THE SOUNDS OF THE AVIAN SYRINX – ARE THEY REALLY FLUTE-LIKE?

Tamara Smyth, Julius O. Smith III

Center for Computer Research in Music and Acoustics
Stanford University
tamara@ccrma.stanford.edu, jos@ccrma.stanford.edu

ABSTRACT

This research presents a model of the avian vocal tract, implemented using classical waveguide synthesis and numerical methods. The vocal organ of the songbird, the syrinx, has a unique topography of acoustic tubes (a trachea with a bifurcation at its base) making it a rather unique subject for waveguide synthesis. In the upper region of the two bifid bronchi lies a nonlinear vibrating membrane – the primary resonator in sound production. Unlike most reed musical instruments, the more significant displacement of the membrane is perpendicular to the directions of airflow, due to the Bernoulli effect. The model of the membrane displacement, and the resulting pressure through the constriction created by the membrane motion, is therefore derived beginning with the Bernoulli equation.

1. INTRODUCTION

Birdsong is commonly associated with the sounds of a flute – evidence of this can be found in the titles many composers have given their works that feature the flute: Vivaldi's flute concerto *Il gardellino* ("The Goldfinch"), Debussey's *Syrinx* for flute solo, and Messiaen's work for flute and piano, *Merle noire* ("Blackbird") are only a few examples.

The pure, often high pitched, tone of the songbird is undeniably flute-like. It is, perhaps, this pure tone quality that led to so much debate in the literature regarding its actual method of sound production. The theory that the sound was produced by an aerodynamic whistle effect has been widely confuted [7, 1, 2]. The avian vocal tract uses a nonlinear vibrating membrane as its primary excitation mechanism. The syringeal membrane, much like the vocal folds in the human vocal tract, form a vibrating valve, the output of which is filtered by the trachea, amplifying and attenuating certain modes of the vibrating membrane.

Though the modeling methods described here are generally used for the synthesis of musical instruments, their use in modeling the bird's unique vocal system offers a configuration of acoustic elements not found in traditional musical instruments.

2. THE AVIAN VOCAL TRACT

The syrinx is the bird's unique vocal organ. It consists of an airway – a trachea which divides into the left and right bronchus at its base (Fig. 1), and two pressure-controlled valves made of flexible membranes, the tensions of which are altered by surrounding muscles [8]. The neural control of the muscles and the bird's respiratory mechanics both greatly contribute to how sound is modulated by the syrinx [2].

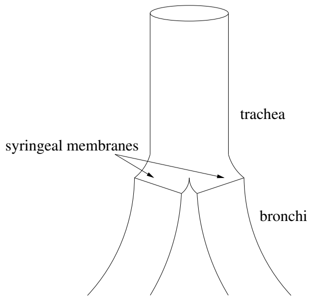

Figure 1: *The syrinx.*

Airflow in the syrinx begins from the lungs and passes through the bronchi, trachea and mouth before radiating from the beak (Fig. 2). On its way, the airflow passes through the bird's primary vocal organ, a non-linear vibrating membrane, situated just below the junction of the two bronchi with the trachea (Fig. 2).

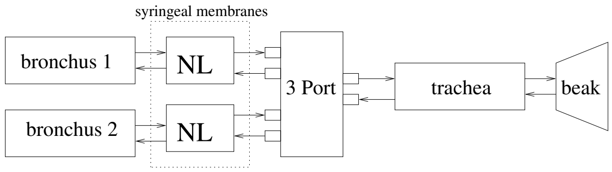

Figure 2: *Block diagram of avian vocal tract.*

2.1. The Syringeal Membrane

When the bird is at rest, the syringeal membrane lies flat on the wall of the bronchus. When singing begins, the membrane vibrates, with motion occurring primarily toward the opposite cartilaginous wall (Fig. 3). This creates a narrowing in the bronchus (with the possibility of closing the air passage completely), and a constriction through which the air flows.

The motion of the syringeal membrane can be described by two of the three configurations for pressure-controlled valves in acoustic tubes [4]. The three possible configurations are: 1) The reed is blown closed (as in woodwind instruments). 2) The reed is blown open (as in the human larynx). 3) The transverse model – the Bernoulli pressure causes the valve to close, perpendicular to the direction of airflow [6, 4]. Though the syrinx uses both configurations 2 and 3, as does the human voice, the third tends to be most significant and is used as the model for implementing the membrane's motion. It is interesting to note that this third configuration does not seem to be employed as significantly by traditional musical instruments [4]. A musical exploration of this model could, therefore, make new contributions to the synthesis of new sounds in computer music.

2.1.1. *The Membrane Modeled as the Transverse Configuration of a Pressure Controlled Valve*

In voiced song (as opposed to whistled song), the membrane is set into motion by airflow. Like any mechanical resonator, it vibrates at a frequency controlled partly by its natural frequency (determined by its mass and tension) and partly by the resonance of the air column to which which it is connected.

The model of the syringeal membrane, the nonlinearity (NL in Fig. 2), determines the pressure on the tracheal side of the constriction (p_1) based on a given input pressure (p_0) from the bronchial side. p_1 becomes the input pressure to the waveguide that simulates the trachea.

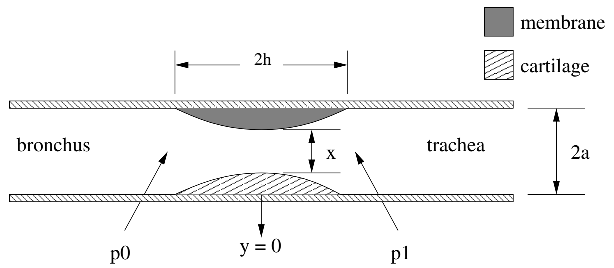

Figure 3: *The pressure-controlled valve in the syrinx.*

The nonlinear junction of the bronchi to the trachea was developed following the acoustic model by Fletcher of airflow through the syringeal constriction [3, 5]. The model has the following four key variables, all of which vary over time during sound production.

p_0 $\triangleq$ pressure on the bronchial side of the constriction
U $\triangleq$ air volume flow through the syrinx
x $\triangleq$ displacement of the membrane
p_1 $\triangleq$ pressure on the tracheal side of the constriction

2.1.2. *Bronchial pressure ($\mathbf{p_0}$)*

The bronchus is considerably shorter than the trachea (approximately one tenth its length). If modeled using waveguide synthesis at audio sampling rates, a delayline of less than 2 samples would be required. It is preferable, therefore, to treat the bronchus as having a volume (V) and an acoustic stiffness $\frac{\rho c^2}{V}$. The rate at which pressure builds up in the bronchus ($\frac{dp_0}{dt}$) is then proportional to the difference in volume velocity flowing *in* ($\frac{p_G - p_0}{Z_G}$) and flowing *out* (U) of the constriction.

$$\frac{dp_0}{dt} = \left(\frac{\rho c^2}{V}\right)\left(\frac{p_G - p_0}{Z_G} - U\right) \tag{1}$$

where,

Z_G $\triangleq$ impedance of the air sacs
ρ $\triangleq$ air density
V $\triangleq$ volume of the bronchus
c $\triangleq$ speed of sound (in air)

2.1.3. *Volume flow ($\mathbf{U}$)*

Since the motion of the membrane causes varying heights in the constriction, Bernoulli's equation (2) is used to determine the pressure at a given point y:

$$p(y) = p_0 + \frac{\rho}{2}\left[v_0^2 - v(y)^2\right] \tag{2}$$

where v is the particle velocity at a given point, and is equal to the volume velocity, U, divided by the cross-section area at that point. That is,

$$v(y) = \frac{U}{2a\left[x + (a-x)(\frac{y}{h})^2\right]} \tag{3}$$

Since U likely forms a jet at $y > 0$, the pressure acting on the entire tracheal half of the membrane is equal to the pressure at the base of the trachea [3]. That is, the pressure at point y = 0 is effectively p_1.

Given that the area of the constriction at point $y = 0$ is $2ax$, (2) becomes

$$p_1 = p_0 + \frac{\rho}{2}\left[\left(\frac{U}{2\pi a^2}\right)^2 - \left(\frac{U}{2ax}\right)^2\right] \tag{4}$$

which can be reduced to

$$p_1 = p_0 - \frac{\rho}{2}\left(\frac{U}{2ax}\right)^2 \tag{5}$$

when $x \ll \pi a$.

The final differential equation governing airflow is obtained by incorporating the force used to accelerate air through the syrinx [3] (this adds a pressure drop of $\frac{\rho}{2\sqrt{ax}}\frac{dU}{dt}$), and by rearranging the equation to isolate $\frac{dU}{dt}$:

$$\frac{dU}{dt} = \frac{2\sqrt{ax}}{\rho}\left(p_0 - p_1 - \frac{\rho}{8a^2x^2}U^2\right) \tag{6}$$

2.1.4. *Membrane motion ($\mathbf{x}$)*

When x, the displacement of the membrane, is equal to zero, the membrane is touching the opposite wall, and the constriction is closed.

Fletcher writes the membrane's motion for mode n as:

$$m_n\left[\frac{d^2x_n}{dt^2} + 2\kappa\frac{dx_n}{dt} + \omega_n^2(x_n - x_0)\right] = \epsilon_n F \tag{7}$$

where

ω_n $\triangleq$ the radian frequency of mode n
m_n $\triangleq$ the effective mass associated with mode n
κ $\triangleq$ the damping coefficient
ϵ_n $\triangleq$ the coupling coefficient between F and mode n

and F is the force driving the fundamental mode of the membrane, expressed as

$$F = ah(p_0 + p_1) - \frac{2\rho U^2 h}{7(ax)^{1.5}} \quad \text{for } x > 0 \tag{8}$$

In order to account for the damping that would occur should the membrane actually touch the opposite wall, a factor E is introduced into (7). Its actual value depends on the stickiness of the contact between the membrane and the wall [3].

$$\kappa \to E\kappa \quad 10 \leq E \leq 100 \quad \text{if } x \leq 0 \tag{9}$$

When the membrane's displacement becomes very large, surrounding tissue will likely take part in the vibration [3]. To account for this, the mass m from (7) is replaced with the following term, so that the mass changes nonlinearly depending on the membrane's position (x).

$$m \to m\left(1 + \eta\left(\frac{x - x_0}{h}\right)^2\right) \tag{10}$$

The 2nd derivative of x_n is isolated by rearranging (7) and, for simplicity, the coupling coefficient (ϵ) is made equal to unity.

$$\frac{d^2x_n}{dt^2} = \frac{\epsilon F}{m_n} - 2\kappa\frac{dx_n}{dt} - \omega^2(x_n - x_0) \tag{11}$$

2.1.5. *Tracheal pressure* ($\mathbf{p_1}$)

The pressure leaving the constriction (p_1) is proportional to the volume velocity (U), scaled by the characteristic impedance of the trachea ($Z_0 = \frac{\rho c}{\pi a^2}$). In order to obtain the true pressure at the base of the trachea, the pressure due to all previous reflections in the trachea must also be considered. The flow diagram in Fig. 4 illustrates how the junction between the constriction and the trachea was implemented.

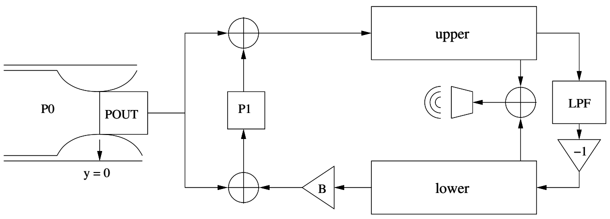

Figure 4: *Model of the interaction between the membrane and the trachea.*

The output of the lower rail is scaled by B to account for wall losses of the pressure wave traveling through the trachea. B is calculated according to the length of the tube and the attenuation coefficient $\alpha = \frac{1.2\times10^{-5}\sqrt{\omega}}{a}$.

Radiation losses at the beak are accounted for using a lowpass filter (LPF) (as wider pipes attenuate high frequencies) and then negated to simulate the reflection of an open end tube.

2.2. Numerical methods

2.2.1. *Backwards difference and the Trapezoid Rule*

The following difference equations are commonly used for numerically integrating differential equations:

$$\begin{aligned} x'[n] &= \frac{x[n] - x[n-1]}{T} \\ x''[n] &= \frac{x[n-2] - 2x[n-1] + x[n]}{T^2} \end{aligned} \tag{12}$$

where $x'[n]$ and $x''[n]$ are the first derivative and second derivatives of $x[n]$ respectively, and T is the sampling period. Since the current model successfully isolates the derivatives of the variables in question (and it is the instantaneous value of these variables that we require), these equations are more conveniently expressed as follows:

$$\begin{aligned} x[n] &= x[n-1] + x'[n]T \\ x'[n] &= x'[n-1] + x''[n]T \end{aligned} \tag{13}$$

The accuracy of the backwards difference approximation is dependent on T (the sampling period). That is, it is first order accurate in T [9]. The first implementation of this model using this method becomes very unstable when discretizing at an audio sampling period ($T = 1/44100$).

The second implementation uses a more sophisticated algorithm, which is second order accurate in T, and uses the trapezoid rule for numeric integration [9, 10]. It yields the following difference equations (comparable to (13)), and solves the problem of stability at larger sampling periods:

$$\begin{aligned} x[n] &= x[n-1] + \left(x'[n] + x'[n-1]\right)\left(\frac{T}{2}\right) \\ x'[n] &= x'[n-1] + \left(x''[n] + x''[n-1]\right)\left(\frac{T}{2}\right) \end{aligned} \tag{14}$$

2.2.2. *Solving the model numerically*

For the purpose of discussion, the model of the vibrating membrane can be reduced to the differential equations (1), (6), (11). Solving the system of equations is done as follows:

All variables and their derivatives are initialized to zero. Pressure from the the air sacs is introduced into the system by setting p_G to some value.

The derivatives of p_0, U, and x are descretized as follows:

$$\begin{aligned} p_0' &= \left(p_0' + \left(\frac{\rho c^2}{V}\right)\left(\frac{p_G - p_0}{Z_G} - U\right)\right)\frac{T}{2} \\ U' &= \left(U' + D\left(p_0 - p_1 - \frac{\rho}{8a^2x^2}U^2\right)\frac{T}{2}\right) \quad 0 < x \leq a \\ x_n'' &= \left(x_n'' + \frac{\epsilon F}{m_n} - 2\kappa\frac{dx_n}{dt} - \omega^2(x_n - x_0)\right)\frac{T}{2} \end{aligned} \tag{15}$$

The result is then added to the variable in question (along with the previous value of that variable).

In the case of membrane displacement x, the displacement is calculated for 2 modes (giving x_1 and x_2) which are added to produce the overall displacement of the membrane.

The displacement (x), and the pressure at the base of the trachea (p_1), is plotted in Fig. 5 along with the audio output (taken at the top of the trachea). It is interesting to observe how the pressure (taken at both the top and bottom of the trachea) relates to the displacement of the membrane. Both pressure waves display a pulsation that clearly coincides with the frequency of the oscillating membrane. When the membrane closes the constriction completely (i.e., when $x = 0$), the amplitude of p_1 is high and its frequency is close to the resonance of a closed tube. When the membrane is open ($x > 0$), the pressure amplitude decreases and the frequency increases to the resonance of an open tube.

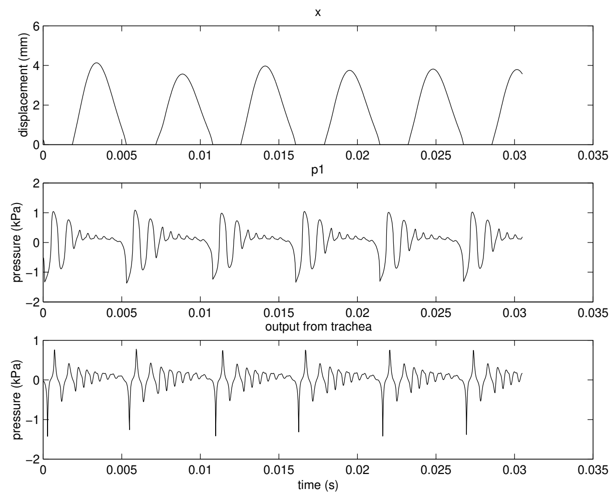

Figure 5: *Waveforms are comparable to those published by Fletcher, but use waveguides with lumped losses and a 2nd-order-error algorithm for numerical integration to achieve stable discretization at audio sampling rates.*

This relationship can also be seen in a sonogram of the pressure wave at the top of the trachea (Fig. 6). The spacing of the harmonics indicates a fundamental frequency that matches the first mode of the vibrating membrane. Very strong frequency components at 1200Hz and 2200Hz show the contribution of the resonance of an open and closed trachea (the latter being lowered by the loading of the constriction [3]).

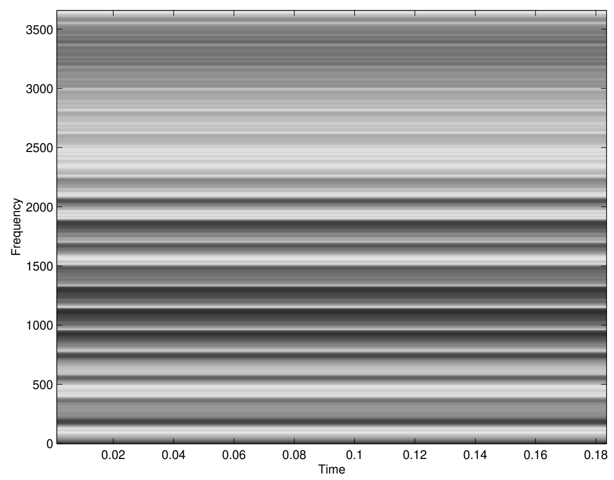

Figure 6: *Sonogram of trachea output.*

3. CONCLUSIONS AND FUTURE WORK

The output of the model is the voiced sound produced by the vibrating membranes in the syrinx (sounding much like a raven). It is very rich in harmonics and is not the flute-like tone associated with the songbird. This again raises the unresolved question of the *actual* sound production mechanism, an issue constantly surfacing in the literature. The suggestion that sound is also produced by vortices set up as air is forced through the constriction created by the membrane may be erroneous scientifically, but the field of music is not constrained to scientific reality. The current model is unable to produce a pure sinusoidal tone; it may therefore be useful musically to extend the model to provide this functionality.

4. REFERENCES

[1] M. R. Ballintijn and C. T. Cate. Sound production in the collard dove: A test of the 'whistle' hypothesis. *Journal of Experimental Biology*, 201:1637–1649, 1998.

[2] J. Brackenbury. *Form and Function in Birds*, chapter Functions of the syrinx and control of sound production, pages 193–220. New York Academic Press, 1989.

[3] N. H. Fletcher. Bird song – a quantitative acoustic model. *Journal of Theoretical Biology*, 135:455–481, 1988.

[4] N. H. Fletcher. *Acoustic Systems in Biology*. Oxford University Press, New York, New York, 1992.

[5] N. H. Fletcher. Private email correspondence, July 2002.

[6] N. H. Fletcher and T. D. Rossing. *The Physics of Musical Instruments*. Springer-Verlag, 1995.

[7] F. Goller and O. N. Larsen. *In Situ* biomechanics of the syrinx and sound generation in pigeons. *Journal of Experimental Biology*, 200:2165–2176, 1997.

[8] A. S. King. *Form and Function in Birds*, chapter Functional Anatomy of the Syrinx, pages 105–191. New York Academic Press, 1989.

[9] J. O. Smith. *Discrete-time lumped models*. Course notes for MUS421/EE367B, Stanford University, May 2002: `http://www-ccrma.stanford.edu/~jos/NumericalInt/`.

[10] J. O. Smith. *Wave digital filters*. Course notes for MUS421/EE367B, Stanford University, March 2002: `http://www-ccrma.stanford.edu/~jos/WaveDigitalFilters/`.