

Dynamical Systems Analysis of Birdsong Generation

Ilknur Icke
Department of Computer Science
University of Vermont
Burlington, VT 05401
ilknur.icke@uvm.edu

ABSTRACT

The study of dynamical systems has proven to be an invaluable tool for understanding how songbirds sing. A very simple nonlinear oscillator model that is similar to a spring-mass system can shed a light onto the physics of the avian vocal organ, the syrinx. The neuronal control mechanisms that generate the songs can be modeled as a central pattern generator, which is a special kind of connectionist model. Furthermore, dynamic variation of the system parameters via temporal hierarchy gives rise to highly precise modeling of birdsongs.

Keywords

birdsong, dynamical systems

1. INTRODUCTION

Songbirds are one of the most colorful ingredients of the nature. Out of the approximately 10,000 bird species, around 4,000 species are known to be songbirds (oscines) that communicate via songs for the purposes of territorial ownership and mating [4, 5]. What is more fascinating from a neuroscience perspective is that oscine birds are capable of learning songs by listening to a tutor which bears certain resemblance to how humans learn to speak. Indeed, the topic of birdsong learning and production has been an active area since the 1990s for many researchers who want to understand how humans learn speech.

The dynamical systems theory has been widely applied to the study of the mechanisms of birdsong generation. Various mathematical models of the avian vocal organ (the syrinx) have been developed in order to describe how the oscillations of the vocal organ creates the rich set of birdsongs. These nonlinear oscillator models have also been incorporated into generative systems as an alternative means to create musical automata as opposed to the symbolic Artificial Intelligence based approaches [3].

This report aims to review one of the simplest nonlinear oscillator models for the syrinx that is a slight variation of

a Van der Pol oscillator and present how the methods of bifurcation theory help us understand the dynamics of birdsong generation. We then briefly discuss how the hierarchical organization of the brain corresponds to the hierarchical structure of the birdsong by means of multiple time scales.

2. STRUCTURE OF A BIRDSONG

Each birdsong exhibits a complex acoustic structure which is made up of basic building blocks that are called the syllables. The duration of a syllable varies from species to species. For instance, canary syllables last between 15-300 ms. The frequency range in a typical birdsong is generally between 2-7Khz. Songbirds have several dozen syllables which are combined to create different songs in their repertoire [4, 6]. A birdsong is a hierarchical structure, repetitions of certain syllables make up the motifs and repetitions of certain motifs make up the whole song (figure 1).

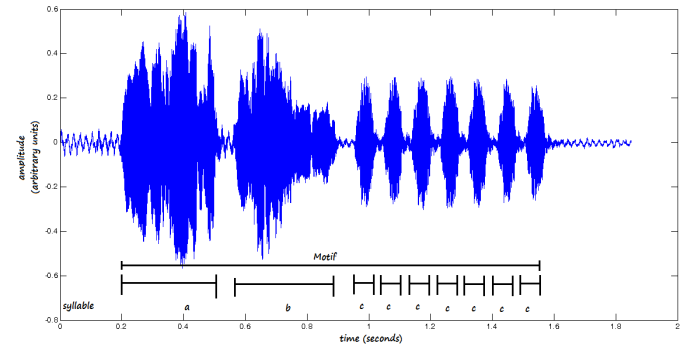


Figure 1: Waveform plot of a Chingolo's song.

Despite being the smallest building block of a birdsong, a syllable is made up of multiple tones. Figure 2 shows the sonogram of the Chingolo's song shown above. A sonogram is a standard tool amongst the researchers to visualize a birdsong in terms of the evolution of the sound frequency across time in seconds. The syllables in the birdsongs are clearly visible as separate markings on a sonogram. As it is evident from the spectral content for each syllable for the birdsong plotted in figure 2, the frequency within a syllable is not static. On the contrary, in most syllables, the frequency fluctuates over time: as an upsweep (as in syllable a) or a downsweep (as in syllables b and c) while pure-tone vocalizations are still possible. It is this richness in the dynamics that makes birdsongs an interesting topic for mathematical modeling.

2012

Unless otherwise specified in the text, all images have been generated by the author.

Copyright Ilknur Icke.

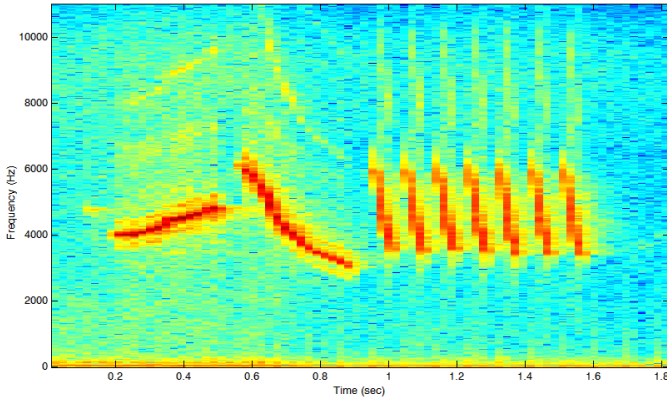


Figure 2: Sonogram of the Chingolo's song.

3. SONGBIRD ANATOMY

Songbirds possess a sound generation mechanism (figure 3) that is driven by the signals from the brain. Namely, the singing starts in the brain and through a neuronal hierarchy, the brain controls the muscles that govern the breathing and tensions of the vocal organs. Located between the lungs and the vocal tract, the syrinx is the sound-producing organ (source) of a songbird. Similar to the larynx in humans, the syrinx contains vibrating membranes that are called labia which oscillate when the air from lungs exerts force on them. The sound waves generated by the syrinx then travel through the vocal tract which serves as the filtering mechanism for producing the vocalizations [1]. Despite the evident similarity in working principles, the syrinx anatomy is fundamentally different than the human larynx. The syrinx contains two separate junctions connecting the vocal tract to the lungs (bronchi). As a result, the songbird has two separate sources of sound generation which can be used separately or concurrently in order to generate complex vocalizations.

The researchers believe that a relatively small set of muscles are in charge of songbird vocalizations. These are the seven muscles in the syrinx (figure 3 c), the muscles controlling the breathing and the muscles controlling the beak [1]. However, this is a high dimensional problem that maps the parameters of each muscle movement into the one dimensional time series of air pressure that is measured by a microphone. In order to understand how muscles move during song generation, researchers placed sensors in live bird's vocal organs in order to measure muscle activity during singing. One interesting finding was that the songbirds take minibreathes in between the syllables. The researchers found that the muscle syringealis ventralis (vS) controls the stiffness of the labia and the tracheobronchialis (TB) and syringealis dorsalis (dS) are influential in controlling the air-flow. Therefore, it was deduced that song generation in songbirds could be mathematically modeled with respect to two variables: the stiffness of the labia and the air pressure generated by the air-flow from the lungs to the vocal tract [1].

The dynamics of the syrinx is in an extremely fast time-scale. For example, it has been observed that a canary could produce up to 27 syllables per second [1]. Scientists were also interested in understanding how the brain executes the motor commands that make the bird sing a whole song that is

made up of complex sequences of syllables. Neuroscientists conducted experiments on songbirds by recording activity of the various brain regions as the bird produced vocalizations. Two distinct forebrain regions were identified as being highly active during birdsong generation. The HVC (high vocal center) projects to the region RA (robust nucleus of the archistriatum) which in turn projects to the brainstem regions that control the syrinx and the respiratory systems (figure 3 b). It has been observed that the HVC-projecting neurons, HVC(RA), fire a single 6-ms burst at exactly the same time during the zebra finch's song motifs and different HVC(RA) neurons fire at different times. The RA neurons on the other hand, fire repeatedly during the song and each time they are driven by a different HVC(RA) neuron. The RA neurons are known to be controlling the vocal tract muscles which are most possibly active at all times during song generation [5]. It has also been reported that the activity of RA neurons was associated with the generation of sub-syllables. Based on the observations of the songbird brain activity during singing, it has been proposed that there is a temporal hierarchical organization that controls the bird-song generation [7].

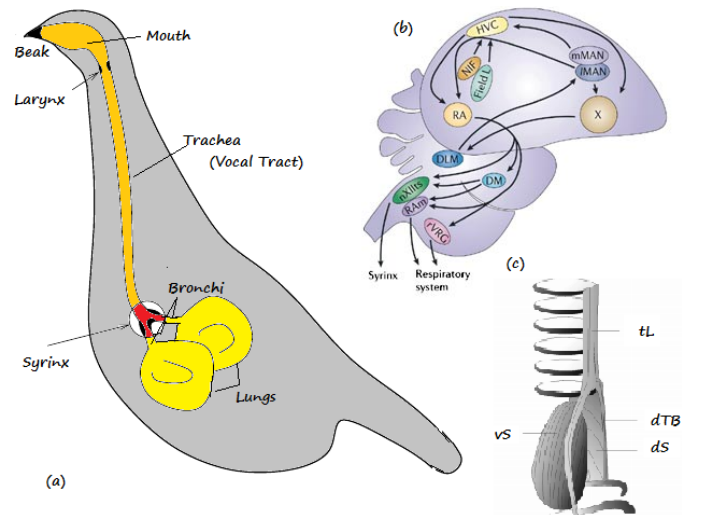


Figure 3: Anatomy of a songbird's sound generation mechanism. (a) The lungs, syrinx and vocal tract. (b) The neural structure that controls the syrinx (c) Anatomy of the syrinx. Image partially taken from [1]

4. MODELING THE SYRINX

4.1 Dynamics of the Syrinx

From physics we know that periodic airflow fluctuations create sound. In songbirds, the syrinx is the organ that generates these fluctuations. In order to study the dynamics of the syrinx, it is necessary to understand the simple oscillatory models. These models that are known as the harmonic oscillators are well-studied models of many natural phenomena such as springs and pendulums. This section first introduces the simple spring-mass oscillator model and presents how a simple model for the syrinx is related to this base model.

4.1.1 Simple Linear and Nonlinear Oscillators

Imagine a small mass attached to a spring which is pulled away from its stable position and then released (figure 1). The spring pulls back the mass by applying a restitution force. The motion of the mass exhibits a periodic pattern as time passes. The strength of the restitution force depends on the initial displacement from the stable position and it is a linear function of this displacement. The larger the displacement, the larger the restitution force that the spring will apply to the mass. The restitution force is also affected by the elasticity of the spring which is represented by a constant K . Given the restitution force $F_{spring} = -Kx$, the acceleration of the mass is defined by

$$a = \frac{d^2x}{dt^2} = \frac{F_{spring}}{M}$$

where

$$M \frac{d^2x}{dt^2} = F_{spring} = -Kx$$

with respect to Newton's Law. Here the - sign indicates that the restitution force is acting in the opposite direction of the mass displacement.

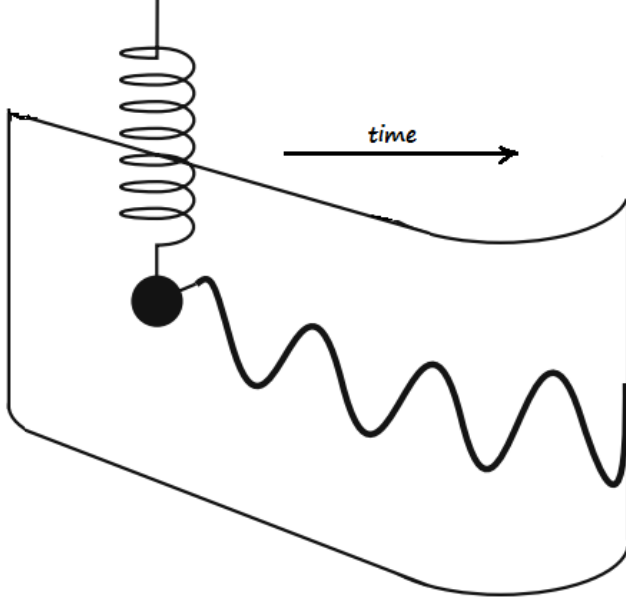


Figure 4: Spring-mass model (adapted from [4])

In this model, there is no energy loss (dissipation) and in theory, the mass should keep oscillating forever. However, this is not the case in reality. Friction is a force that acts on moving bodies in proportion to their velocity and defined as

$$F_{friction} = -B \frac{dx}{dt}$$

The frictional force makes sure that the mass eventually settles down to its stable position instead of oscillating forever. It is possible to sustain the oscillations by applying an external force at every moment that would compensate for the effects of friction. If the external force is greater than the frictional force, then the amplitude of the oscillations

will increase. The external force is defined as

$$F_{external} = C \frac{dx}{dt}$$

It is generally the case that there are extra constraints on the motion of the mass. Physical constraints such as walls bound the motion of the mass. Similar to friction, this dissipative force acts on moving bodies and more intensely as the moving body approaches the wall. This bounding force is no longer linear and a simple model for this function is

$$F_{bounded_motion} = -Dx^2 \frac{dx}{dt}$$

The spring-mass model including all considered forces defined as above can be written as

$$F_{total} = F_{spring} + F_{friction} + F_{external} + F_{bounded_motion}$$

Using Newton's Law we have

$$M \frac{d^2x}{dt^2} = F_{total} = -Kx - B \frac{dx}{dt} + C \frac{dx}{dt} - Dx^2 \frac{dx}{dt}$$

dividing both sides by M we get

$$\frac{d^2x}{dt^2} = -kx - b \frac{dx}{dt} + c \frac{dx}{dt} - dx^2 \frac{dx}{dt}$$

where k, b, c and d are the re-dimensionalized variables as a result of dividing by the mass.

By assigning variable $\dot{x} = y$ where $\dot{x} = \frac{dx}{dt}$ we can write the following 2D system of equations:

$$\begin{cases} \dot{x} = y \\ \dot{y} = -kx - (b - c)y - dx^2y \end{cases}$$

Figure 5 shows examples of the system behavior with respect to various values of k, b, c and d parameters governing the contribution of the respective forces. As it is evident from the waveforms, the addition of the nonlinear term creates complex oscillations which can be decomposed into their harmonic components. Note that the 2D system given above is a slightly modified version of the Van der Pol oscillator whose dynamics are well-studied in the dynamical systems literature.

4.1.2 Modeling Syrinx as a Van der Pol Oscillator

The crosssection of the songbird avian organ, the syrinx, is shown in figure 6. The syrinx is made up of two junctions connecting each bronhus to the trachea. Each junction contains two tissue structures that are called the labia. The labia are flexible tissue structures which can be stretched or compressed. Similar to the spring-mass system, the tissue recovers owing to the restitution force. Moreover, additional forces due to friction and bounded motion also apply to the labia. When compressed, the labia approach the cartilaginous tubes that they are attached to and when stretched, the two labia collide [4]. An external force is necessary in order for the labia to sustain oscillations. It is evident that this external force is related to the airflow between the bronchi to the trachea.

A quantitative model of how the airflow generates labia motion has been studied by Ingo Titze in 1998. Even though the original model was based on human vocal folds, later observations via videography showed that same model could also be used for syrinx [4]. Titze proposed that the complex

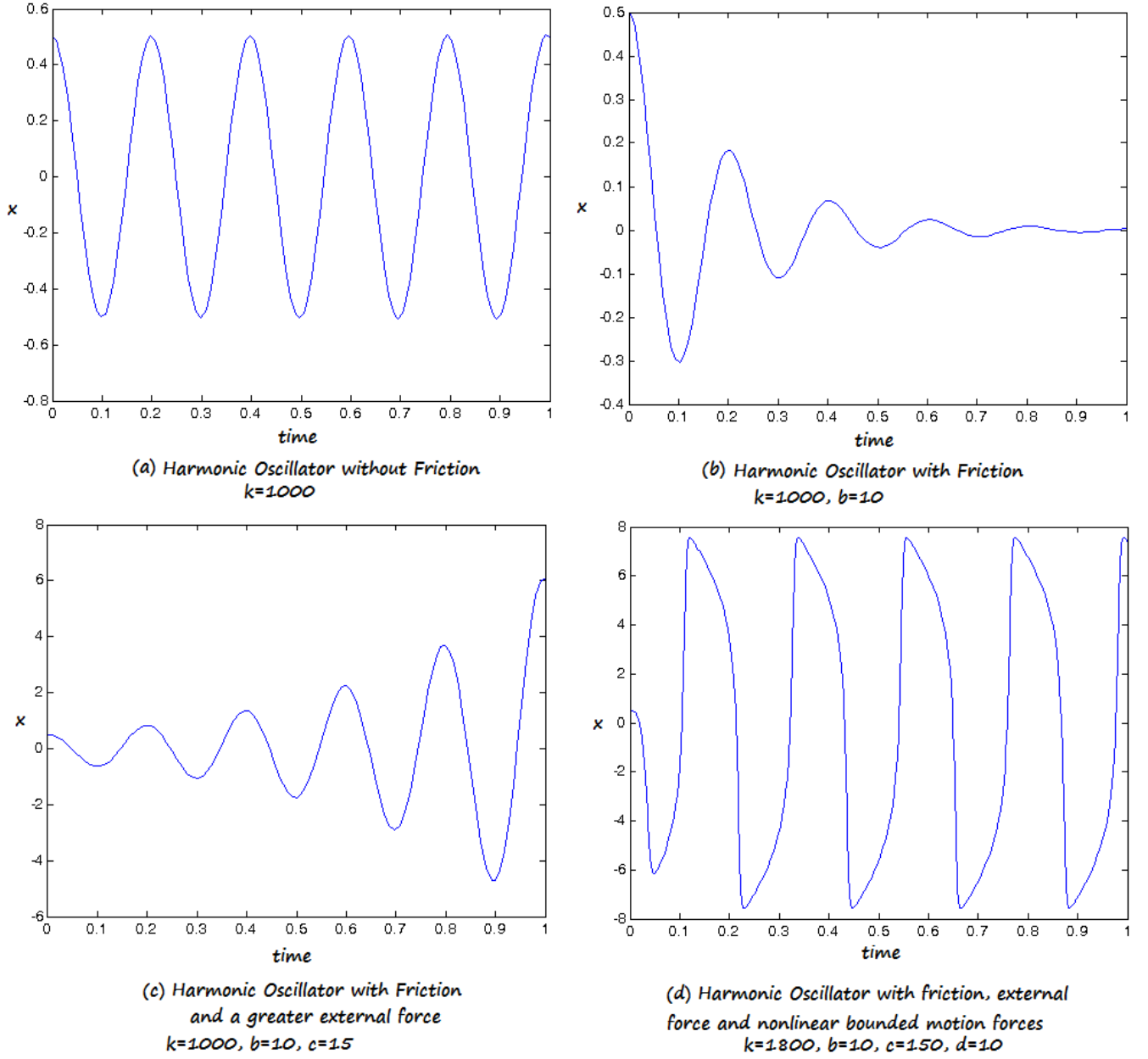


Figure 5: Oscillations of the spring-mass model. The ODEs are numerically integrated using Matlab ode15s function starting from $x=0.5$ and $y=0$.

motions of the two labia as the air passes between them could be reduced to two simple motions. The first motion is described by two tissue masses either moving towards each other or away from each other. The second motion is described by the upper edges of the labia either approaching or moving away from each other while the lower edges are moving away or approaching each other respectively (figure 7).

The first motion is called the lateral displacement and the second motion is called upward-propagating wave. In order for the sustained oscillations to take place these two types of motions should be coordinated in a way that the air flow exerts a force on the moving labia in the same direction as their motion [4]. Under this condition the 2D system govern-

ing the motion of the labia is very similar to the spring-mass system presented in the previous section

$$\begin{cases} \dot{x} = y \\ \dot{y} = -kx + (p - b)y - dx^2y \end{cases}$$

where k is the elasticity of the labia and p is the sublabial air pressure per unit mass. This is the simplest model that is supported by experimental evidence on songbirds. Here, the variable x represents the mid-points of the labia. Despite its simplicity, this model was successful in producing birdsong like vocalizations. Note that this model assumes the lateral mode, that is, it explains the dynamics of one junction of the syrinx. Extended versions of the systems de-

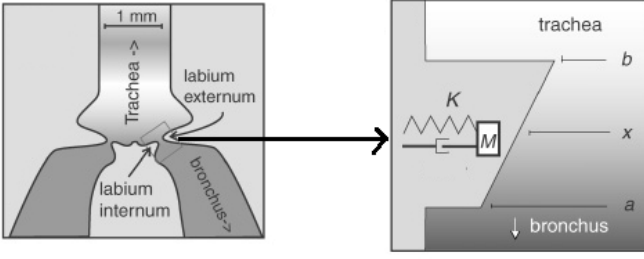


Figure 6: The structure of the avian vocal organ (syrinx). Image taken from [6]

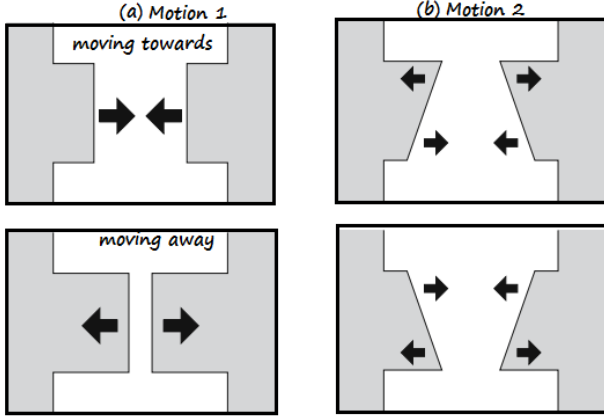


Figure 7: The two modes of motion of the labia. Image adapted from [4]

scribing the sequential and concurrent motion of the labia in each junction can be found in [1, 4].

This system is a slight variation of the Van der Pol Oscillator. Different parameterizations of this system exhibit different behavior and generate vocalizations with different acoustic features. In order to analyze the dynamics of this system, we identify the fixed points and study the stability of the system around the fixed points using the linearization method. The system has a trivial fixed point at $x_* = 0, y_* = 0$ for all parameters k, b, p and d . Note that $k, b, p, d > 0$.

The Jacobian matrix is

$$J = \begin{bmatrix} 0 & 1 \\ -k - 2dxy & (p - b) - dx^2 \end{bmatrix}$$

plugging in the fixed point, we get

$$J_{(0,0)} = \begin{bmatrix} 0 & 1 \\ -k & (p - b) \end{bmatrix}$$

The characteristic polynomial is

$$\lambda^2 - \tau\lambda - \Delta = \lambda^2 - (p - b)\lambda + k$$

The eigenvalues are

$$\lambda_{1,2} = \frac{(p - b) \pm \sqrt{(p - b)^2 - 4k}}{2}$$

In analyzing this system, we are interested in finding out at what values of the parameters the system will exhibit sus-

tained oscillations instead of converging to the fixed point. The reason for this requirement is that sound is only generated when sustained oscillations are present. Otherwise no phonation will be generated by the syrinx. In that regard, we look for the parameters that would cause a Hopf bifurcation which manifests itself as purely imaginary eigenvalues. Thus, when $(p - b) = 0$ the system will switch into generating sustained oscillations.

The bifurcation diagram given in figure 8 summarizes the dynamics of the simple syrinx model in terms of the two main parameters that are controlled by the songbird. Regardless of what the labia elasticity is, unless the air pressure is strong enough to overcome the dissipation (b), the labia will be at rest. As soon as the air pressure from the songbird's lung gets strong enough, the oscillations begin. The frequency of these oscillations depend on the labia elasticity. High frequency oscillations will be observed as k gets larger.

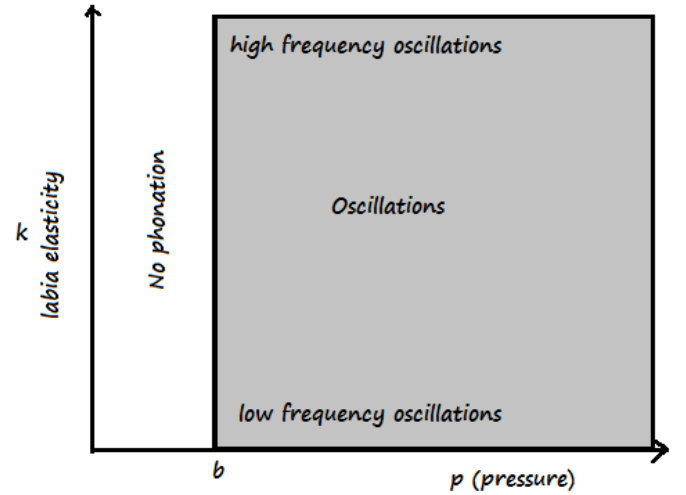


Figure 8: The bifurcation diagram of syrinx dynamics with respect to the two system parameters

A number of examples of system behavior are presented in figure 9. For regions in the parameter space such that $p < b$, no sustained oscillations occur. The phase portrait exhibit a stable spiral. The oscillation of the mid-points of the labia (x) settles down to the fixed point (no displacement) after a short time. As soon as $p = b$, the system enters the oscillatory state. In this case, the phase portraits exhibit the limit cycles. Note that the shape of the oscillating waves gets more complex (in terms of sub-tones) as p increases where all other parameters are kept constant. Also, as k increases while all other parameters kept constant, the frequency of the oscillations increase.

4.2 Syrinx Dynamics and Syllables

In the previous section, the presented model assumed that the variables k (labia elasticity) and p (pressure) are kept constant across time. This is hardly the case as the songbird vocalizes. As it is discussed in section 2, the frequency generally fluctuates within a syllable. In order to model a syllable in terms of the syrinx dynamics one needs to incorporate time-varying k and p variables instead.

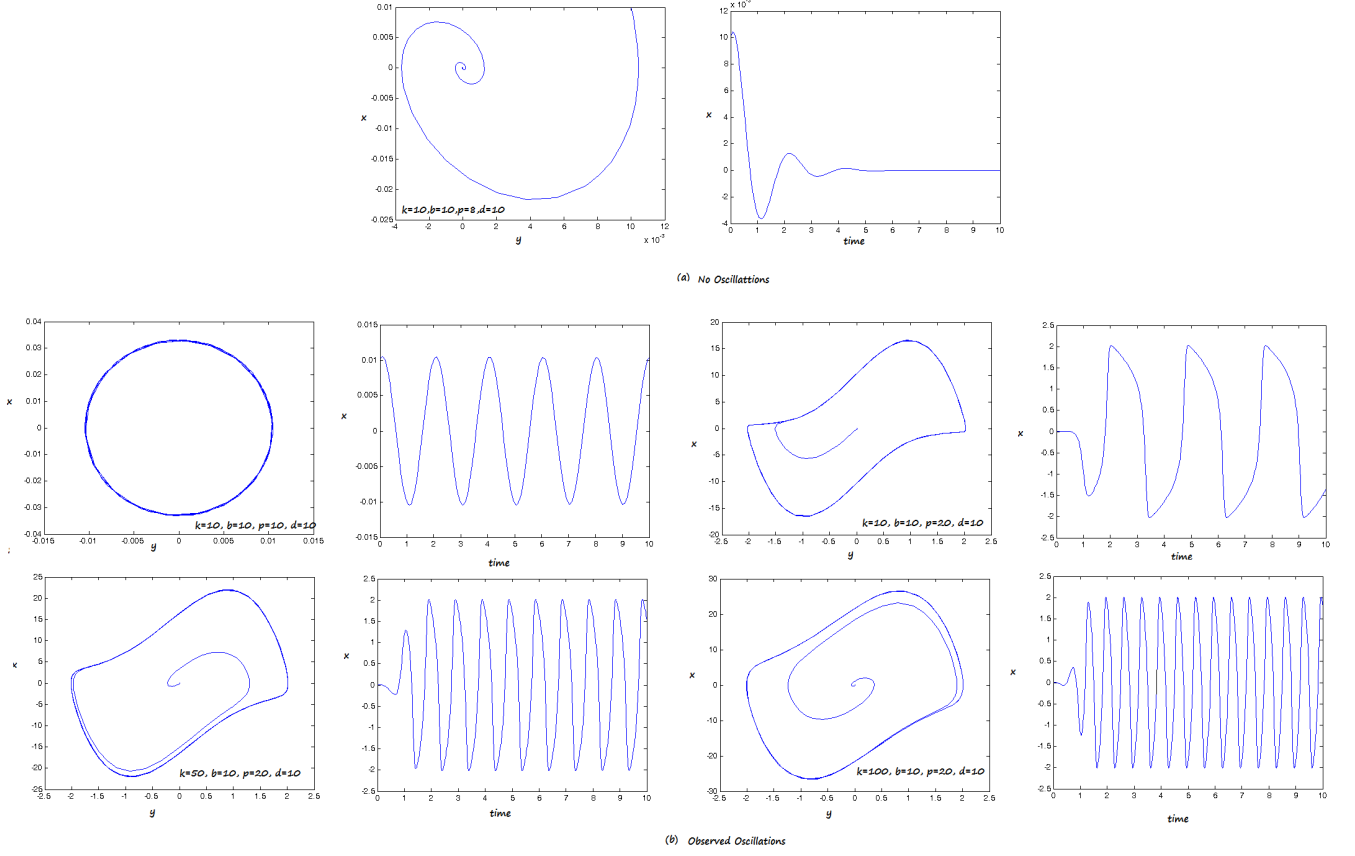


Figure 9: Various examples of the syrinx model. Oscillations occur where $p \geq b$ and the frequency of the oscillations increase as k increases.

$$\begin{cases} \dot{x} = y \\ \dot{y} = -k(t)x + (p(t) - b)y - dx^2y \end{cases}$$

It is possible to define directed paths in the p - k parameter space that would dictate how p and k will vary over time. Figure 10 shows one such path. The frequency changes evident in the sonogram are in direct agreement with our observation about the influence of k on the frequency.

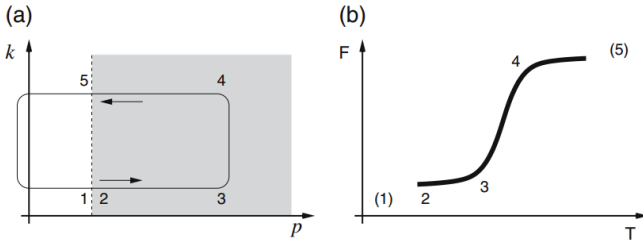


Figure 10: One possible path in p - k parameter space and the spectrum of the generated vocalization. Image taken from [4]

5. THE HIERARCHICAL NEURAL CONTROL OF SONG GENERATION

As it was argued in the previous section, the parameters governing the inner workings of the syrinx are time-varying so that the rich vocalizations can be achieved. Based on the finding from the neuroscience field that a temporal hierarchy is responsible for the birdsong generation, a number of researchers have attempted to build mathematical models of this neural hierarchy that would produce realistically sounding synthetic birdsongs. One such attempt has been reported in [2], where the authors model a population of the RA neurons using a central pattern generator given by

$$\begin{cases} \dot{x}_p = 30\{-x_p + S(\rho_1 + Ax_p - By)\} \\ \dot{y} = 30\{y + S(\rho_2 + Cx_p - Dy + \alpha x_k)\} \\ \dot{x}_k = 120\{-x_k + S(\rho_3 + Ex_k - \beta y)\} \end{cases}$$

where S is the sigmoid function. In this neuronal model, the ρ_i s are the excitatory inputs from the RA-projecting HVC neurons and x_k and x_p are the neurons that control the syringial and respiratory muscles respectively. The output of this neuronal model (figure 11) controls the two parameters for songbird vocalization: the labia elasticity (k) and the air pressure (p). The authors define paths in the p - k parameter space based on the excitatory input ρ_2 which creates a different syllable for each value.

A synthetic birdsong generated by using this technique is presented in figure 12. All parameter values were assigned

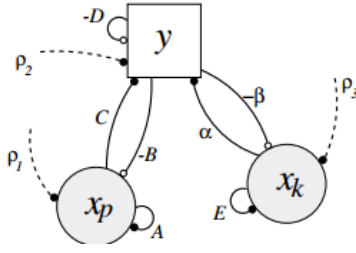


Figure 11: The central pattern generator used in [2] to model the RA region of the songbird brain

as specified by the authors in their paper and a short song consisting of the motif a-b-c-c has been generated using the model. As it can be seen on the sonogram, the syllable c is a complex two-part vocalization. This serves as an indication that this model can generate a rich set of syllables for mimicking birdsongs.

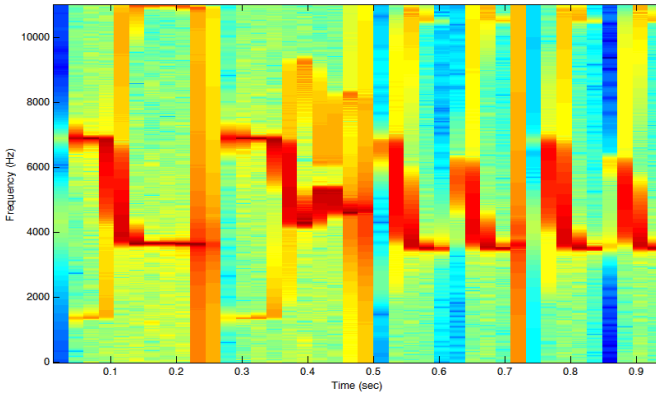


Figure 12: A synthetic birdsong generated using the model described in [2]. The song contains syllables a-b-c-c where $\rho_2 = -11.0$, -11.8 and -7.1 respectively

6. CONCLUSION

Dynamical systems theory has been a great tool for investigating how songbirds sing their songs. In this report, we provided a very brief overview of the main components, which are the anatomy and the dynamics of the vocal organ and the neuronal structure that controls the vocal organ. The models discussed here have been successfully used by various researchers. Although a number of variations to the basic models presented here have also been proposed over the years, the main techniques (such as bifurcation theory) to analyze and study dynamical systems remain widely applicable to these cases.

7. REFERENCES

- [1] Rodrigo Laje, Timothy J. Gardner, and Gabriel B. Mindlin. Neuromuscular control of vocalizations in birdsong: A model. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 65:051921–8:051921–8, 2002.
- [2] Rodrigo Laje and Gabriel B. Mindlin. Diversity within a birdsong. *Physical Review Letters*, 89:288101–1:288102–4, 2002.

- [3] David M. Michael. Musical automata as dynamical systems. *Adaptive Behavior - Animals, Animals, Software Agents, Robots, Adaptive Systems*, 14(2):139–146, June 2006.
- [4] Gabriel B. Mindlin and Rodrigo Laje. *The Physics of Birdsong*. Series in Biological and Medical Physics, Biomedical Engineering. Springer, 2005.
- [5] Fernando Nottebohm. Birdsong's clockwork. *Nature Neuroscience*, 5:925–926, 2002.
- [6] M. A. Trevisan and G. B. Mindlin. Philosophical transactions of the royal society a. *New perspectives on the physics of birdsong*, 367:3239–3254, 2009.
- [7] Albert C. Yu and Daniel Margoliash. Temporal hierarchical control of singing in birds. *Science*, 273:1871–1875, 1996.

APPENDIX

https://github.com/ilknuricke/birdsong_generator_CPG

```
%syrinx.m
%The bird syrinx model used in
%From Laje and Mindlin(2002), Diversity within a Birdsong, Physical Review
%Letters
%pt = interlabial air pressure
%k:labia elasticity
function yprime = syrinx(t,y,flag,tra,p,k)
b=1000 %(dissipation-friction coefficient) parameter taken from the paper
d=power(10,8) %(nonlinear dissipation coefficient for bounded motion) parameter taken from
the paper
yprime(1) = y(2)
pt=interp1(tra,p,t,'spline')
kt=interp1(tra,k,t,'spline')
yprime(2)=(pt-b)*y(2) - kt*y(1) - d*(power(y(1),2)*y(2))
yprime = yprime'

%birdbrain.m
%The CPG (Central Pattern Generator) Neural Network
%Modeling the RA brain region of a songbird
%From Laje and Mindlin(2002), Diversity within a Birdsong, Physical Review
%Letters
%the parameter rho2 is varied where all other parameters are fixed in
%generating birdsong syllables
function yprime = birdBrain(t,y,flag,rho2)
rho1=0
rho3=6
A=10
B=10
C=10
D=-2
E=4
alpha=2
beta=20
yprime(1)= 30 * ( y(1) + ( 1 / ( 1 + exp(-1 * (rho1 + A * y(1) B * y(2)))) ) ) %xp
yprime(2)= 30 * ( y(2) + ( 1 / ( 1 + exp(-1 * (rho2 + C * y(1) D*y(2) + alpha * y(3)
)))) %y
yprime(3)= 120 * ( -y(3) + ( 1 / ( 1 + exp(-1 * (rho3 + E* y(3) beta*y(2)))) ) ) %xk
yprime = yprime

%ingsyllable.m
function [t,song] = singSyllable(rho2)
Fs = 22050; %sampling
t = 0:1/Fs:0.24;
[t1 y] = ode15s( birdBrain ,t,[0.01 0.01 0.01],[],rho2);
kt=1.4*power(10,9)*y(:,3) + 4.8 * power(10,8);
pt=7000*y(:,1) 2200;
[m z1] = ode15s( syrinx ,t,[0.01 0.01],[],t,pt,kt);
song=z1(:,1);

%createSong.m
clear all;
Fs = 22050; %sampling
%syllable a
disp(Generating Syllable a (rho2=-11));
[m1 s1]=singSyllable(-11.0);
disp(Syllable a generated);
%play the syllable
sound(s1(:,1),Fs);
%syllable b
disp(Generating Syllable b (rho2=-11.8));
[m2 s2]=singSyllable(-11.8);
disp(Syllable b generated);
%play the syllable
sound(s2(:,1),Fs);
%syllable c
disp(Generating Syllable c (rho2=-7.1));
[m3 s3] = singSyllable(-7.1);
disp(Syllable c generated);
%play the syllable
sound(s3(:,1),Fs);
%stich the song up
disp(Stiching the song up a-b-c-c);
song = [ s1(:,1) ;s2(:,1) ;s3(:,1) ;s3(:,1)];
disp(Saving wav file );
wavwrite(song,Fs,mindlin2002_song.wav );
disp(Playing wav file );
sound(song,Fs);
```