.e., the linear equations which identify each plate on the plane $xz$. Particular values of $a(z)$ are $a_1, a_2$ (syringeal areas at 1st and 2nd mass heights, $d_1$ and $d_1 + d_2$) and $a_M$ (syringeal area at height $d_M$). The minimum syringeal area $a_{min}$ is defined as $a_{min} = \min\{a_1, a_2\}$.

### 4.2. Calculation of forces

In the traditional two-mass model, the area exposed to pressure (or collision) is always rectangular and normal to the motion of the masses [36]. In order to derive the pressure forces $F_1, F_2$, we need to multiply the pressure with the corresponding area. The collision forces $I_1, I_2$ can be written as a linear function of the areas $a_1, a_2$.

In our trapezoidal model, most of the quantities have to be generalized to adapt to the new model geometry. First, we assume for simplicity that all forces above the imaginary horizontal midline at $d_M = d_1 + \frac{d_2}{2}$ act on the upper mass whereas all forces below that midline drive the lower mass (midline assumption). Note that in Lous et al. [21], $F_1$ and $F_2$ are derived from a balance between forces and torques leading to a more detailed derivation of the force terms.

Second, we define the syringeal area at height $z$ as $a(z) = 2 lx(z)$, where $x(z)$ is obtained by means of plate equations, i.e., the linear equations which identify each plate on the plane $xz$. Particular values of $a(z)$ are $a_1, a_2$ (syringeal areas at 1st and 2nd mass heights, $d_1$ and $d_1 + d_2$) and $a_M$ (syringeal area at height $d_M$). The minimum syringeal area $a_{min}$ is defined as $a_{min} = \min\{a_1, a_2\}$.

### 4.2.1. Pressure force

By means of the Bernoulli equation, jet separation assumption and $a(z)$ defined above, we can calculate the pressure $P = P(z)$ in the syringeal lumen at height $z$:

$$P(z) = \begin{cases} \frac{P_s}{z_{min} - z} & \text{if } a_{min} > 0, \\ \frac{P_s}{z_{min} - z} & \text{if } a_{min} \leq 0, \end{cases}$$

where $z_m$ is the ordinate at which $a_{min}$ is found and, in case of collision ($a_{min} < 0$), $z_{min}$ is the minimum collision ordinate, i.e. the minimum ordinate $z$ for which $a(z) \leq 0$.

If $z_0$ and $z_1$ are two generic ordinates of points belonging to the same plate and $\alpha$ the angle formed by that plate with the $z$ axis, the pressure force acting on the plate area between $z_0$ and $z_1$ will be by definition $F(z_0, z_1) = \int_A P(z) dA$, where $A$ is the plate area

$$A = \frac{z_1 - z_0}{\cos \alpha} l$$

Because no vertical motion is supposed, the component along the $x$ axis of this pressure force reads:

$$F_x(z_0, z_1) = \cos \alpha F(z_0, z_1)$$

$$= \cos \alpha \int_0^{z_1} \left[1 - \frac{z}{l}\right] dA = \int_0^{z_1} P(z) dz.$$

Consequently, and by means of the midline assumption, the forces $F_1, F_2$ read:

$$F_1 = F_x(0, d_1) + F_x(d_1, d_M).$$

$$F_2 = F_x(d_M, d_2).$$

### 4.2.2. Collision force

Traditional two-mass models do not require the calculation of contact area, because the projected area is rectangular and there is no gradation in opening and closing [36]. We define a collision force that is consistent with (7) and admits a gradual variation of contact area in time. First, we remark that each collision force $I_i$ is zero if:

$$I_i = 0 \iff a_i \geq 0 \text{ AND } a_M \geq 0, \quad i \in \{1, 2\}$$
If the collision force $I_i$ is nonzero, we define respectively:

$$I_1 = \frac{1}{\delta(0, dM)} \int_0^{dM} -\Theta(-a(z))a(z) \frac{c_i}{2z} dz,$$

$$I_2 = \frac{1}{\delta(dM, d_2)} \int_{dM}^{d_2} -\Theta(-a(z))a(z) \frac{c_i}{2z} dz,$$

where $\delta(z_1, z_2)$ is the distance (on the line $x = 0$) between $z_1$ and $z_2$ along which $a(z) < 0$.

We chose to normalize $I_1$ and $I_2$ using distance $\delta$ in order to keep $c_i$ of the dimension of $g/ms^2$ and to obtain a generalization of the previous collision force. The latter is found comparing it with the degenerate case $d_1 = 0$. For the implementation of $I_i$, we take into account that the integrals in eq. (10) and (11) are proportional either to triangular or trapezoidal areas of colliding masses.

### 4.3. Bifurcations in the trapezoidal model

As in section 3.2, we study the onset of oscillations and the strength of harmonics in our trapezoidal model. The upper panel in Figure 6 shows that oscillations can be obtained at fairly low subsyringeal pressures (> 200 Pa). Increasing stiffness leads to an almost linear increase of the onset pressure as observed earlier in other models [5, 37]. An essential modification of the standard two-mass model is the smoother geometry, i.e. a non-vanishing height $d_1$. It turns out that in a certain range around our default parameter ($d_1 = 0.004$ cm) oscillations are easily obtained (lower graph in Figure 6).

Figure 7 shows a one-dimensional bifurcation diagram for increasing subsyringeal pressure $P_s$. Figure 8 shows a more detailed bifurcation diagram near the onset visualizing coexisting limit cycles and negative values of the areas corresponding to colliding tissues. We observe in Figure 7 a sudden onset of oscillations around 300 Pa via a subcritical Hopf bifurcation. The amplitude of the resulting limit cycle increases and no collision occurs (see Figure 8). At 425 Pa another limit cycle with larger amplitudes is observed exhibiting collisions of the upper part described by $x_2$ (corresponding to negative values of $a_2$ in Figure 8). Furthermore, Figure 7 implies that quite distinct vibration patterns coexist at the same pressure. The large limit cycle contains stronger harmonics than the smaller one (Figure 9). These observations resemble observations in previous experiments of the chest to falsetto transitions in excised larynx studies [38]. Several register-like transitions were also found in simulations of an extended two-mass model [36]. Because the suppression of collision and the register transitions are novel features of our trapezoidal model, we discuss these phenomena in some detail.

### 4.4. Avoidance of collisions

As illustrated in Figure 2, the standard two-mass model is characterized by strong harmonics due to collision even at small and medium pressures. However, the lower limit cycle of our trapezoidal model is collision-free even at high pressures and consequently has only weak harmonics (lower panel in Figure 9). This is due to the increasing steady state area (see Figure 8) which pulls apart the masses. The equilibrium in the standard two-mass model, however, is constant, i.e. it does not depend on the subglottal pressure. If we set the derivatives with respect to time in the equations (1)–(4) to zero, we obtain in the two-mass model $x_2 = x_1 k_c/(k_2 + k_c)$. For a rectangular
we always have a1 > a2, i.e. a convergent configuration.

Because for a convergent shape the suprasyringeal pressure has little effect on the masses, such a persistent convergent configuration might lead to a reduced source-tract interaction. It is clearly visible that both areas remain positive for the small limit cycle, while the large limit cycle exhibits negative values of a2 corresponding to colliding tissues.

4.5. Register transitions via subharmonics and chaos

As described above, we found coexistence of a small and a large limit cycle in the range of 425-535 Pa. This implies that different initial conditions lead to distinct vibration patterns.

Furthermore, small perturbations can induce sudden jumps from one attractor to another. In this section we analyze the transition from the large attractor with collisions to the small limit cycle due to a slow increase of the suprasyringeal pressure. Figure 10 shows the spectrogram of the sound pressure generated during the transition from the large limit cycle (with collisions) to the small limit cycle caused by a slow increase of the suprasyringeal pressure. It is evident that there is a jump from harmonic rich spectra to a more sinusoidal oscillation.

Furthermore, subharmonics and noise-like components are visible. In Figure 11, high-resolution spectra² confirm the appearance of subharmonics and chaos. Using phase portraits and Poincaré sections (not shown), we have confirmed the existence of deterministic chaos in our trapezoidal model.

5. Discussion

Our simulations show that rescaled biomechanical models originally developed to describe mammalian vocal fold vi-

² Integration time from 1000 to 2000 ms in order to avoid transients and time step of 0.0005 msec (2000 kHz sampling rate).
brations can be adapted to model the bird syrinx. We assumed that both sound producing organs are excited by the same principle: in the opening phase a high pressure drives the vibrating structure apart and during closure the pressure is reduced due to the Bernoulli force. The fundamental frequency is governed by the mass and stiffness of the vibrating tissue.

Our simulations represent symmetric vibrations. Interestingly, the same model equations can be used to model a single vibrating structure (a “hemi-syrinx”). In this case only the sound intensity is reduced but onset pressure or intensity of harmonics are identical.

Control of harmonics in bird songs is widely debated [23, 24, 25, 26, 27, 28, 29, 30]: some species have whistle-like songs and not much energy is found in the harmonics [26], whereas other species, such as the zebra finch (Tae nopygia gut tata), display strong harmonics during song or calls.

In all two-mass model versions almost pure tones are found near the onset of vibrations (i.e. near the Hopf bifurcation line shown in Figure 2). In the rescaled two-mass model, strong harmonics appear at higher pressures due to collisions. For a small upper mass and a rectangular geometry, collisions leading to strong harmonics can be avoided only near the phonation onset. At higher pressures countering forces would be required to diminish collisions. We hypothesize that the avoidance of strong collisions in song birds might be achieved by the medial tympaniform membranes (MTM) that are continuous with the inner vibrating labia [19]. This possible function of the MTM will be discussed in more details in a forthcoming study.

In our model of the ring dove syrinx no collisions occur at default parameters. Consequently, harmonics are fairly weak. The smoother configuration and equal upper and lower masses counteract collisions even at relatively high pressures. This is presumably due to a stronger effect of the subsyringeal pressure acting on both masses. In a recent experimental paper [23], Riede et al. showed how varying suprasyringeal configurations can suppress the second harmonic in ring doves, and Fletcher et al. [29] showed that the combined influence of trachea, glottis and upper esophagus acts as an effective band-pass filter that eliminates higher harmonics generated by the dove syrinx. Our simulations reveal that the configuration of the syrinx influences the intensity of overtones. Therefore, the amount of energy in the harmonics could also be controlled by syringeal muscles that directly affect the configuration of the syrinx [39, 40].

Around our default parameters given in Table I we found no voice instabilities. On the other hand, we have shown that our trapezoidal model exhibits coexistence of attractors and jumps from harmonic rich spectra to more sinusoidal oscillations of the radiated sound pressure.

A previously published two-mass model of the songbird syrinx [12] demonstrated many instabilities such as period doublings, transitions from periodic to chaotic dynamics, as well as mode locking transitions. Unfortunately, we are unable to compare the behaviour of the two models in more details, because the governing equations and the parameter settings were not provided. Ring doves already exhibit stronger harmonics during inspiratory phonation compared to expiratory phonation even at low intensities [23, 27, 41, 42]. This implies that asymmetries between outflow and inflow of the air have to be taken into account. This will be treated in a more sophisticated model. Instabilities such as frequency jumps are commonly observed in the ring dove coo [43, 41]. Even our current model described in this paper exhibits already coexistence of different “registers”.

In this paper we have shown that the geometry and the rest position of the syrinx can influence the harmonic spectra drastically. Our simulations are a first step towards more realistic modelling of the syrinx. In subsequent studies we will incorporate the MTM and the dynamic control of associated superfast syringeal muscles. This will allow a quantitative comparison of observed bird songs and simulations.

Acknowledgements

The authors would like to thank two anonymous reviewers for their comments to improve the manuscript. We acknowledge support from the Deutsche Forschungsgemeinschaft (grant He 2168/7).

References


