APPLICATIONS OF BIOACOUSTICS TO MUSICAL INSTRUMENT TECHNOLOGY

MODELS FOR SOUND SYNTHESIS AND MUSICAL CONTROLLERS BASED ON ANIMAL SOUND PRODUCTION MECHANISMS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MUSIC
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Tamara Smyth
March 2004

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Julius O. Smith III
(Principal Adviser)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Neville Fletcher

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jonathan Abel

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jonathan Berger

Approved for the University Committee on Graduate Studies:

Abstract

In this research, the scientific and artistic contributions of animal sound production—or *bioacoustic*—systems to the field of music technology is explored through the development of musical controllers and real-time physics-based synthesis models. Though bioacoustic and musical acoustic systems share similar physical elements, such as vibrating plates and membranes and acoustic tubes and cavities, the coupling and functioning of these elements as acoustic and mechanical resonators can be markedly different. Two such examples that have become the focus of this work are the cicada's efficient buckling tymbal and the bird's vocal organ, the syrinx.

The cicada is of interest because of its ability to produce very loud sustained tones in spite of its small size. A physical model was developed of the cicada's tymbal—a mechanical device with a series of four ribs and a vibrating plate—and its coupling to the abdominal air sac—an acoustic or Helmholtz resonator. The model was successful both in producing the desired sound and in revealing the uniqueness and efficiency of the cicada's sequential buckling mechanism as a means of exciting a mechanical resonator.

A scaled mechanical model of the tymbal was then constructed as an input device to the computer synthesis model to determine whether the buckling rib mechanism could be used effectively by a human. The controller has four spring-loaded ribs, each equipped with sensors to obtain the energy generated when a rib is buckled by the downward force of a player's finger. Computer software was developed to receive the impulsive signal associated with the buckling rib and to input it directly into the computer model. One of the benefits of this controller is that sound can be manipulated using a single gesture rather than changing an overwhelming number of knobs and buttons. A trained percussionist who was videotaped while using the device found that the *upbuckling*—which produces a sort of rebound effect—allowed the ribs to be played very rapidly and with an efficiency that is very helpful to the percussionist's technique.

The second bioacoustic system, the songbird's syrinx, is of particular interest because, though it resembles some wind instruments and even more the human larynx, it has a structure that allows the bird to produce extremely virtuosic song, rapidly switching from low to high registers and often playing both simultaneously. The sounds produced by the syrinx, ranging from pure flute-like tones to a discordant cacophony, have clearly been a musical inspiration to musicians throughout history and therefore any study on the applications of bioacoustics to music would be incomplete without

including the syrinx.

The bird's airway consists of a trachea which divides into left and right bronchi at its base, and a membrane forming a valve at the top of each bronchial lumen. A physical model of the bird's vocal tract was developed following an article published by Neville Fletcher that provided a detailed quantitative description of birdsong. The model presented in this work uses waveguide synthesis techniques for the bronchi and trachea tubes and finite difference methods for the nonlinear vibrating syringeal membranes. Problems of numerical integration such as bringing sampling rates down to an audio range and reducing aliasing by "feathering" the collisions of the beating reed-like membrane were solved.

The model's produced sound is modified in real-time by changing two important continuous parameters: the air pressure from the bird's lungs and the tension of the syringeal membrane. The final contribution of this work presents a method for extracting these parameters from recorded birdsong so that the model can be more easily controlled. A look-up table pairs combinations of pressure and tension parameters with the model's corresponding output power spectra. At each time frame, a generalized likelihood ratio fills a pressure-tension matrix indicating similarity between the birdsong power spectrum and the tabulated spectra. Successive pressure-tension matrices are stacked and points exhibiting a good fit to the data align to form trajectories corresponding to changes in pressure and tension over time. These trajectories can then be used to control the model. In the event a range of trajectories matches the data well, the one selected is that of least action. This aspect of the research serves two purposes: 1) to judge the model's ability to produce actual birdsong and 2) to restrict the parameter space and improve the user's ability to interact with the model by, for example, having a controller that follows predetermined trajectories through the matrix.

There are three main aspects of instrument design being addressed in this research: 1) sound production, which involves the use of physical modeling techniques to develop sound synthesis models, 2) sound control, the building of a mechanical haptic human interface that allows for intuitive user input and 3) mapping, which determines relationships between the input of the performer, the synthesis parameters and the intended produced sound. The bioacoustic umbrella under which this research is conducted aims to expand the availability of quality musical sounds by searching for models that go beyond traditional musical instruments.

Preface

This work is very interdisciplinary in nature and is the culmination of the many activities and interests since, and prior to, beginning my doctoral studies at Stanford University's Center for Computer Research in Music and Acoustics (CCRMA). Though its greatest substance delves into the more technical aspects of musical instrument development involving signal processing and a scientific understanding of sound production and bioacoustics, *music* is the thread directing all of its parts.

Covered in this dissertation are many important issues of *sound production* and *sound control* involved in developing musical instrument technology. They include acoustics, synthesis techniques—with particular attention to physical modeling—input devices and parameter estimation methods for controller mapping. Bioacoustics, the umbrella under which these ideas are developed, serves to expand the ideas of musical sound beyond existing traditional music instruments.

Designers and researchers have frequently sought inspiration from natural systems to solve problems which arise in technology. This work continues this trend for the purpose of developing music technology so that it may evolve in a way that makes it as rich and lasting as its acoustic counterpart.

Acknowledgments

I thank my advisor Julius Smith who has been a tremendous source of guidance since the day I inquired about the PhD program at CCRMA. Julius was also very supportive when I chose this less conventional topic for my dissertation research and for that I am very grateful. I am also very much indebted to Jonathan Abel, both as a friend and a scholar, whose interest, enthusiasm and involvement greatly contributed to the completion of this work. Many thanks to the members of my oral defense committee, David Beach, Julius Smith, Jonathan Berger, Chris Chafe and Jonathan Abel, all of whom had an impact on my experience and my progress while at Stanford. In addition, I am very grateful to Neville Fletcher for reading this work (though distance prevented him from attending the defense) and for taking the time at conferences and via email to discuss issues related to acoustics and my dissertation topic. Thanks to the teaching assistants at Stanford University's Product Realization Lab (PRL) who provided endless help and suggestions during the development of the *Tymbalimba* controller. I extend this thanks also to Bill Verplank, the instructor for the Human Computer Interaction class at CCRMA, who also provided many insightful comments on the *Tymbalimba* and generously gave up much of his time to help make video demonstrations on the controller in action. Thanks to Brian Alegant, a former professor at McGill University, for his musical guidance and his encouragement when I applied to Piano Performance at McGill. I would also like to thank Dan Levitin who generously provided me with office space at McGill University during my summer stays in Montreal so that I could continue working on my research. Many thanks to my very dear friends, Tommy Babin, Kim Smyth, Agnieszka Roginska, Stefania Serafin, Patty Huang, Niki Bern, Debbie Driscoll, Adam Hacking, Martin Jolicoeur, Pamela Greenwell, Julien Bensa and to my family, Mom, Dad, Kim and Jeff, for their encouragement, friendship and support. Thanks to the folks at CCRMA and the Music Department who made my stay in California a rich and unforgetable experience. A final thanks to Pete "the pest control guy" whose help is probably best left unmentioned but whose words I am likely never to forget: "It's probably best to look away ma'am, this ain't gonna be pretty".

Contents

List of Tables

List of Figures

Chapter 1

Animal Acoustics and Musical Instrument Technology

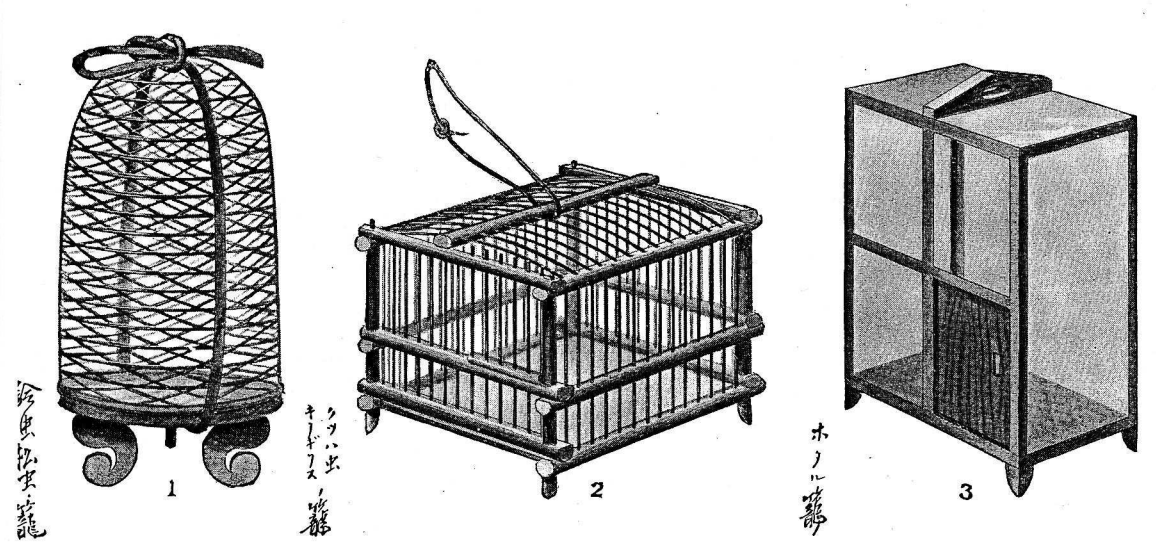
The idea of looking to nature for musical and scientific inspiration is not new. So many ideas, concepts and theories in science, technology and the arts, have been advanced by observing and understanding how things work in our natural environment. It seems only consistent therefore, that developments in a field where music and technology are so closely linked, might also benefit from this knowledge.

Sound produced by animals has been appreciated by many listeners and has been a source of inspiration to many musicians (in western culture and beyond). In 1898, Lafcadio Hearn published an essay called "Insect-Musicians" in a book entitled *Exotics and Retrospectives* [85]. In this essay he speaks of a temple-festival in Japan called *en-nichi*, where he finds the booth of a vendor of singing insects. The booth was "stocked with tiny wooden cages out of which an incomparable shrilling proceeds". The sound of which he speaks is made by insects.

At the time of Hearn's writing, the song of insects was as much respected in Japanese culture as that of thrushes, nightingales and canaries in Western civilization [85]. To the Japanese, "there is as much difference in the notes of insects as there is in those of larks and sparrows" [85]. Insect musicians were so appreciated that it was customary to visit the countryside on the weekends to listen to their songs. Eventually the insect-musicians were captured and bred for the insect-trade making this trek no longer necessary. The Japanese were willing to pay considerable sums to keep these "short-lived insect-pets" in cages within the home. Typical cages used to contain these singing insects are shown in Figure 1.1.

According to Hearn, the first definite mention of cages for singing-insects appears in a work by *Chomon-Shū* [85, page 43]:

> On the twelfth day of the eighth month of the second year of Kaho, the Emperor ordered his pages and chamberlains to go to Sagano and find some insects. The Emperor



1. A FORM OF INSECT CAGE. 2. CAGE FOR LARGE MUSICAL INSECTS, — *Kirigirisu, Kutswamushi, etc.*
3. CAGE FOR SMALL MUSICAL INSECTS, OR FIRE-FLIES

Figure 1.1: Typical cages for keeping musical insects [85].

> gave them a cage of network of bright purple thread. All [...] went on horseback to hunt for insects. [...] In the evening they returned to the palace. They put into the cage some *hagi* and *ominameshi* [for the insects]. The cage was respectfully presented to the Empress. There was *saké* drinking in the palace that evening; and many poems were composed.

In Japan it was common to listen to the sound of certain insects as though it were music. Insect-musicians would typically be present at parties and social gatherings and both poetry and music would be written while listening to their song. Some examples of insect-musicians are illustrated in Figure 1.2. The *matsumushi* (top left of Figure 1.2) was much esteemed for the peculiar clearness and sweetness of its notes. Hearn describes this sound as "resembling that of an electric bell heard from a distance" [85]. The *suzumushi* (top right of Figure 1.2), whose name signifies "bell-insect", sounds like a very small bell (or a bunch of little bells) but can also be mistaken for the sound of rapids [85]. The *kanētataki* (bottom left of Figure 1.2) is also known for its bell-like sound. The *kirigirisu* (bottom right of Figure 1.2) is a much-prized insect fetching very high prices. "They are large vigorous insects, uttering very clear notes" [85] and are said to utter the Japanese words *"Tsuzerē—sasē! sasē!"* (Torn clothes—patch up! patch up!).

The purpose of this essay, as stated by Hearn himself, is to show how superficially travelers to Japan might unconsciously judge the most interesting details of Japanese life [85, page 41]. It also shows how something which can seem so bland and commonplace to one, may be imbued with

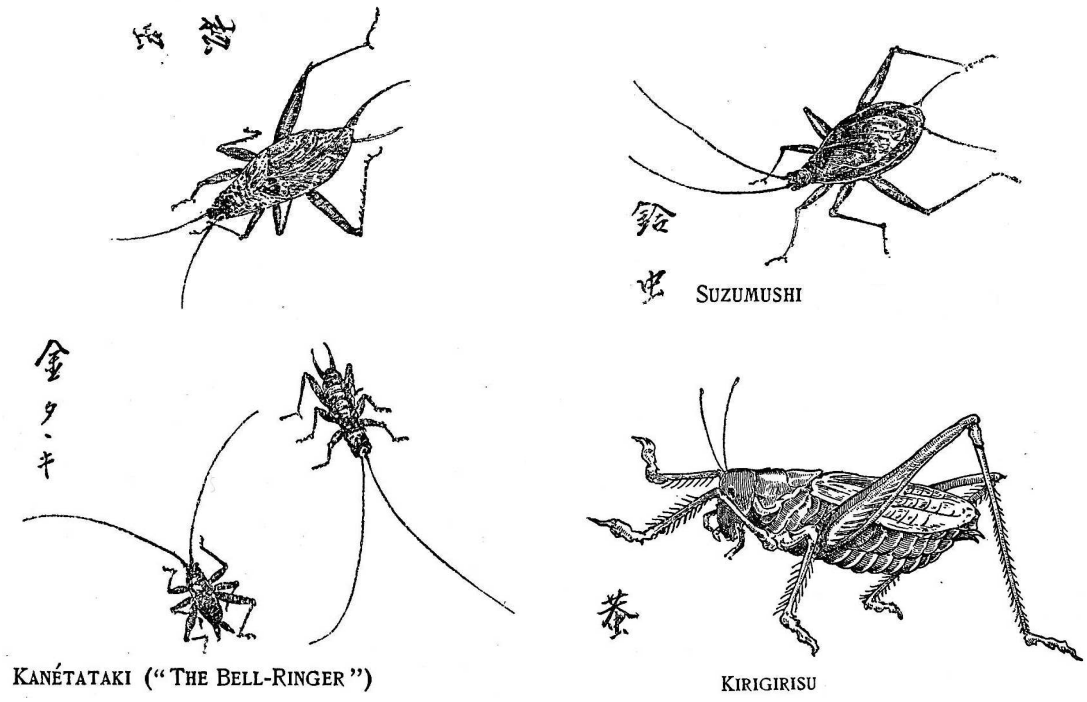


Figure 1.2: Popular insect musicians from top left (clockwise): *matsumushi*, *suzumushi*, *kanétataki* and *kirigirisu* [85].

influence and richness to another. What attracted the Japanese listener to the insect's song? What music did they hear? "The national liking for caged insects does not mean a liking for mere noise; and the note of every insect in public favor must possess either some rhythmic charm, or some mimetic quality celebrated in poetry or legend." [85, page 42].

In many civilizations, past and present, the vocalization of insects, birds and other animals has a tremendous luring quality, with the ability to captivate listeners in different ways. The purpose of this research is to look at the mechanism by which this sound is made (concentrating on the sound production mechanisms of both the cicada and the songbird) to determine if it too can influence various aspects of computer music instrument technology. If the sound has been deemed musical, perhaps the sound production mechanism too has musical applications.

1.1 Musical Instrument Building Blocks

There are many musical sound sources in our environment besides those produced by traditional musical instruments. What is lacking for musicians is an effective and intuitive means of controlling them. As a result, many composers have resorted to sampling and non-real-time processing of these

sounds in order to achieve something close to their desired effect. It would be preferable however to provide musicians with the ability to manipulate these sound sources in a more natural and meaningful way—much in the same way a musician plays a traditional acoustic musical instrument.

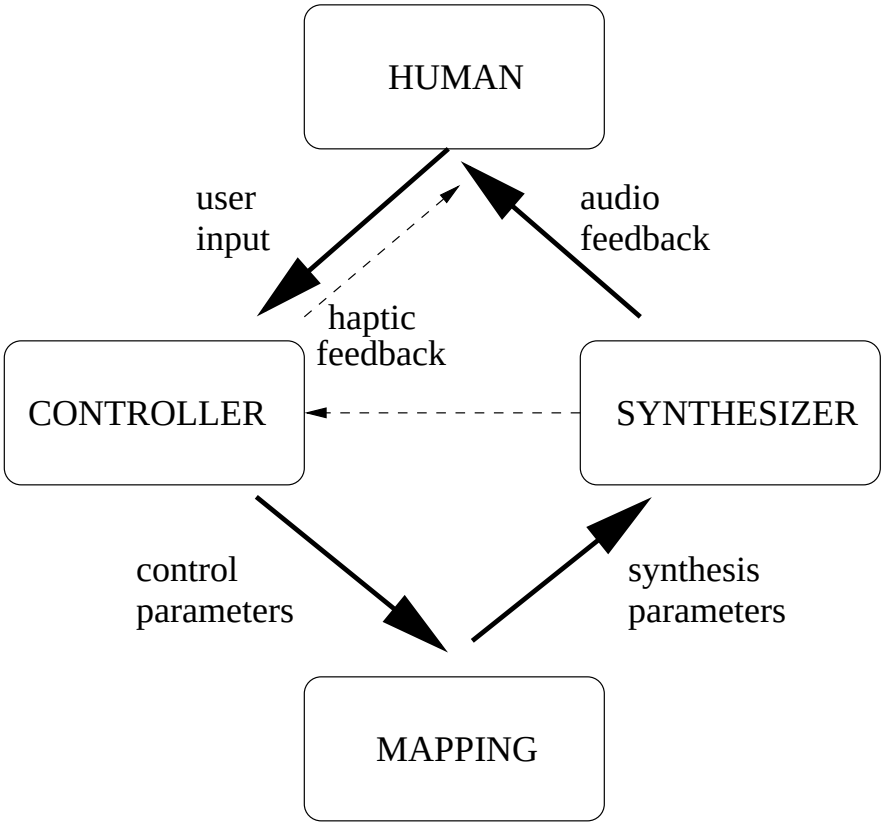


Figure 1.3: The building blocks and signal flow of a computer/electronic musical instrument.

Computer music instrument technology—in the scope of this work—refers to the development of sound synthesis and computer sound generation, interfaces for human interaction (or *controllers*) and a mapping between the two, linking human input to the synthesis control parameters. As illustrated by the diagram in Figure 1.3, a human user directly interacts with a *controller* (or input device) which, through some predetermined *mapping*, causes changes in the parameters of a *synthesizer*. The synthesizer changes the produced sound according to the synthesis parameters corresponding to the user's input and provides audio feedback to the user—usually through loudspeakers or headphones. Synthesizers may also convey information back to the user through adjustments made to the touch or *feel* of the controller. This process, referred to as haptic feedback (represented by the dashed line in Figure 1.3), has been deemed a very important aspect of musical instrument interaction [64]. It occurs naturally in most acoustic musical instruments because of the close integration between the mechanics of sound *production* and sound *control*. Perhaps one of the reasons why

electronic/computer instruments tend not to survive the evolutionary process of natural selection is because this aspect of human-instrument interaction is often overlooked.

1.1.1 Synthesizer

A musical instrument requires a means for producing musical sound. In computer/electronic music instruments this is done using synthesis algorithms (or models) with input parameters allowing for change in the produced sound over time. The quality of the model is usually determined both by the character of the sound and its response to human input.

The developer of synthesis instruments endeavours to make them rich in possibilities yet not so complex they are impossible to play. Naturally, computer instruments will be compared to their acoustic counterparts which have established excellence in sound production and playability and which musicians have devoted countless hours practicing, playing and perfecting. This is perhaps one of the reasons it is so difficult to look beyond traditional instruments when trying to design new ones.

Because of its richness in both sound quality and input parameters, physical modeling [11] has become a very popular synthesis technique. As the name suggests, it attempts to capture the physical behaviour of an existing system through a series of equations that are discretized and run from within a computer. A possible drawback of this otherwise successful synthesis technique is its reliance on an existing model which often tends to be limited to existing musical instruments.

Physical modeling has been, in large part, a technique developed by engineers and used very little by composers and musicians. This is perhaps because physical models tend not to stray too far from scientific reality making composers and musicians unable to expand their existing tools (one of the reasons they would likely be interested in computer music in the first place). Though it is very worthwhile to understand and reproduce existing musical instruments, if the purpose is to evolve the direction of musical instrument technology, it may be beneficial to search for models beyond existing traditional instruments.

Though the physical models described in this work are based on existing acoustic systems, they are not systems with which humans typically interact. They are models of animal sound production mechanisms and in particular, the sound mechanism of the cicada and the songbird.

Chapter 2 presents a model of the cicada's use of a vibrating tymbal plate coupled with an abdominal air sac. The model uses a series of biquadratic resonant filters for each of the modes of the vibrating tymbal plate and the abdominal acoustic resonator. The cicada uses a rapid sequence of buckling ribs to initiate and sustain vibrations in the plate. This unique and interesting aspect of the cicada's mechanism is modeled using actual data provided by Henry C. Bennet-Clark [81].

In Chapter 4, a physical model of the song bird's vocal organ, the *syrinx* is presented. This model, which is considerably more involved than that of the cicada, uses waveguide synthesis [11] for the bronchi and trachea tubes and finite difference methods for the pressure-controlled valve

formed by nonlinear vibrating membranes. Like the cicada model, it runs in the real-time computer environment Pd [88] and is equipped with control parameters allowing the user to interact with the model and manipulate the sound over time.

Though most successful when physics is used as a base or starting point, computer models are far less constrained by scientific reality than acoustic instruments. The numerous parameters with which they are equipped offer a wide variety of possible sounds by extending and exploring the model.

1.1.2 Controller

The synthesizer, and the ability to play sounds from loudspeakers, has given rise to the need of a device which will allow human interaction with the produced sound from which s/he has effectively been disconnected [62]. Much like the sounds created by physical modeling synthesis, developers of music controllers have not really looked beyond traditional musical instruments for design ideas. Midi keyboards are perhaps still the most popular controller for synthesizers, yet midi wind and string controllers are also readily available to instrumentalist who, when experimenting with new sounds, wish to apply the technical expertise they have spent so many years refining.

The disconnection of sound source from sound control can however, also be advantageous since it is no longer necessary to adhere to older musical instrument paradigms. Like computer synthesis models, controllers are also less constrained by physics, allowing the player to be free of large, burdensome, unergonomic designs that are made solely because they are required by the acoustics of the sound production mechanism (a small hand held input device for example, is as capable of controlling the low thunder of a tuba as it is the higher pitches of a piccolo). The controller can devote its design exclusively to catering to the needs of the player. In spite of this, little has been done to develop mechanisms which will actually help players overcome their technical hurdles.

When the sound source is no longer the controller (as it is in acoustic musical instruments) it can also be a tremendous disadvantage to the player who loses valuable information normally acquired by touching a resonating instrument. When the player interacts directly with the sound producing elements, their physical states change in a way that makes their playability more of less difficult (or simply different). This information is processed by the player to gain mastery over his/her instrument, allowing him/her to use subtle or more drastic changes in gesture to alter the produced sound. Ideally controllers should also capture human gesture (rather than discrete on-off triggers) but this is made more difficult by the separation between player and sound source. It becomes necessary to measure and translate human motion so that it can be linked to the sound engine through some predetermined mapping.

In Chapter 3, a mechanical controller called the *Tymbalimba* is presented, offering a natural *haptic* feedback to the performer and allowing notes to be played in extremely rapid succession. The purpose of the controller—based on the cicada's efficient sequential buckling mechanism—is

to determine whether the mechanism which allows the cicada to use discrete impulses to sustain vibrations of the tymbal plate, can be used as effectively by a human playing a musical instrument.

1.1.3 Mapping

Mapping is a difficult issue in developing musical instruments and is rarely adequately addressed—perhaps because it is the one building block that cannot so easily duplicate existing acoustic instruments where the mapping usually falls out of the instrument design *naturally*. In acoustic instruments, the distinction between the sound generator and the sound controller is harder to see [62], and therefore the task of creating a link between the two is practically non existent.

It should not be assumed (nor required) that the performer have the same knowledge of the physics of the synthesis model as does the developer. It is a rare performer that can explain in technical terms the acoustics governing the sound production of his/her instrument. This level of detail is simply not necessary to perform on the instrument successfully. Rather, it is sufficient for the instrumentalist to be aware of how his/her physical motions effect the instrument and the resulting produced sound (this, of course, can be enforced by a deeper understanding of the instrument's physics but it does not *dependent* on it). Because there is such a tight connection between sound and physical gesture, this ability can be developed through trial and error, intuition, training and significant good practice.

In software however, there is more effort involved in predetermining intuitive relationships between the input gestures of the performer, the synthesis parameters and the intended produced sound. Frequently, the easiest most intuitive approach for the developer does not produce satisfactory results for the performer. For example, a one-to-one mapping, where the perform must tweak a knob, press a button or perform some other action to alter one parameter is simple for the developer but often not very efficient or intuitive for the user of the instrument.

The links among intention, input and produced sound can be accomplished by making synthesis parameters functions of each other, where multiple parameters are controlled using a single gesture. Another approach may be to significantly limit the control parameter space in a way that won't necessarily remove functionality. A piano, for instance, effectively has two control parameters, pitch and amplitude (if pedaling is omitted). Yet a great deal of variety can be obtained simply by modifying these two parameters over time during a performance. In a synthesis model of the piano however, there are innumerable possible piano parameters that *could* be afforded to the pianist [4]. If these parameters are not managed through an intuitive mapping however, they will likely become overwhelming and burdensome to the performer.

In Chapter 6, a method for determining likely control parameters for the physical model of the syrinx is introduced. One of the more difficult tasks with physical models is determining the parameter values which will make it sound as the user intends. It is not until these parameters are explored that the quality of the model can really be determined. As a means of making the model

sound more birdlike (a starting point from which the player can then deviate), the model is calibrated to recorded birdsong to determine values over time for the two primary control parameters: pressure from the lungs and tension in the vibrating valve. With these trajectories in place, it is possible to determine the range of values for these control parameters, their rate of change, and which sound most bird-like. This information is very useful in guiding mapping strategies.

1.2 Acoustics, Biology and Music

In Florence, circa 1623, the Vatican imposed a censorship on Galileo, forcing him into a promise not to promote Copernican worldview. Refusing to be completely silenced, Galileo published a nontechnical book on science for a wider reading audience called *The Assayer*.

The following passage from *The Assayer* (taken from *Galileo's Daughter* [91]) is a parable showing Galileo's enchantment with natural sound and in particular with the song of the cicada. It also demonstrates the lengths to which one will go to satisfy an insatiable curiosity driven by the sounds of nature.

> Once upon a time, in a very lonely place there lived a man endowed by Nature with extraordinary curiosity and a very penetrating mind. For a pastime he raised birds, whose songs he much enjoyed; and he observed with great admiration the happy contrivance by which they could transform at will the very air they breathe into a variety of sweet songs.
>
> One night this man chanced to hear a delicate song close to his house, and being unable to connect it with anything but some small bird he set out to capture it. When he arrived at a road he found a shepherd boy who was blowing into a kind of hollow stick while moving his fingers about on the wood, thus drawing from it a variety of notes similar to those of a bird, though by quite a different method. Puzzled, but impelled by his natural curiosity, he gave the boy a calf in exchange for this flute and returned to solitude. But realizing that if he had not chanced to meet the boy he would never have learned of the existence of a new method of forming musical notes and the sweetest songs, he decided to travel to distant places in the hope of meeting with some new adventure.
>
> As the man roved, the encountered song made by a bow sawing upon some fibers stretched over a hollowed piece of wood, by the hinges of a temple gate, by a man rubbing his fingertip around the rim of a goblet, and by the beating wings of wasps.
>
> As his wonder grew, his conviction proportionately diminished that he knew how sounds were produced; nor would all his previous experiences have sufficed to teach him or even allow him to believe that crickets derive their sweet and sonorous shrilling by scraping their wings together, particularly as they cannot fly at all.

> Well, after this man had come to believe that no more ways of forming tones could possibly exist, he suddenly found himself once more plunged deeper into ignorance and bafflement than ever, for having captured in his hands a cicada, he failed to diminish its strident noise either by closing its mouth or stopping its wings, yet he could not see it move the scales that covered its body, or any other thing. At last he lifted up the armor of its chest and there he saw some thin hard ligaments beneath; thinking the sound might come from their vibration, he decided to break them in order to silence it. But nothing happened until his needled drove too deep, and transfixing the creature he took away its life with its voice, so that he was still unable to determine whether the song had originated in those ligaments. And by this experience his knowledge was reduced to diffidence, so that when asked how sounds were created he used to answer tolerantly that although he knew a few ways, he was sure that many more existed which were not only unknown but unimaginable.
>
> Galileo Galilei, *The Assayer*

Acoustic musical instruments seem to be the epitome of perfection in human control over sound. They have controllers which, in their various mechanisms, are extremely responsive to input and they have an ideal sound quality, void of artifacts characteristic of band-limited signals. It is no wonder it is so difficult to expand upon our current repertoire of musical instruments (one of the last widely used musical instruments being developed by Adolph Sax a whole century and a half ago).

It is not however, only in human craftsmanship that such excellence in music technology exists. There are other acoustic systems that possess these musical qualities such as the instruments and mechanisms used by animals, including humans and insects, to produce sound. It is not necessary to look farther than one's backyard (quite literally) to hear that nothing requisite is wanting in natural sounds—the sounds produced by the many animals in our environment, including our own voices. Like musical instruments, these sound production mechanisms are also very closely intertwined with the way in which they are controlled (with physical control elements directly modifying the sound producing elements).

Aspects of biological material such as shape, density and elasticity are not as clearly defined as the materials in musical instruments (though even in this case much of the material is selected empirically) [31]. In addition, animals and insects must often compensate for the lack of size with a greater use of nonlinearity [59], adding a complexity to their instruments that can make them more difficult to control. Pitch and intensity, as seen in the syrinx for example (see Chapter 4) are never independent.

Yet is there something that can be learned from bioacoustic sound production mechanisms that can be directly applied or partly adapted for the purpose of developing musical instrument technology? All three aspects of instrument design as outlined in Figure 1.3 are addressed in subsequent

chapters to respond to this question, considering two bioacoustic systems in particular, the cicada's tymbal and the bird's syrinx.

Chapter 2

The Sound of the Cicada

It was discovered in the Amanab district of Papua New Guinea that melodies could be generated by coupling the mechanical resonance of a buzzing beetle to the changing resonance of the mouth. By attaching the insect to a stick and holding it in front of the mouth, certain harmonics could be selected by changing the volume and shape of the vocal cavity—a technique similar to playing a Jew's harp.

2.1 The Sound Mechanism of the Cicada

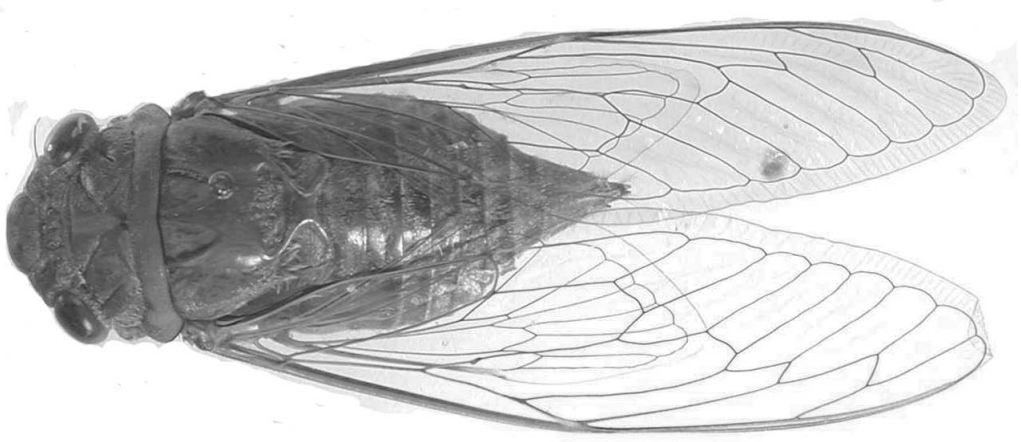


Figure 2.1: The dundubia species of cicada.

The cicada produces its characteristic loud song (the loudest measured from any insect [60]) using a very similar mechanism in which the energy of an extremely fast contracting muscle is converted to sound energy using both a mechanical and an acoustic (or Helmholtz) resonator [61]. In fact, the

coupling of mechanical and acoustic resonators is the method used by most musical instruments—including both strings and winds—to allow for the vibrations of a primary mechanical resonator to propagate through the surrounding air and reach the human ear.

The cicada's primary mechanical resonator is called the *tymbal* and there are two, one located on either side of the insect's abdomen (see Figure 2.2). The vibrating tymbals drive the *abdominal air sac* which acts as the secondary acoustic resonator. The pureness and strength of the song largely depends on the quality factor or Q of this secondary acoustic resonator. The large eardrums, or *tympana*, provide the aperture through which the insect's song may radiate. These main components of the cicada's sound producing mechanism, the tymbal, the abdomen and the tympana, are illustrated in Figure 2.2.

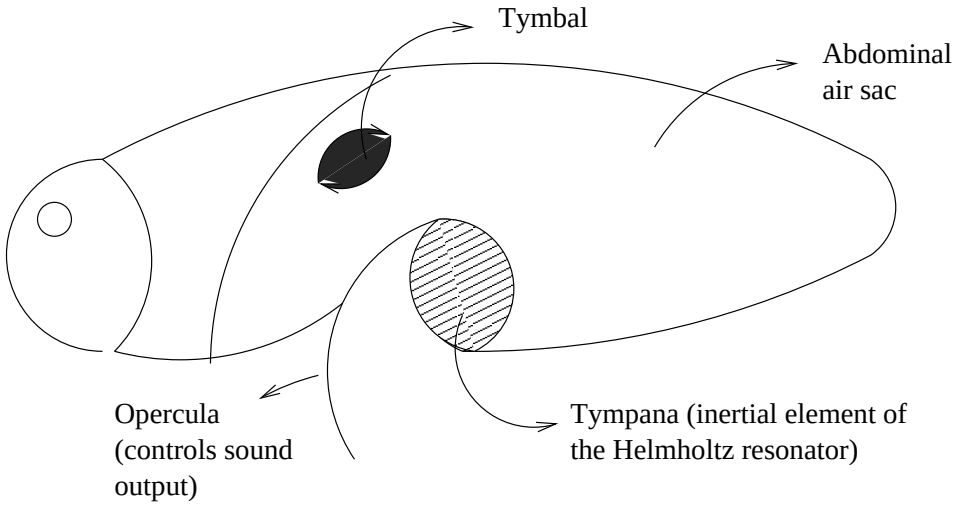


Figure 2.2: A simplified diagram of the cicada showing the location of its sound production elements: the tymbal (of which there are two—one on either side of the abdomen), the tympana, the opercula and the abdominal air sac.

2.2 Resonators

Many biological and musical sounds are produced by driven resonances: vibrations that are initiated and sustained by repetitive re-excitation at an appropriate frequency and phase [43]. Once there is an excitation mechanism in place (a mechanism by which energy is introduced), the fundamental frequency of oscillation of the electrical, mechanical or acoustic resonator, is determined by the interaction of reactive elements: an analogue of compliance (or elasticity) and an analog of mass.

In an electrical circuit (shown on the left of Figure 2.3), the capacitance is the compliant element and is analogous to the volume of a fluid filled cavity in the acoustic (or Helmholtz) resonator (shown in the center of Figure 2.3) and to the stiffness of a spring in a mechanical system (shown on the

right in Figure 2.3). Likewise, the inductance of the electrical system is analogous to the neck of the Helmholtz resonator and to the the mass in a mechanical resonator.

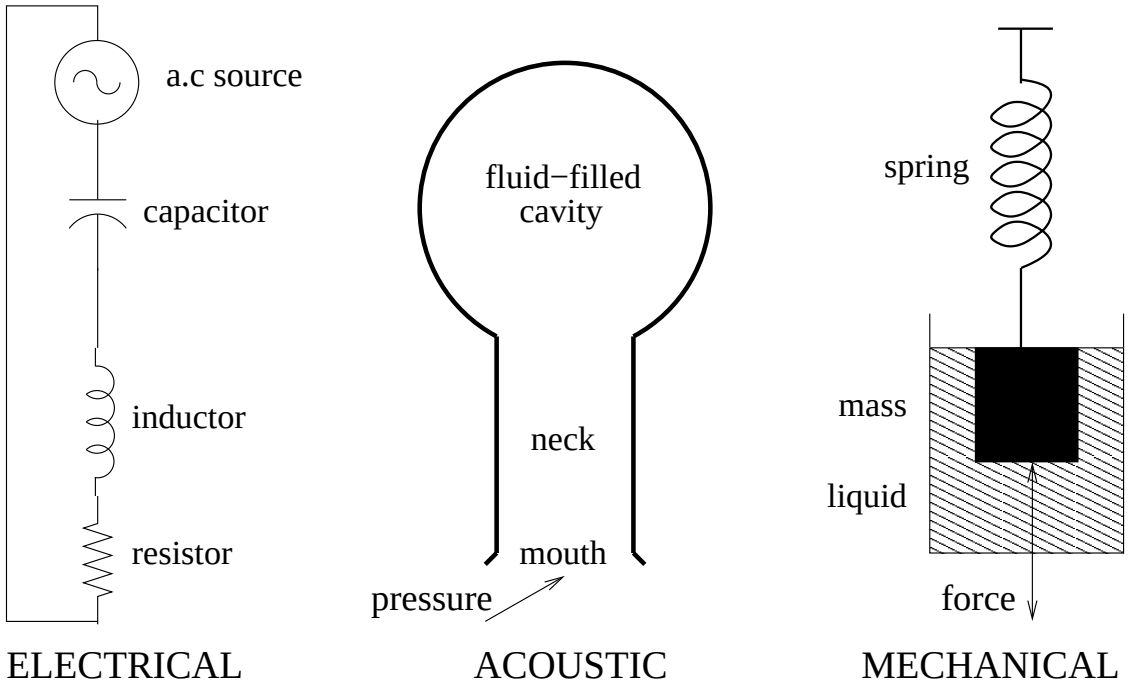


Figure 2.3: The fundamental frequency of an electrical, mechanical and acoustic oscillator is determined by the interaction between reactive elements: an analog of compliance (or elasticity) and an analog of mass [43].

2.2.1 Quality Factor or Q of the Resonator

In practical oscillators, the energy of oscillation is dissipated by a viscous or resistive load (as, for example, the mouth of the Helmholtz resonator) which causes the amplitude of oscillation to decay exponentially. The sharpness of the resonance is indicated by the quality factor or Q of the resonator. The Q of a resonant system is given by the ratio of the resonant frequency to the 3dB bandwidth [43], that is,

$$Q = \frac{\text{resonant frequency}}{\text{bandwidth at -3dB}}. \tag{2.1}$$

If the Q of the resonator is high, energy will be confined to a narrower band and the sound produced will be closer to a pure tone. Likewise, if the Q is low, short-lived impulses used to excite the resonator will not allow vibrations to reach and/or maintain the fundamental frequency long enough to be perceived as having a pitch. Rather, the resulting sound will be more percussive in nature.

2.3 The Tymbal as a Mechanical Resonator

The *tymbal* is the cicada's primary resonator. It is a small, hardened, dome-shaped membrane, equipped with a plate and a series of four long convex ribs (see Figure 2.4).

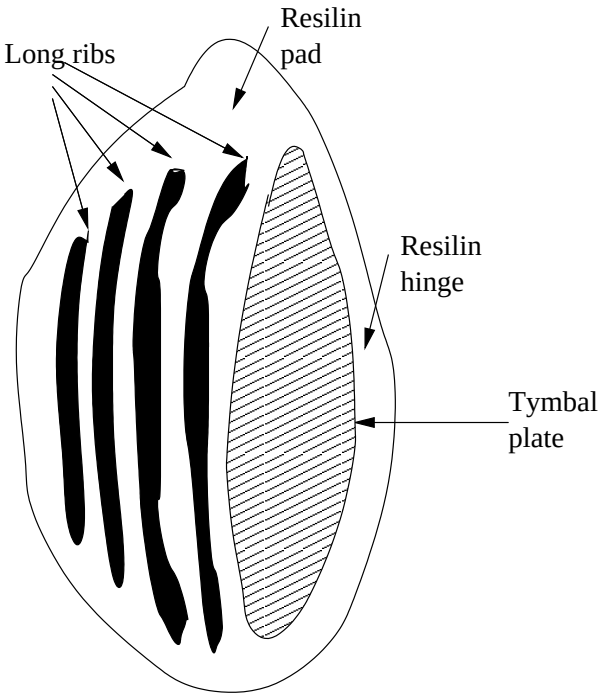


Figure 2.4: A simplified diagram of the tymbal showing the position of the four long concave ribs with respect to the tymbal plate.

When the cicada contracts its tymbal muscle, the hinged tymbal plate is pulled inward placing a downward force on the center of the first rib to which it is connected, until the rib collapses or *buckles* into a V-shape. Since all the ribs are connected with a springy organic resilin, when the first rib collapses, it pulls on its neighbouring rib and applies a similar downward force causing it also to collapse. This continues until up to all four ribs (depending on how much energy is released by the buckling of the first rib) in the sequence buckle, initiating and sustaining vibrations of the tymbal plate. A sequence of two buckling ribs as seen by the center curved region of the rib is shown in Figure 2.5.

When the first rib buckles, energy stored in the elastic resilin (the tymbal's compliant element) is released and the tymbal plate, along with the first buckled rib, begin to vibrate. The tymbal behaves like a mass and spring system (see Section 2.2), with the reactive elements being the mass of the tymbal plate (and any ribs that are buckled) and the stiffness of the tymbal's dorsal support [41]. The buckling of each individual rib contributes to the overall effective vibrating mass of the

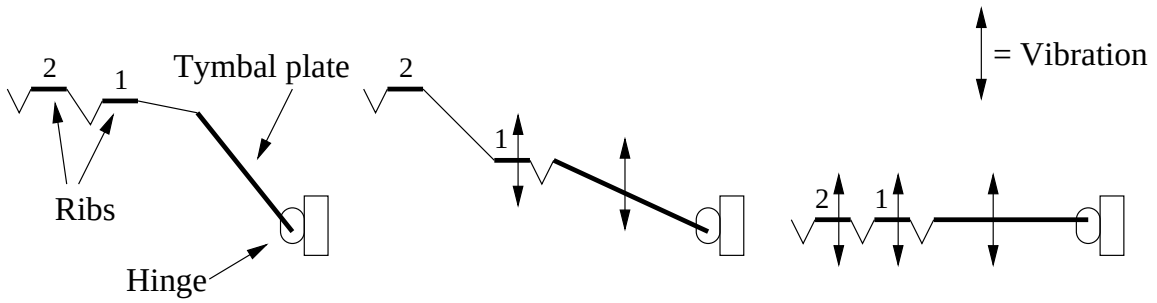


Figure 2.5: The center buckling region of two ribs during a buckle cycle [43]. The leftmost diagram shows the ribs and plate in their resting position. As the cicada contracts its muscle, the tymbal plate is pulled downward. In the middle diagram the first rib is buckled and in the rightmost diagram, two ribs are buckled.

tymbal plate during a buckle cycle (see Table 2.1).

This increase in the plate's mass after an inward buckling rib causes a decrease in the fundamental frequency of the vibrating tymbal according to the equation

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}}, \tag{2.2}$$

where k is the stiffness factor, and m is the changing mass. During a buckle cycle therefore, there is a sort of chirping effect created by a slight stepwise decrease in frequency (visible in Figure 2.6).

Tymbal Element	Mass (μg)	Total Vibrating Mass (μg)	Fundamental Frequency (kHz)
plate	550	—	—
load	380	—	—
1^{st} rib	100	1030	3.5
2^{nd} rib	85	1115	3.37
3^{rd} rib	80	1195	3.26
4^{th} rib	55	1250	3.18
out buckled ribs	550	550	6.54

Table 2.1: This table shows the mass of the individual tymbal elements and the increasing total mass of the vibrating tymbal plate after the buckling of each additional rib. The result is a slight decrease in the fundamental frequency of the resonator during one buckling cycle, creating a sort of chirping effect.

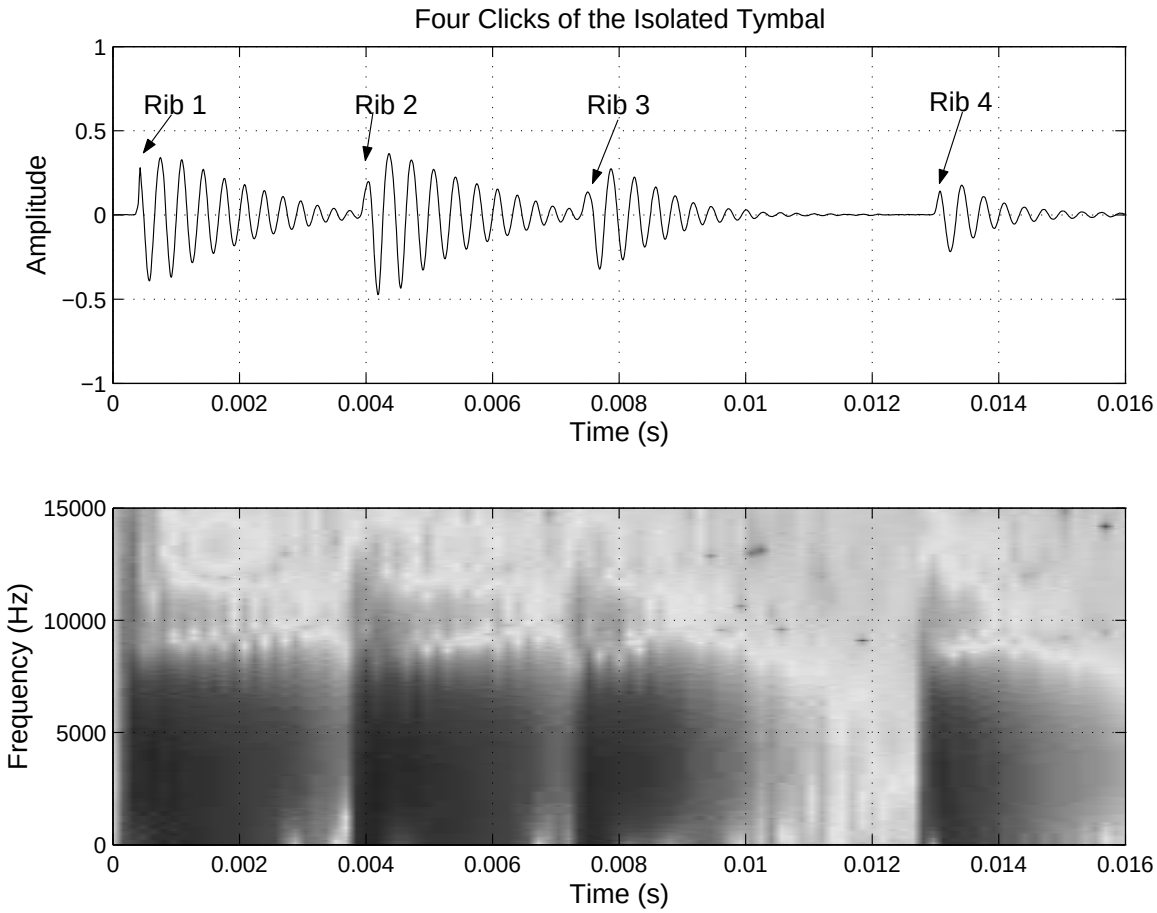


Figure 2.6: Measured signal (top) and its spectrogram (bottom) of one buckling sequence (data provided by [81]).

2.3.1 The Tymbal's Sequential Buckling Mechanism

The rapid sequence of buckling ribs is the sound excitation mechanism of the cicada. Buckling is a nonlinear phenomenon that results when a downward force applied to a convex shape causes it to spring into a concave shape (see Figure 2.7). When the force is no longer present, as is the case when the cicada relaxes its muscle, the rib is free to spring back to its original shape—ready to be buckled again.

Once the first rib buckles, energy is immediately introduced into the system and the tymbal plate begins to vibrate. The cicada can change the amplitude of this vibration by changing the curvature in the tymbal (using a tensor muscle) [44]. The greater the convexity of the tymbal, the greater the amount of energy required to buckle the ribs (which are also made more convex), and the more acoustic energy released as a result [44]. This ability of the cicada to change the geometry of the

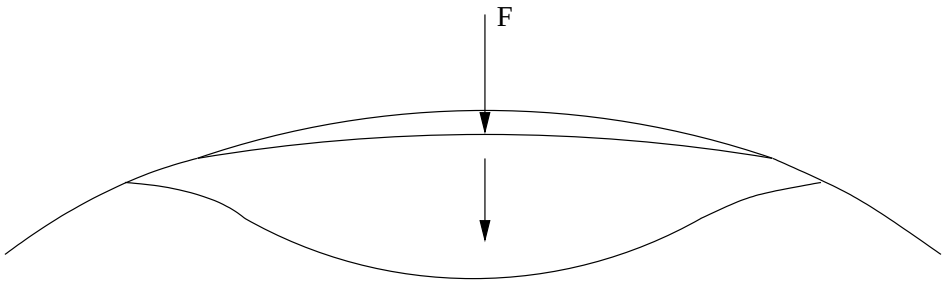


Figure 2.7: Inward buckling motion of a tymbal rib under a central force

ribs inspired the *stop* mechanism in the *Tymbalimba* (described in Chapter 3). The stops allow the user to change the feel (and thus the resulting produced sound) of the instrument by pulling out the ribs to flatten them or by pushing them in so their v-shape is more acute.

The efficiency of the sequential buckling mechanism lies in the fact that for one motion, that is, one muscle contraction, the cicada gets a sequence of up to four *IN* pulses (a result of the inward buckling of four ribs) along with the same number of *OUT* pulses (a result of the rebound outward buckling). This makes up to eight bursts of energy that sustain vibration of the tymbal. It is, in part, this efficient use of energy that allows some species of cicada to use a rapid sequence of discrete impulses to produce a perceptually sustained tone.

The cicada has two tymbals located on either side of its abdomen which can be controlled independently by the cicada [61]. Free vibrations of each tymbal decay according to (2.1). Both tymbals drive the abdominal air sac which serves to smooth asymmetries between the individual tymbal vibrations, filter the frequencies to a narrower band, amplify the sound and improve sound radiation through the abdomen's tympana.

2.4 The Abdominal Air Sac as a Helmholtz Resonator

The tymbal drives a secondary abdominal air sac which can be modeled as a Helmholtz resonator. As shown in Figure 2.3, a Helmholtz resonator consists of an enclosed volume, an open neck and a mouth through which the sound may propagate [47]. The reactive elements of the abdominal air sac, the compliance of the abdomen's volume and the inertance of the tympana or eardrums (the apertures through which the sound propagates), determine the fundamental frequency of resonance according to the following equation:

$$f_o = \frac{c}{2\pi}\sqrt{\frac{A}{LV}} \tag{2.3}$$

where c is the speed of sound,where A is the mouth area, L is the length of the neck (of a bottled shape resonator) and V is the volume of the cavity. In the case of the cicada, which has two tympana and therefore two sonic apertures or holes, the effective area, A is the combined area of

the two tympana. The "neck" of the abdominal resonator is not so clearly defined as in a bottled shape resonator; its effective length therefore, is calculated according to

$$L = \frac{16r}{3\pi}, \tag{2.4}$$

where r is the radius of one tympana [42].

The increase in the amplitude of vibration at the fundamental frequency—and the rate at which such a change occurs—is determined by the quality factor or Q of the resonator (previously discussed in Section 2.2). Though the dimensionless quantity was given in (2.1), the Q of a Helmholtz resonator is also given by

$$Q = 2\pi\sqrt{\frac{L^3V}{A^3}}, \tag{2.5}$$

where A is the combined area of the two holes and L is defined by (2.4) [40]. Note that Q is made large when the area, A, is small or the volume, V, is large. When applied to a cicada-like system, equation (2.5) ignores any mechanical damping within the tymbals or abdominal walls [82].

Though the abdominal air sac is generally tuned to the same fundamental frequency as the tymbal, adjustments can be made by extending the abdomen to adjust its volume and thus its resonant frequency. The tuning of the tymbal and the abdomen is a very important factor in the nature of the resulting sound. If the two resonators are well tuned, the abdomen provides an amplification system for the mechanical vibrations of the tymbal. In addition, the abdomen will smooth the more impulsive waveform of the tymbal's pulse train and asymmetries that will likely occur between the independently driven left and right tymbal. The effect of an acoustic resonator with different Q values is shown in the model output as seen in Figure 2.10.

2.5 The Tympanum

The extremely loud sound, for which the cicada is often renowned, is also aided by the inertial element of the Helmholtz resonator—the large eardrums or *tympana* which provide a sonic aperture through which sound propagates. The tympana, forming part of the auditory system as well as the system for sound production, vibrate significantly during vocalization and, in fact, the loudest sound is produced in their vicinity [39] making them, not the tymbals, the main sound radiators.

Because the tympana are much larger than the tymbals (see Figure 2.2) they improve the acoustic impedance matching[1] between the insect and the surrounding air. This allows the sound to radiate more freely, enabling the cicada to project an extremely loud song [42].

As seen by (2.5) and (2.3), the tympana also plays an important role in determining both the Q and the fundamental frequency of the abdominal air sac by changing the effective surface area, A.

[1] For good impedance matching, the linear dimensions of the sound-radiating structure should vary in direct proportion to the to the sound wavelength.

In addition to changing the volume of the abdomen, the cicada can open and close the opercula (see Figure 2.2) to change the area of the sonic aperture and thus the frequency and Q of the resonator.

2.6 The Cicada Model

The cicada model consists of buckling ribs, a vibrating tymbal and a Helmholtz resonator for the abdominal air sac. Once the frequency of the tymbal and the abdomen is determined using (2.2) and (2.3), they can be simulated using a simple two-pole two-zero resonant digital filter. In the case of the abdomen, where more tuning is required, a higher order filter is used. This can be implemented by placing several such filters in series.

2.6.1 Bandpass/Resonant Filters

The transfer function of a bandpass/resonant filter, where the poles are chosen to be a complex conjugate pair [74], is given by

$$\begin{aligned} H(z) &= G\frac{(1-\sqrt{R}z^{-1})(1+\sqrt{R}z^{-1})}{(1-Re^{j\theta_c}z^{-1})(1-Re^{-j\theta_c}z^{-1})}, \\ &= G\frac{1-Rz^{-2}}{1-2R\cos\theta_c z^{-1}+R^2z^{-2}}. \end{aligned}$$

Because both the numerator and denominator are quadratic polynomials in z, this filter is often called a *bi-quadratic* filter or simply a *biquad* [76]. The corresponding difference equation is given by

$$y(n) = Gx(n) - Rx(n-2) + 2R\cos\theta_c y(n-1) - R^2 y(n-2). \tag{2.6}$$

The radius of each pole, R, is set according to the desired bandwidth or Q of the resonator using the following approximate relation [74]

$$R = e^{\pi BT} \tag{2.7}$$

where B is the bandwidth at -3dB in Hz (its relationship to Q is given in 2.1) and T is the sampling period. The angle θ_c determines the fundamental frequency according to

$$\theta_c = 2\pi f_c T. \tag{2.8}$$

The two main parameters for these filters, the fundamental frequency f_c and the Q (or alternatively, the bandwidth), determine the characteristic of the resonator and can be changed according to other model parameters in real-time, making this an efficient and ideal implementation for performance situations. If a resonator requires more than one mode (as is the case for the tymbal discussed below), multiple *bi-quad* filters can be placed in parallel (one for each mode) with the overall output being the sum of the filter outputs.

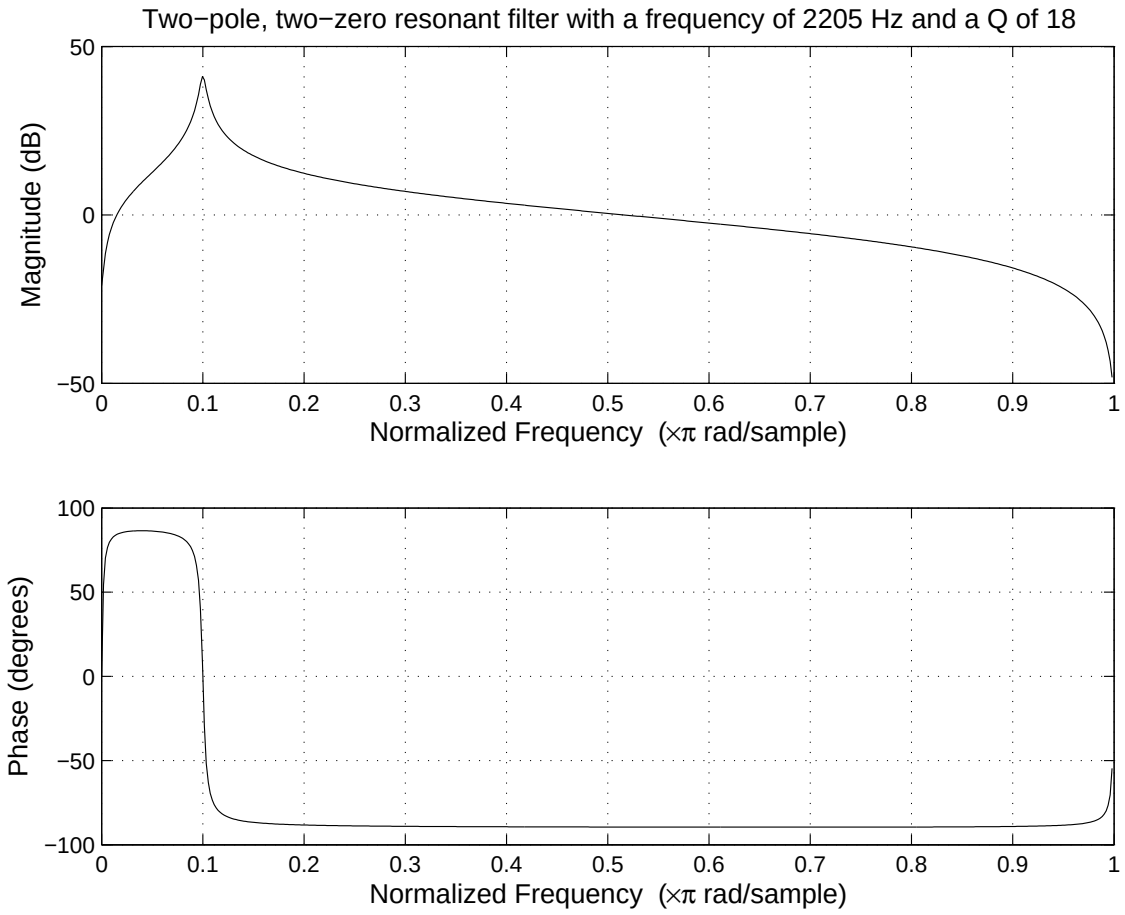


Figure 2.8: Magnitude and phase of the two-pole, two-zero biquadratic filter. The resonant frequency is set to 2205 Hz (with a sampling rate of 44100) and the Q to 18.

2.6.2 The Tymbal Model

There are two prominent modes in the measured spectrum of the tymbal with the fundamental being near 4 kHz. Given the small size of the tymbals, and therefore the relatively small number of modes, it is unnecessary to implement the motion of the vibrating tymbal using finite difference methods or a more complex algorithm such as the two-dimensional waveguide mesh (a common method for modeling membranes, plates or shells where vibration occurs in two-dimensions). Rather, the tymbal is more simply modeled using two biquadratic filters in parallel, one for each of the two modes in the spectrum. The output of the tymbal model is the sum of the output of each of the two biquads.

The fundamental frequency of the resonator, θ_c is calculated according to (2.2). Since the mass is determined by Table 2.1, the frequency of the model is controlled primarily by changing the stiffness parameter.

Like any resonator, the oscillating output of the tymbal is contingent on a source of energy, that is, it needs to be driven by some kind of excitation mechanism. Therefore, in order for this resonator to produce sound there must be some kind of input, whether it be the output of another resonator or an initial impulse.

In the initial model, a variable width hamming impulse is sent through the two biquadratic filters modeling the tymbal, each time a rib buckles. This input excitation is later improved (as discussed in Chapter 3) by a mechanical input device, the *Tymbalimba*, which measures the energy generated by a buckling spring-loaded aluminum rib and provides a more accurate sequence of excitation impulses. However, the hamming impulse is still quite satisfactory as an input to the tymbal model and with the width and height of the impulse as parameters, the corresponding produced sound can change considerably.

For each *IN cycle* (referring to the sequence of buckling ribs produced by one muscle contraction), the changing mass and frequency from Table 2.1 is considered, and the filter coefficients are updated accordingly. Because of valuable data provided by Bennet-Clark [81], it was possible to determine the elapsed time between each point of buckling and the relative amounts of energy released by each rib in the buckling sequence (see Figure 2.6). In order to produce the buckle sequence, impulses are generated according to this data while ensuring proper phase matching between each buckling point. An example output waveform of the tymbal resonator can be seen at the top of Figure 2.10.

The *OUT cycle* refers the point of muscle relaxation, when the ribs spring back to their convex shape. This event occurs much more quickly and only the mass of the plate is used in calculating the frequency at this point.

Quantity	*Ciclochila australasiae*	*Macrotristria angularis*	Unit
*Body length	45.3	44.5	mm
*Body width	17.6	16.8	mm
*Tympanal width	8.37	7.0	mm
*Area of one tympanum	43.3	28.2	mm^2
*Tympanal thickness	1-2	—	μm
Volume of abdominal air sac	1.93	1.67	cm^3
Hole radius from area	3.70	3.0	mm
Effective neck length	6.3	5.1	mm
Abdomen resonant frequency	4.56	4.41	kHz
*Song carrier frequency	4.3	4.0	kHz

Table 2.2: Example parameter values for males in two cicada species [39]. Data marked with * is taken from [60].

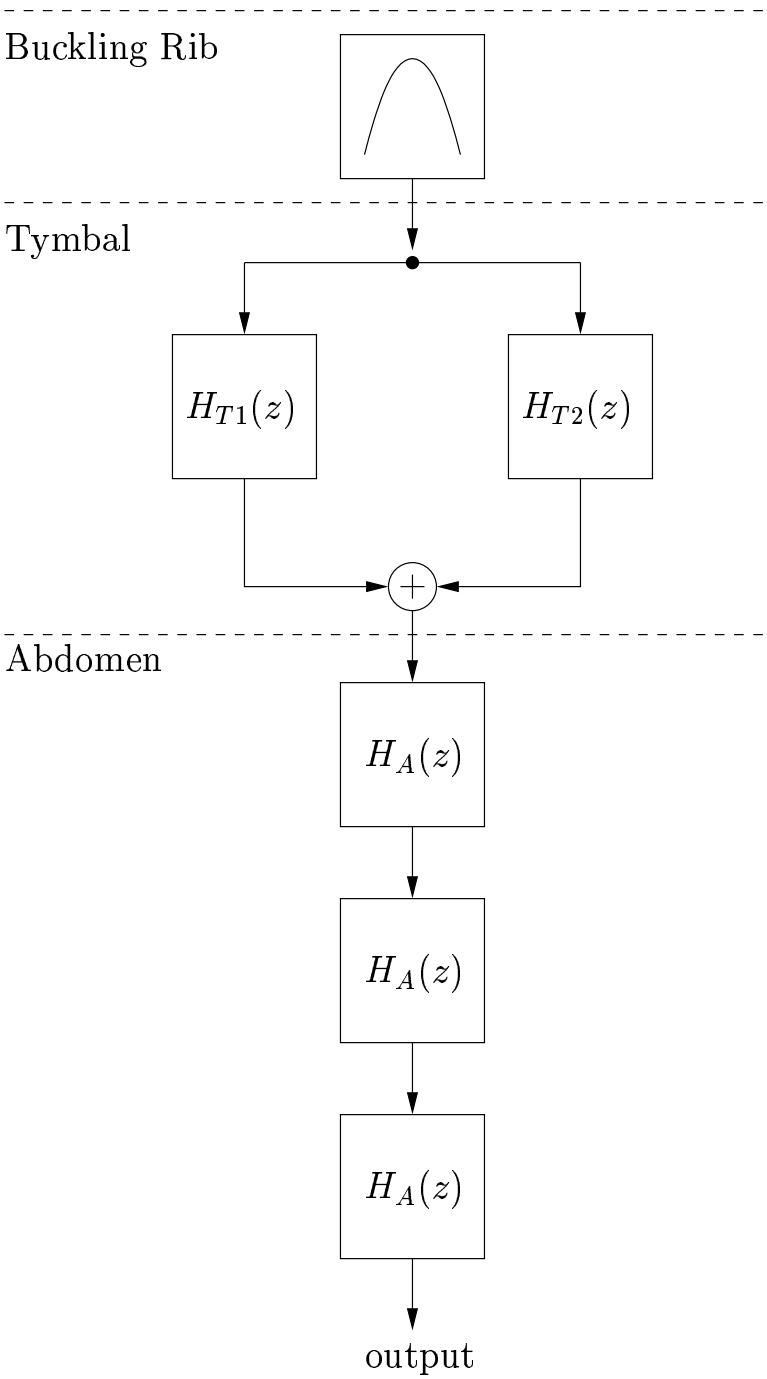


Figure 2.9: This signal flow diagram of cicada model shows arrangement of resonant filters. An impulse is input into the system each time a rib buckles. Two resonant filters $H_{T1}(z)$ and $H_{T2}(z)$, one for each mode, are placed in parallel to model the tymbal while several resonant filters $H_A(z)$ are placed in series to achieve finer tuning in the abdomen.

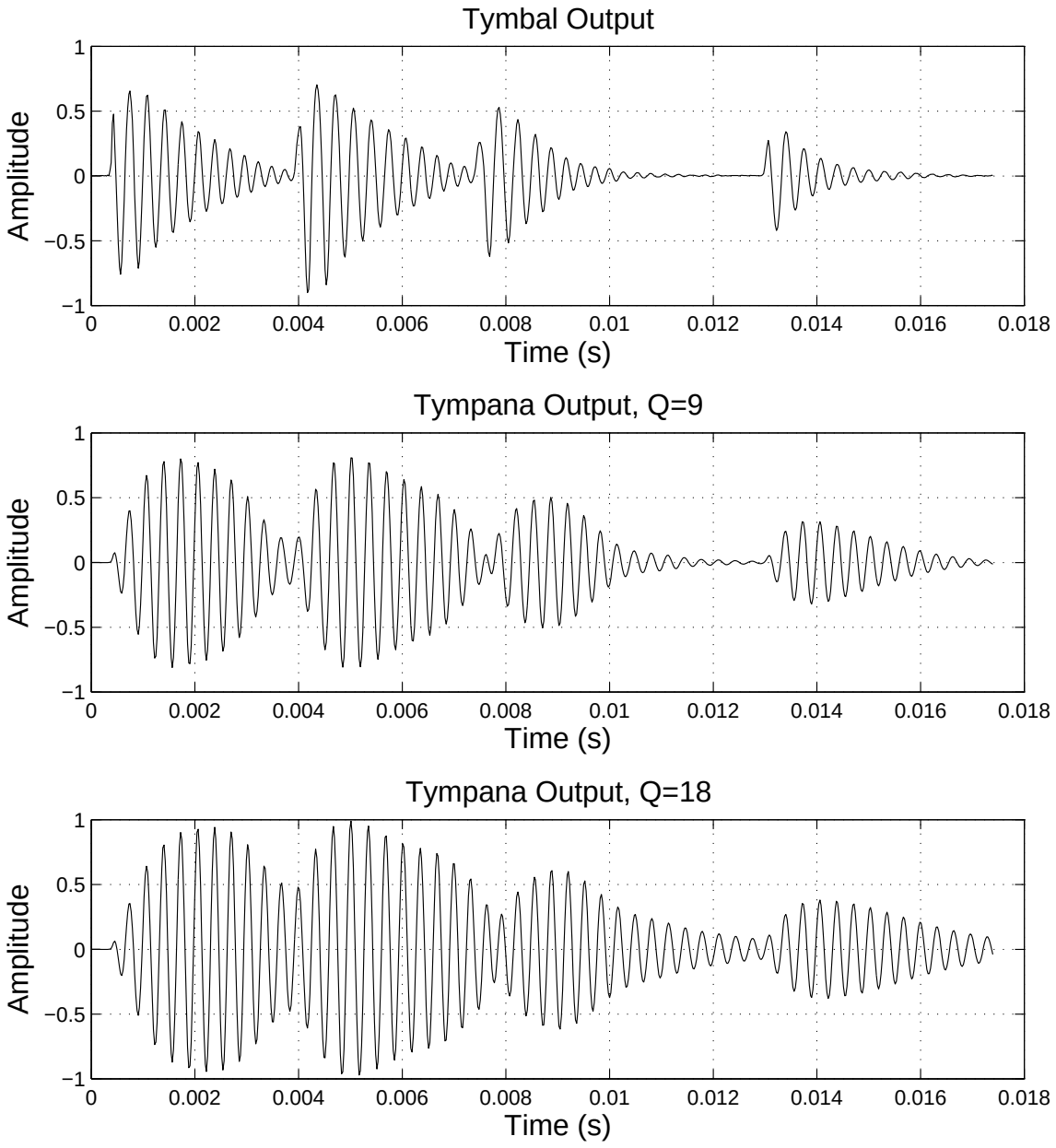


Figure 2.10: The waveform produced by the tymbal model (top) and after being passed through a Helmholtz resonator with a Q of 9 (middle) and a Q of 18 (bottom) representing the cicada's abdominal air sac.

2.6.3 The Abdomen Model

The output of the tymbal is put through a similar resonant filter (or *biquad*) modeling the abdominal air sac. Since finer tuning is required, a higher order filter is used.

Using the body dimensions of two different species of cicada as seen in Table 2.2, a range of possible Q values can be calculated according to (2.5). The Q and the fundamental frequency of the resonator can both be altered by changing the volume of the abdomen and by adjusting the opening of the opercula (effectively increasing the area of the tympana). Both actions will require recomputing the filter coefficients.

When the output of the tymbal is filtered by the abdomen tuned to the same fundamental frequency, the resulting waveform is greatly amplified as well as being significantly smoother and more coherent (with fewer discontinuities). Notice in Figure 2.10 that the moments at which each individual rib buckles is less discernible with an increase in the Q of the abdominal resonator.

2.7 The Cicada Instrument

The cicada model is implemented as an an external object (in the C programming language) within an interactive software environment called Pd (pure data) [88]. A screen shot of the cicada object in Pd is shown in Figure 2.11. This environment was selected for its real-time interactivity, its inlets and outlets allowing for compatibility with other objects and the overall ease with which it can be used by musicians in live musical performances.

2.7.1 The Controller

The original controller for the cicada was a pair of electronic gloves which allows for simple communication, via MIDI, with the cicada object in the Pd environment (a screen shot of the cicada object in this environment is shown in Figure 2.11). The tip of each finger in the glove, as well as the heel on each hand, is equipped with a force sensing resistor or FSR. There is a variable voltage (between 0-5V) output from each sensor which is proportional to the amount of pressure applied to a finger or heel. This analog signal is sampled using an analog to digital converter (MAX1270) before being sent on its own channel to the Basic Stamp (a programmable micro-controller). The Basic Stamp is programmed to receive and process any data collected on the chip's channels. The final output of the stamp is a continuous controller MIDI message with a MIDI channel for each FSR and a value between 0 and 128 indicating the amount of pressure applied to that sensor. This is then read by Pd's midi object and the data is mapped to the various parameters of the model.

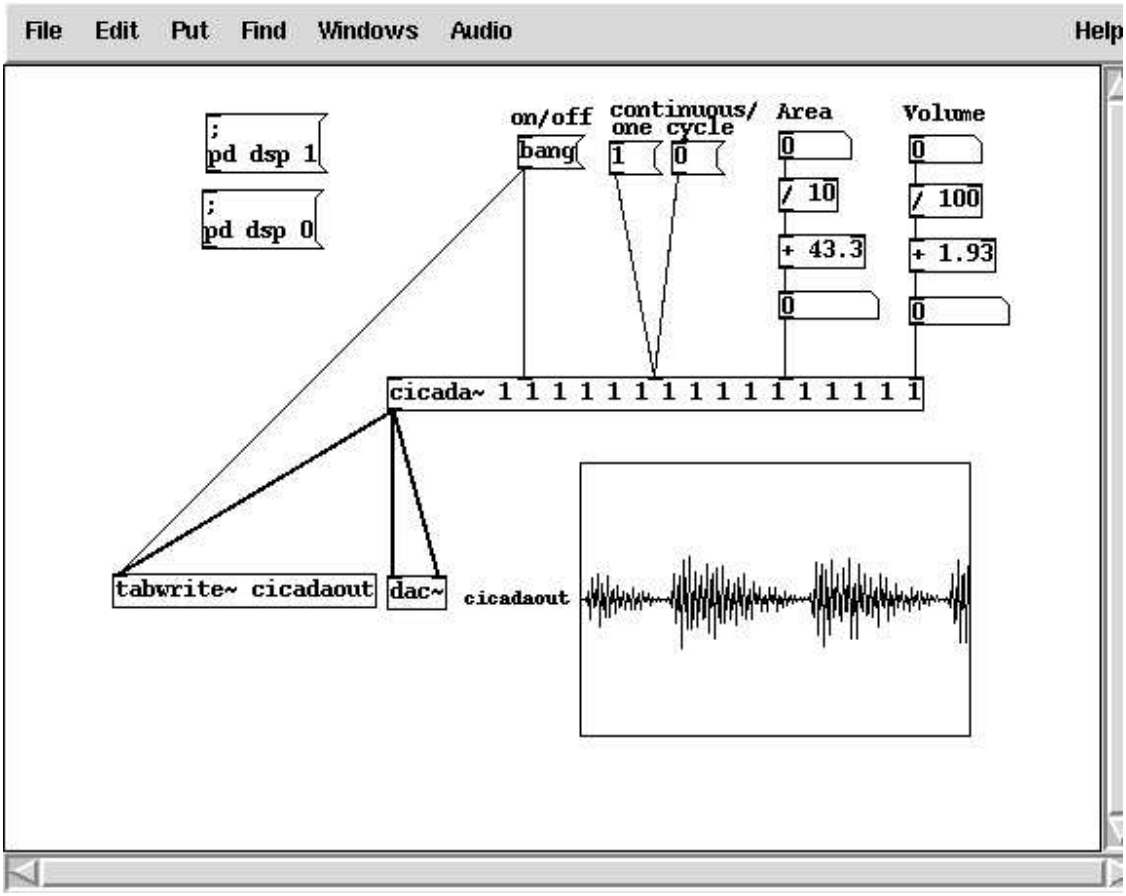


Figure 2.11: A screen shot of the cicada model in the Pd environment.

2.7.2 The Control Parameters

The three most important parameters which contribute to the nature of the cicada's sound are the rate of contraction of the tymbal muscles, the structure of the tymbals which includes the number of ribs on the tymbal and the stiffness of the dorsal plate, and the tuning between the tymbal and the abdominal air sac [44].

Rate of Muscle Contractions

Depending on the species of Cicada, the rate of muscle contractions can vary a great deal. However, in one of the species on which this model is based, the neurogenically excited muscle contractions can occur as frequently as 117 times per second [41]. This rate is, of course, extremely fast and beyond the locomotion ability of any human.

The player specifies the rate of contractions therefore, and the computer handles the repetition

Figure 2.12: A Max (MSP) patch is used to demonstrate the sensitivity and resolution of the sensors. Each finger and heel of the glove is mapped to a level meter, and gives visual feedback to the amount of pressure being applied.

of the buckling ribs with some random adjustments of the timing. This leaves the player free to control other input parameters. When playing the instrument therefore, the player is not expected to produce four impulses per muscle contraction and 117 muscle contractions per second. One of the many conveniences of computer music instruments is that humanly impossible tasks can be left to the computer.

Structure of the Tymbal

The structure of the tymbal refers primarily to the curvature of the ribs (which would result in varying excitation signals when each rib buckles) and the number of ribs on the tymbal (which varies between different species). Also important is the stiffness of the dorsal plate.

The curvature in the ribs is accounted for in the model by allowing the user to change the shape, including height and width, of the hamming impulse that is sent to the tymbal model each time a rib buckles. Before the *Tymbalimba* controller—which will be described in Chapter 3—this was the primary way in which the player controlled the variation in the excitation mechanism.

The number of ribs on the tymbal can vary among species—and even within one species there is no certainty that *all* ribs will buckle during one cycle (as this depends on the energy generated by the buckling of the first rib). The number of buckling ribs is also a changeable parameter, though experiments were done to randomize this number by the computer to create a more interesting output.

Tuning

The fundamental frequency of the tymbal resonator is determined by (2.2). Recall that the tymbal behaves like a mass and spring system, with the reactive elements being the mass of the tymbal plate, and the stiffness of the tymbal's dorsal support. Since the buckling of each individual rib contributes to an increase in the overall mass of the tymbal plate according to Table 2.1, the fundamental frequency is determined primarily by the stiffness (which is controlled by the player). Besides simply changing the pitch, this will result in more or less consonant tuning with the secondary abdominal resonator, creating a variable complexity in the tone.

The fundamental frequency of the abdomen model can be controlled by changing either the area of the tympana (or equivalently, the opening of the opercula) or the volume of the abdominal air sac according to (2.3). Therefore, just as the cicada has the ability to fine tune the frequency of its song by making slight adjustments to the volume of the abdominal air sac, the instrument permits the same control. As (2.5) implies, changing the abdomen frequency by changing the its volume instead of the tympana area will affect the Q of the resonator differently. This adds another interesting control over the sound output.

The coupling of the tymbal to the abdominal air sac, tuned to the same frequency, allows the signal to project very strongly. In addition to increasing the volume, the secondary resonator acts as a filter to smooth the discrete pulses from the output of the primary resonator (see Figure 2.10. The nature of the sound can be changed significantly by altering the tuning of the two resonators and, depending on their relative frequency, the secondary abdominal resonator will have a varying effect on the resulting sound.

2.8 Conclusion

Though the sound output of the cicada model is very interesting and certainly has musical potential—particularly when varying the tuning between the tymbal and the abdomen—the most striking aspect of this sound production mechanism is the efficiency with which the cicada is able to use a rapid sequence of buckling ribs to create perceptually sustained tones.

Though the glove controller is very useful for exploring the parameter space and the tunings, it is not sufficient for exploring this particularly unique mechanism. Mapping each of the parameters to the heel of the hand, or to a finger, is a rather cumbersome mapping strategy and does not provide

the user with an intuitive sense of what s/he is controlling.

The *Tymbalimba*, the subject of Chapter 3, addresses this issue by constructing a controller which provides the computer synthesis model with an input signal that corresponds more closely with the user's input and musical intentions.

Chapter 3

The *Tymbalimba*

3.1 The Need for a Haptic Controller

There is a strong need in the field of computer and electronic music to create usable instruments with which musicians and artists can interact with the same depth and richness as is currently being done with acoustic instruments. Though digital technology can be very empowering by allowing the musician's creativity to expand beyond the constraints imposed by physical reality, some aspects of physical reality are still necessary if the human player be able to effectively interact with the virtual instrument.

One of the most widely used electronic musical instruments, the Theremin [84], has a simplicity in design, a limited control parameter space—pitch controlled with one hand and volume with the other as seen in Figure 3.1—and a responsiveness to motion that allows the player to use intuitive gestures to control the produced sound that is potentially very musical. In spite of this, the Theremin is often criticized for its free space control lacking haptic feedback which makes it very difficult to play notes with precision.

When this control is mastered however, the sounds produced by the Theremin, as demonstrated by virtuosic players such as Clara Rockmore and Lydia Kavina, come very close to the quality of traditional acoustic instruments. This instrument is rarely mastered however, and as a result, the music produced by the Theremin is often limited to eerie ghost-sounding glissandi effects for B movie soundtracks. The Theremin is exemplary of the importance of a good input device—even in the presence of limited control parameters and an intuitive mapping—which will allow the player to use natural and unstrained movements that intuitively correspond to his/her musical intentions.

The synthesis model of the cicada described in Chapter 2 requires such an effective musical controller. Though the sound produced by the model is entirely satisfactory, the controller mentioned in Chapter 2 does not make use of the cicada's unique sound mechanism—a sequence of efficiently buckling ribs that enable sustained vibration of the tymbal, each rib having a rebound or *upbuckling*

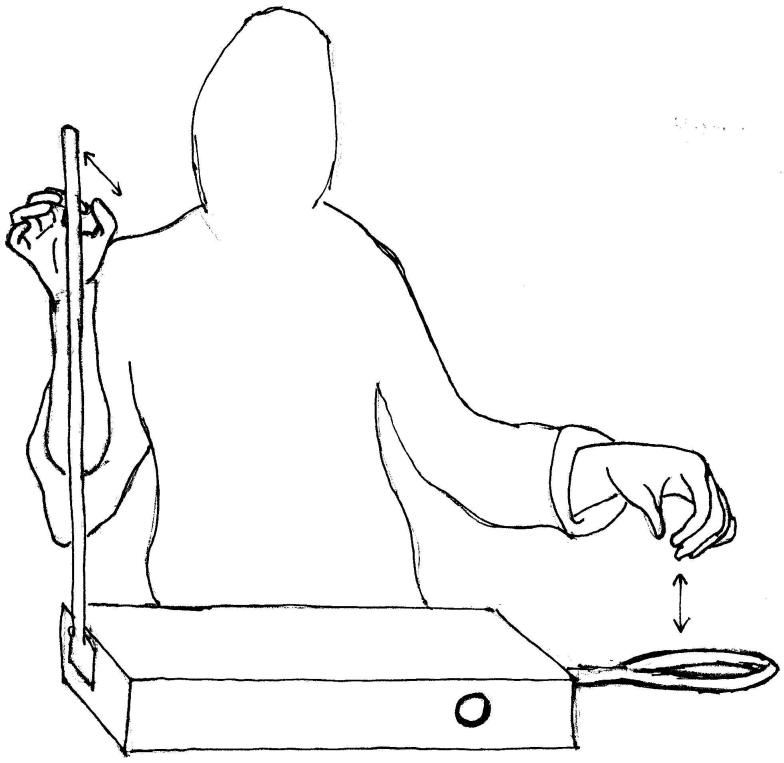


Figure 3.1: Two-handed free space interaction with the Theremin.

mechanism making the repetition of the buckle cycle more efficient. In this chapter, a scale model of the ribs is created and its use as a musical controller, interfaced to the synthesis model, is explored.

3.2 A Controller Inspired by the Cicada's Rapid Buckling Mechanism

Musical instruments generally require that the player supply continuous energy to sustain a tone, as in the continuous motion of the arm while bowing a string instrument or the continuous blowing required to sustain a note on a wind instrument. Discrete excitations mechanisms such as plucking or striking produce tones that decay after a period of time. And though it may be common to play a rapid tremolo to simulate a sustained tone such as in the rapid strumming of a mandolin, this still involves discrete excitation with each attack requiring some action on the player's part, and overall, demanding significant energy and virtuosity.

Buckling mechanisms that initiate and sustain mechanical vibrations have been used in noise makers and other sound making devices yet do not seem to be employed by traditional musical instruments. The uniqueness of this mechanism in the cicada lies both in the way the mechanism functions sequentially, where one buckling rib caused by one muscle contraction can initiate a sequence of excitations, and also in the rebound or *upbuckling* which allows the cicada to repeat the action extremely rapidly. Both these mechanism contribute to the efficiency with which the cicada initiates and sustains tones during vocalization.

This chapter describes a music controller inspired by the cicada's buckling mechanism. The *Tymbalimba* was developed to determine whether the buckling mechanism could be used as efficiently by a human, and whether it could be adapted to function effectively as a musical instrument. The device provides a mechanical user interface to the existing synthesis model of the cicada's sound production mechanism described in Chapter 2, and allows the user to manipulate the computer model in a meaningful way using intuitive gestures.

3.3 The Cicada's Sequential Buckling Mechanism

As described in Chapter 2, the cicada's primary sound production organ, the tymbal, is equipped with a plate and a series of connected convex ribs that collapse or *buckle* under the force of a contracting muscle [47].

When the tymbal muscle contracts, the tymbal plate is pulled inward and a central downward force is placed on the adjacent rib to which it is connected. This force is sufficient to cause the first rib to buckle inward and, since all the ribs are joined by an elastic resilin, a similar force is exerted on the next rib in the sequence, causing it to buckle. Depending on how much force is generated by the first buckling rib, a number of subsequent ribs will buckle in a domino effect, where each rib buckles under the force of its neighbouring rib. During the very short time intervals between muscle contractions when the tymbal muscle is at rest, each rib that is in an inward-buckled position springs back to its original shape, readying the ribs so the sequence can be repeated.

After the first rib in the sequence buckles, the tymbal plate is immediately set into vibration. The buckling of succeeding ribs supplies the system with bursts of energy allowing the vibration of the tymbal plate to be sustained [42].

3.4 The Mechanical Model

The Tymbalimba, shown in Figure 3.2, is similar to the cicada's tymbal in that it consists of a sequence of four convex ribs capable of buckling into a concave shape. The arrangement of the aluminum ribs in a sequence is reminiscent of the keys on a kalimba with each rib fitting comfortably under the hand's fingers, the spacing between the ribs being adjustable to suit the user.

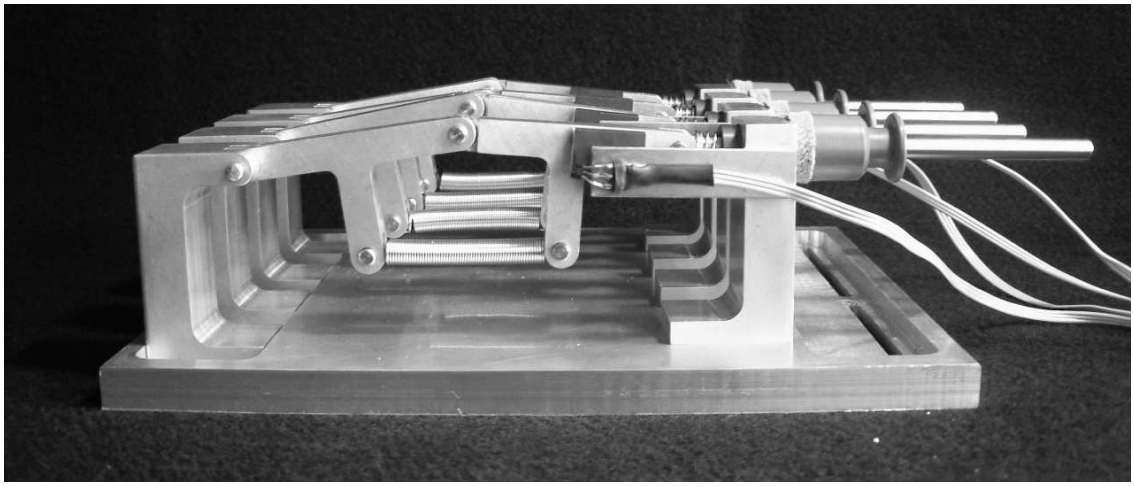


Figure 3.2: The *Tymbalimba*.

When the user applies a sufficient downward force to a rib, it buckles as shown in Figure 3.3. The energy signal generated by this motion will depend on the applied force and the velocity of the user's finger. As with the tymbal ribs, when the user removes his/her finger, thus effectively removing the applied downward force, the model's ribs will spring back to their original shape.

Each rib is assembled from machined aluminum pieces (see Appendix A for a description of the machining methods) and springs empirically selected to achieve the desired haptic response from the device. As illustrated in Figure 3.4, the primary spring controlling tension in the ribs is spring A (though there is also some contribution from spring B), a compression spring which cannot be too stiff or too much effort would be required to buckle the rib. Spring B in Figure 3.4 is an expansion spring responsible for the rebound or up buckling of the rib. Without spring B the rib would remain in its inward-buckled position and would not be restored to its original shape once the player's finger is released and the downward pressure removed. In order for the mechanism to operate correctly, a change in spring A (primarily effecting stiffness) will likely require a change in spring B (to achieve the rib rebound). Much experimentation involving several different users was required before settling on the combination of springs that provided the right feel and responsiveness to the device.

Though the springs are pre-selected, there is a mechanism in place which allows the user to adjust the stiffness and rebound (or bounce) of the ribs. The *stops*, shown in Figure 3.5, allow the mechanical ribs to be pulled out or pushed in, effectively changing the angle of their curvature. Just as a cicada may change the amount of energy released by a buckling rib by changing the curvature in the ribs [44], so can the user change the stiffness of the mechanical rib by adjusting the stops at each rib end. Letting the stops out will flatten the rib, making it less stiff and easier to buckle by making it more "bouncy". Though this will produce a lower amplitude impulse, it has the benefit of allowing the user to play notes in more rapid succession. Likewise, pushing a stop inward increases

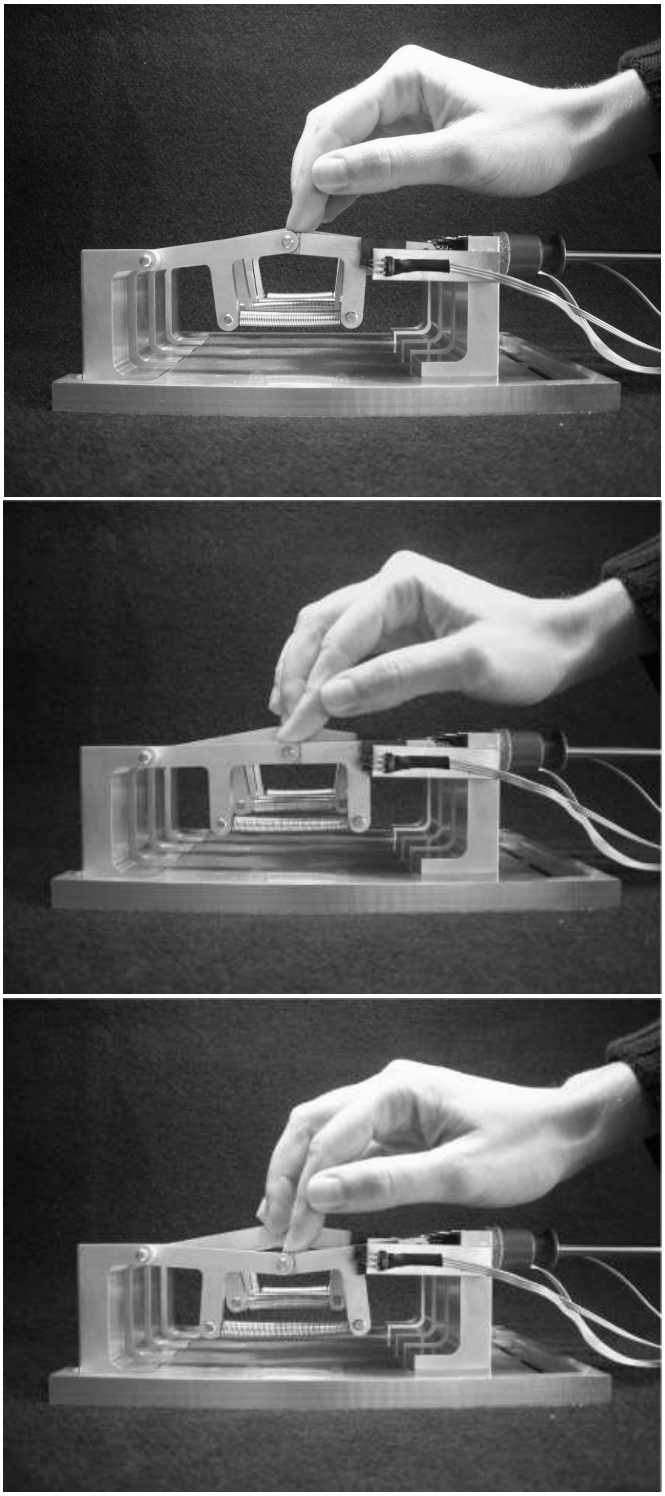


Figure 3.3: The controller's buckling motion.

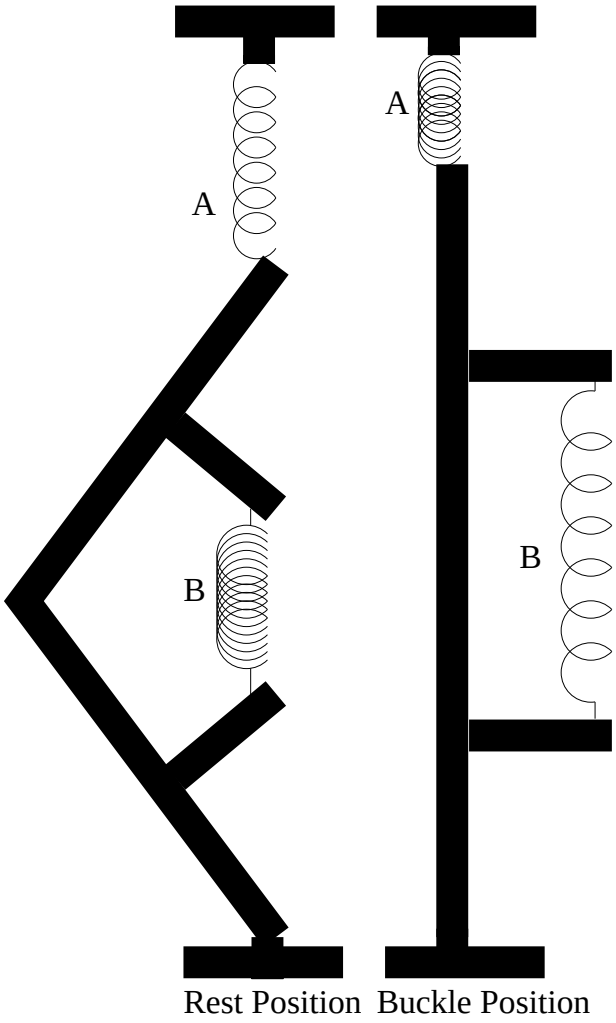


Figure 3.4: The expansion and compression of the springs during buckling.

the slope in the curvature of the rib, making it slightly more difficult to buckle yet generating a higher amplitude signal at the point of buckling. This adjustment then will have an effect on the feel of the instrument and, because of the way it varies the energy released by a buckling rib and thus the number of ribs that buckle in the sequence, it has an effect on the overall produced sound.

Depending on the species of cicada, muscles may contract as frequently as 200 times per second. Of course, the neurogenically initiated muscular contractions of the cicada give it a significant advantage over less agile human fingers. It would be absurd to expect a human to buckle the mechanical ribs with a repetition rate comparable to that of an insect, as it certainly wouldn't be long before the user suffered from repetitive strain injuries. The mechanical controller therefore captures the energy generated by the buckling of the first rib, and based on this signal, the computer

Figure 3.5: Adjusting the *Tymbalimba*'s stops.

model accurately produces the sequence of impulses generated by additional buckling ribs.

The additional advantage to this design is that all four ribs are free to be controlled by the user—effectively providing the user with multiple tymbals that can be tuned differently and played independently. Just as the cicada has two tymbals—one on either side of its abdominal air sac—that can be controlled with separate muscles and tuned to different frequency, it was desirable that this feature also be made available to the user of the Tymbalimba. The independence of the ribs—made possible by having the buckling sequence generated by the computer model—provides the user with four independent tymbals, all individually controllable with the motion of a finger (this number could, of course, be expanded in a future version of the controller). Depending on how the ribs are mapped in the computer model, gestural motions such as drumming the fingers over all four ribs can create a wide range of different sounds.

A public demonstration of the controller at an early stage in its development showed that users were captivated by the responsiveness of the mechanism itself—even though at this stage they were not actually controlling any sound output. The way the spring-loaded mechanism caused the ribs to bounce back up at the user every time they were buckled created an animated excitement in some users and clearly showed the importance of physical stimuli in haptic user interfaces—something which is far too often overlooked in the design of musical controllers. Since this demonstration and the addition of sound controlling capabilities, experiments have shown that the rebound buckling, or outbuckling, is also very useful for rapidly repeating notes. This is suggested by the blurred motion of a player's hand in Figure 3.6, captured from a video.

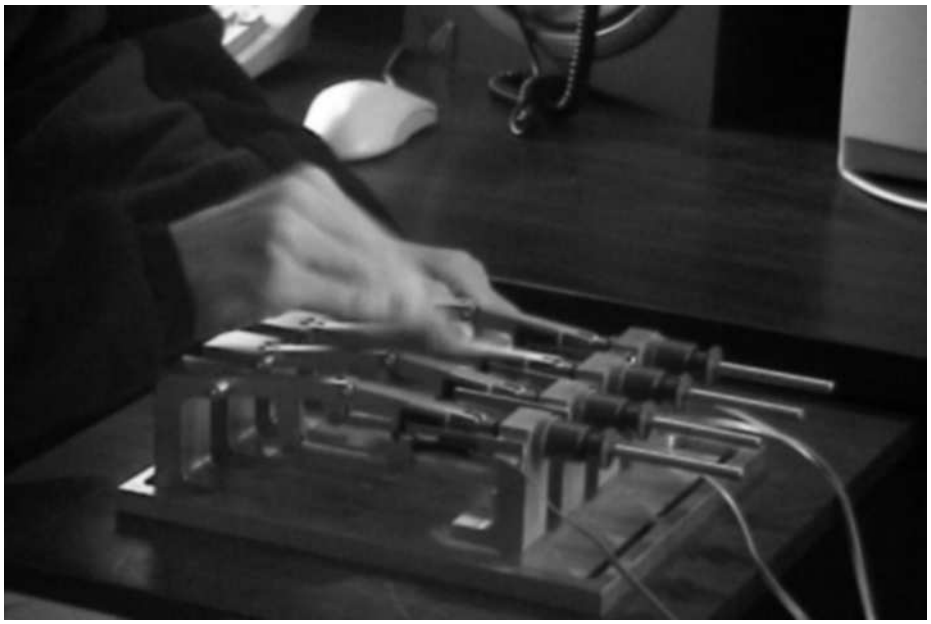


Figure 3.6: The blur from the finger's motion suggest the agility with which the user is able to buckle the ribs.

The player can change the signal generated by the buckling of a rib by varying the gesture of the attack, that is, by changing the velocity and/or force of the finger motion. Many of the parameters that significantly change the produced sound of the computer model are determined by the nature of this one attack. For example, the number of ribs that buckle in the sequence and the width and height of the generated impulses is dependent on the amplitude and shape of this initial input. Though this interdependence of parameters may seem like a reduction in control available to the user, it is preferable to an arrangement whereby the user must be aware of and control an overwhelming myriad of individual parameters by assigning each to one finger (as was done in the original glove controller described in Chapter 2). As with other successful musical instruments, the control parameters are locked in a certain relationship determined either by the physics of the system or by some preprogrammed mapping. The result is that the user must work on creating different gestures and learning the corresponding produced sound to play effectively. Ideally, with this paradigm, the player need not even be aware of the existence of other parameters during performance in order to create a wide variety of possible sounds.

3.5 Interfacing to the Cicada Synthesis Model

Rather than measure the force or energy input of the user (as many controllers do), the Tymbalimba measures the energy generated by the ribs that are buckled by the user. This is a benefit in having

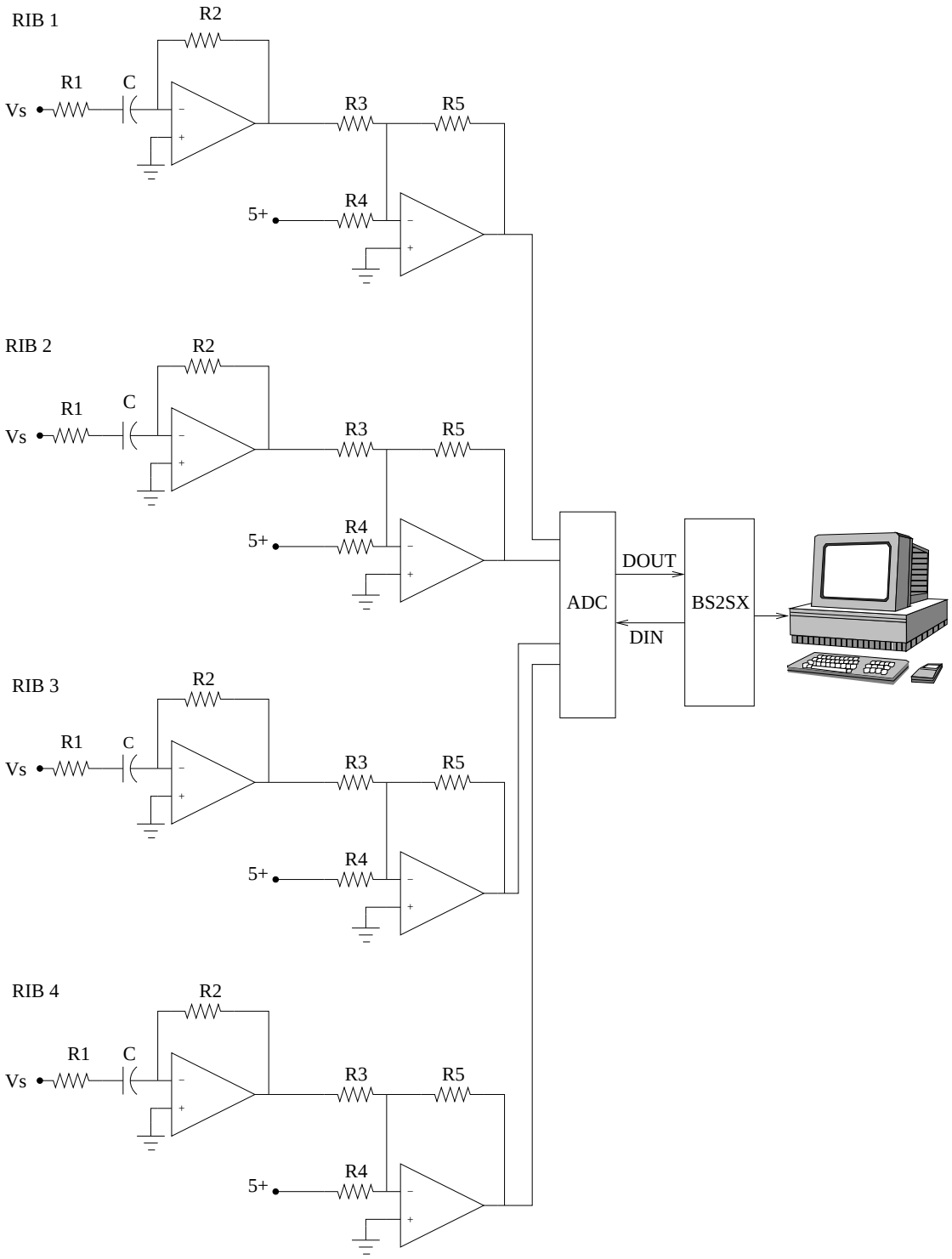


Figure 3.7: The signal conditioning circuit for the Hall Effect sensors. The components surrounding and including the first op-amp make up a differentiating circuit to obtain a signal representing velocity.

a mechanical controller: energy in the system can be both fed back to the user to provide valuable haptic information and also measured to obtain an accurate input signal to another system, in this case a physics based computer model.

The input signal is obtained by placing Hall Effect sensors to the side but slightly away from each rib and a magnet directly on the rib itself. During buckling, the magnet passes over the sensor causing it to produce a voltage representing its displacement from the magnet. The displacement is passed through the analog differentiating circuit illustrated in Figure 3.7 to obtain velocity which, since proportional to the square root of energy, can be used to obtain the energy input signal to the physical model.

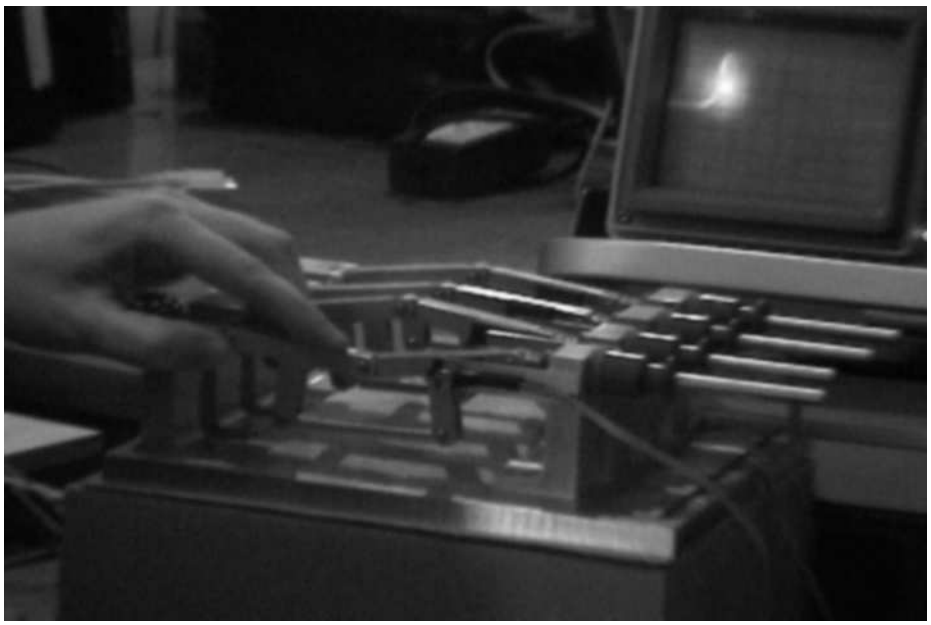


Figure 3.8: The signal (as viewed on the background oscilloscope) after passing through the signal conditioning circuit in Figure 3.7 representing the energy generated by the *downbuckling* of a rib.

The voltage from the sensors (V_s on the far left of Figure 3.7) represents the displacement of the ribs during buckling. The electronic components surrounding and including the first op-amp in Figure 3.7 make up a differentiating circuit. The displacement voltage from the sensors passes through the differentiator to obtain a signal representing velocity. Once this signal passes through an analog to digial converter (ADC in Figure 3.7) it is sent to be processed by the Basic Stamp micro-controller (see Appendix A for Basic Stamp code used to acquire this signal). Since positive numbers are more easily handled from the Basic Stamp chip, additional circuitry (seen around the second op-amp in Figure 3.7) is used to add a DC component to the signal. An analog to digital converter reads the voltage values from four channels (one for each rib) and sends it to the Basic Stamp (BS2SX) microcontroller where it is redirected in digital form, via the serial port, to the

Figure 3.9: The signal (as viewed on the background oscilloscope) after passing through the signal conditioning circuit in Figure 3.7 representing the energy generated by the *upbuckling* of a rib.

computer.

Figures 3.8 and 3.9 illustrate the velocity voltage produced when a rib is buckled downward and upward as measured by an oscilloscope (seen in the background of both figures). During *downbuckling*, an upward impulse appears on the oscilloscope in Figure 3.8 and during *upbuckling*—occurring when the finger is raised from a rib—there is a very clear downward impulse on the oscilloscope in Figure 3.9.

Figure 3.10 shows a waveform representing a sequence of down and up pulses as obtained from within Pd. The serial data is interpolated and resampled to create an audio control signal which is used as an input signal to the synthesis model.

Different sounds will result from different gestures used to buckle the ribs. Different input gestures will result in different excitation signals and, like any resonator with a varying impulse excitation mechanism, the overall produced sound will vary accordingly. In the case of the cicada, varying the buckling gesture on the mechanical rib will produce varying sequence of impulses effecting the timbre of the sound. Figure 3.10 shows an example output of the model for the tymbal plate in response to one controller rib being buckled (only the downbuckle motion is shown) with enough force to set off four ribs in the sequence. The initiation and decay of the four impulses is very apparent in the resulting waveform. It is interesting to observe the decrease in amplitude by the time the final rib buckles—showing an apparent loss of energy in the system during one buckling cycle.

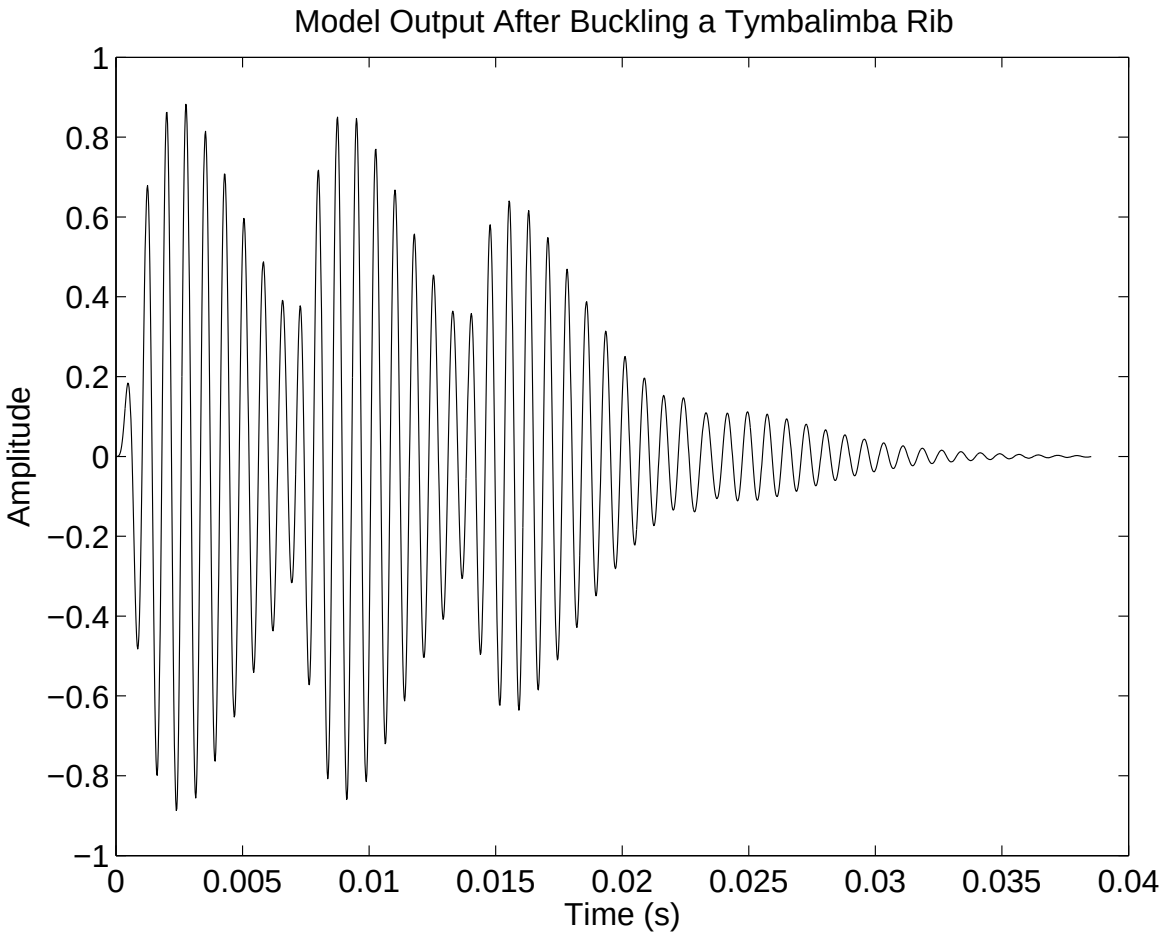


Figure 3.10: Output produced by the tymbal synthesis model after buckling a *Tymbalimba* rib.

3.6 Conclusion

The Tymbalimba provides a more accurate input signal to the physical model than the arbitrary impulse signals previously being used in Chapter 2. Also, because the user is playing a real mechanical device, s/he is not being denied important haptic information when playing a musical instrument [64]. Changing input impulse is all the user needs to do to get a responsive change in the produced sound and this can now be done using a single gesture with certain adjustments to the device to get the desired feel.

Physical modeling synthesis tends to be extremely parameter-rich, often requiring too much attention to individual variables to be suitable for real-time performance. The previously developed physical model of the cicada described in Chapter 2 suffered from this problem when determining the parameters of a rather complex system of buckling ribs. How many ribs should buckle? What

Figure 3.11: A sequence of up and down pulses can be seen on the oscilloscope (in the background) during normal playing.

is the rate of buckling? What is the rate of muscle contractions (the time before re-starting the sequence of buckling ribs)? What is the energy difference between the in and out buckling cycle or between the buckling of the first and last rib?

Just as the cicada varies each of these parameters during vocalization, so should the user be able to manipulate, in real-time, all the factors that determine the physical model's produced sound. Of course, it is impossible to be aware of so many individual parameters when playing any musical instrument (electronic or acoustic)—yet this does not mean it is impossible to control them. Since it is generally not desirable to remove functionality from a physical model by eliminating some of its parameters, the Tymbalimba obtains these parameters from functions representing one single action or gesture which the player may learn in various ways to produce different results.

The motion of buckling one rib produces a signal that is sent to the physical model. From this input signal, the physical model derives other parameter values without relying on additional user input. From one gesture, the model obtains information about the envelope of the impulse which will effect the timbre, volume and sustain of the sound. Since each rib is mapped to a pitch, frequency information can also be obtained. Since the parameters are determined by one motion they will change only as the motion changes. With time and practice, the user will eventually learn how varying the motion alters the produced sound and will be much more successful at playing the instrument.

The Tymbalimba goes beyond on/off triggering devices. Measuring the energy generated by

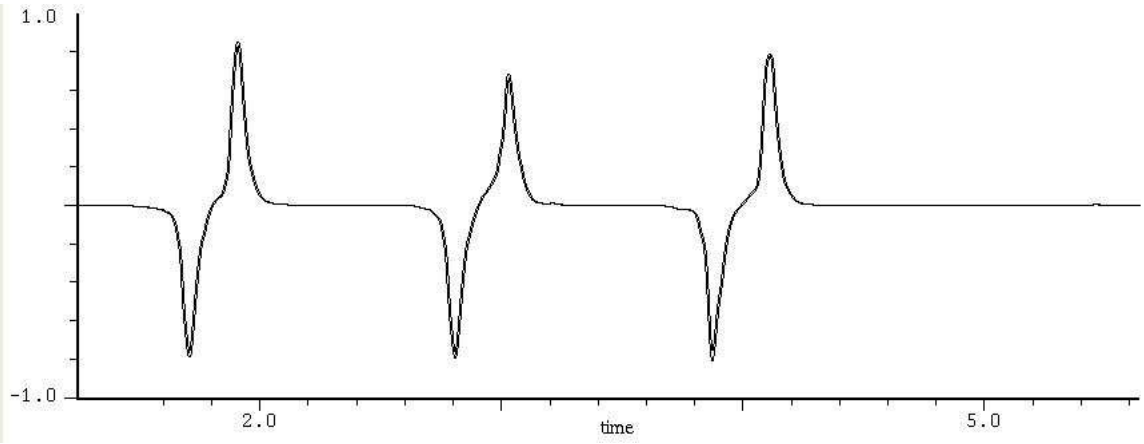


Figure 3.12: The signal generated by a series of up and down buckling, as measured from the computer.

the mechanical controller provides the computer model of the cicada with an input signal that accurately represents the user's gestures, as well as the buckling motion of the ribs. Not only does this eliminate the encumbrance of controlling the numerous individual parameters normally associated with physical models, it allows the user the satisfaction of playing a responsive haptic interface while hearing the results of the intriguing sustained tones of the cicada's vocalization.

3.7 Future Directions

An addition to the instrument, currently in progress, will allow the user to change the sounding pitch with the left hand, while controlling the buckling mechanism with the right. It is an extension to the base plate and will hold four linear force sensing resistors. Specially formed grooves in the plate will allow the user to easily locate the tempered placement of pitches on the FSR while still permitting the finger to slide between the notes.

Future work may also include creating different sizes of the controller. For example, it may be interesting to create a scaled down version that easily fits in one hand, and buckling occurs when one makes a fist. Another alternative may be to increase the number of ribs, and therefore the size of the device, so that a full keyboard range is supported.

A modified version of the Tymbalimba, with the left hand control for pitch is clearly reminiscent of a guitar model with the ribs taking the place of the strings. The buckling ribs have the effect of plucked strings with aftertouch, allowing them to be plucked or strummed very rapidly with the player's finger never really leaving the rib [92].

Other techniques may involve pressing on the ribs away from the buckling point to change the feel. It may also be possible to use the free hand to adjust tension in the ribs during performance.

Chapter 4

The Sound of the Avian Syrinx

Birdsong has always been very influential to musicians and any study on the applications of bioacoustics to music would be incomplete without including bird vocalization.

The sound of the songbird is often associated with the pure, high-pitched tone of the flute. This is exemplified by the titles many western composers have given their works that feature the flute such as Vivaldi's flute concerto *Il gardellino* (The Goldfinch), Debussy's *Syrinx* for flute solo, and Messiaen's work for flute and piano, *Merle noir* (Blackbird). Yet the theory that the songbird produces its sound using an aerodynamic whistle effect (much like in the flute) has been widely confuted [53, 37, 38]. Rather, it has been determined that the avian vocal tract uses a nonlinear vibrating membrane as its primary excitation mechanism. Much like the vocal folds in the human vocal tract, the syringeal membrane forms a vibrating valve, the output of which is filtered by the trachea which amplifies and attenuates certain modes.

Some of the waveguide synthesis methods described here have been used extensively for synthesizing wind instruments and, in fact, have also been used by Kahrs and Avanzini to model birdsong [8]. Their use in modeling the bird's vocal system provides a unique configuration of acoustic elements not found in traditional musical instruments, potentially providing a new computer musical instrument based on a sound mechanism that has always intrigued human listeners.

4.1 The Avian Vocal Tract

The bird's airway consists of a trachea which divides into the left and right bronchi at its base, and a membrane which forms a pressure-controlled valve near the top of each bronchus (see Figure 4.1). During voiced song, the membranes are set into motion by air flow, vibrating at a frequency determined partly by the mass and tension of the membrane and partly by the resonance of the air column to which it is connected [48]. The neural control of the muscles surrounding the syrinx, the pressure in the interclavicular air sac which encases the syrinx, and the bird's respiratory mechanics

all greatly contribute to how sound is modulated by the syrinx [38].

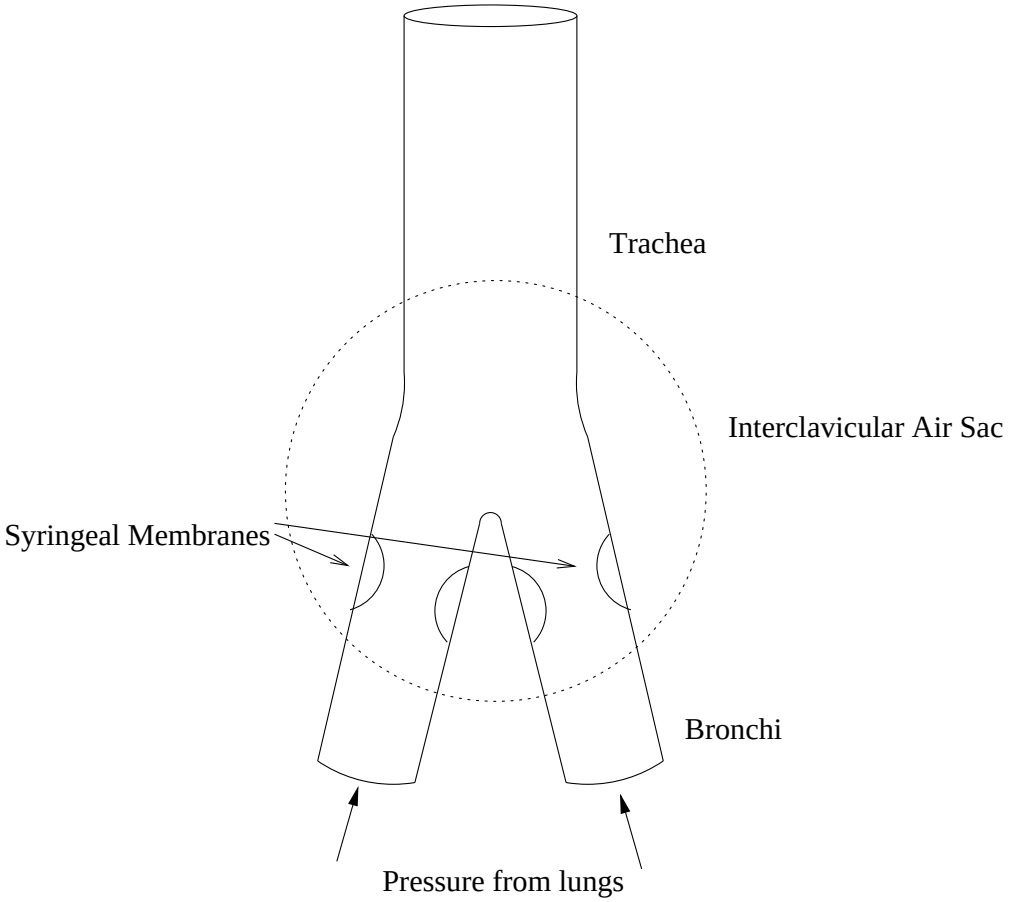


Figure 4.1: The syrinx.

4.2 The Syrinx Valve

Airflow in the syrinx begins from the lungs and passes through the bronchi, trachea and mouth before radiating from the beak. On its way, the airflow passes through the bird's primary vocal organ, the syrinx, made of non-linear vibrating membranes which lie within each bifid bronchi just below the junction with the trachea (see Figure 4.1).

During vocalization, the rise in pressure in the interclavicular air sac and the contraction of the syringeal muscles causes the third bronchial syringeal cartilage to rotate counterclockwise, bringing the lateral labium closer to the membrane and creating a constriction in the bronchial lumen (see Figure 4.2). This constriction along the membrane creates a pressure-controlled valve. When the bird is at rest, the syringeal membrane either lies either flat on the wall of the bronchus or

with a bulge protruding by some amount into the bronchial lumen (this is determined by the membrane's equilibrium position—its position in the absence of volume flow). When singing begins, the membrane vibrates, with motion occurring primarily toward the opposite cartilaginous wall of the bronchus (see Figure 4.4), creating a further narrowing in the bronchus through which the air flows, with the possibility of closing the air passage completely.

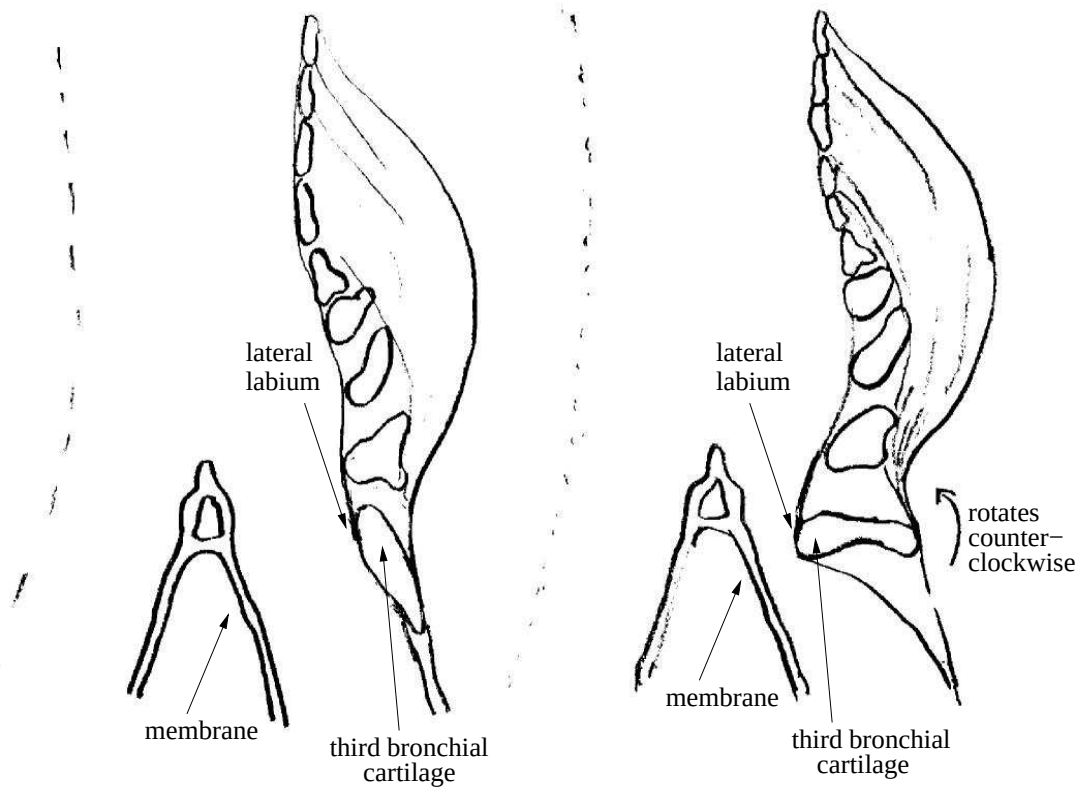


Figure 4.2: On the left the syrinx is in its normal respiratory position and on the right it's in its sound producing position.

4.2.1 Configurations of Pressure-Controlled Valves

The motion of the syringeal membrane can be described by two of three configurations for pressure-controlled valves in acoustic tubes [47]. The classification of the valve, thoroughly described by Fletcher in [29], is determined by the effect of an additional pressure applied to either of its two sides and defined by the couplet (σ_1, σ_2). The symbol σ_1 specifies the upstream behaviour of the valve in the presence of an additional upstream pressure while σ_2 specifies the downstream valve behaviour in the presence of an additional downstream pressure. A $\sigma_{1,2}$ value of $+1$ signifies that

side of the valve opens further, while a value of -1 signifies that it closes. This construction is very useful when evaluating the force used to drive the fundamental mode of the vibrating membrane (see Section 4.2.2).

The three possible configurations of a pressure-controlled valve are therefore:

1. $(-,+)$: the valve is blown closed (as in woodwind instruments or reed-pipes of the pipe organ).
2. $(+,-)$: the valve is blown open (as in the simple lip-reed models for brass instruments, the human larynx, harmonicas and harmoniums).
3. $(+,+)$: the transverse model where the Bernoulli pressure causes the valve to close perpendicular to the direction of airflow [31, 47].

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 [47]. A musical exploration of this model could, therefore, contribute to the synthesis of new sounds in computer music.

4.2.2 The Bernoulli Effect in the Syrinx Valve

In voiced song (as opposed to whistled song), the membrane is set into motion by air flow, creating variable heights in the valve channel. If the channel is leak-free, a fixed volume of air must pass any given point per second during air flow [26]. When a fluid (such as air) flows rapidly through a constricted aperture therefore, its flow velocity must increase. This causes an increase in the kinetic energy which cannot come from the fluid since energy in the fluid must be conserved. Since air tends to flow from high pressure regions to regions where the pressure is low, it can be deduced that the pressure in the valve channel is lower than that of the lower bronchus (the entrance to the valve). The increased energy therefore comes from this pressure difference between the upstream part of the bronchus and the pressure in the aperture. Pressure is due to kinetic energy, in fact, pressure is precisely proportional to kinetic energy density [92]. When a gas flows, a portion of its pressure may be converted to the kinetic energy of the flow [92].

The model of the valve displacement and the resulting pressure through the constriction is developed following an acoustic model by Fletcher [46] and is based on the mechanical properties of the membrane and the Bernoulli equation for air flow.

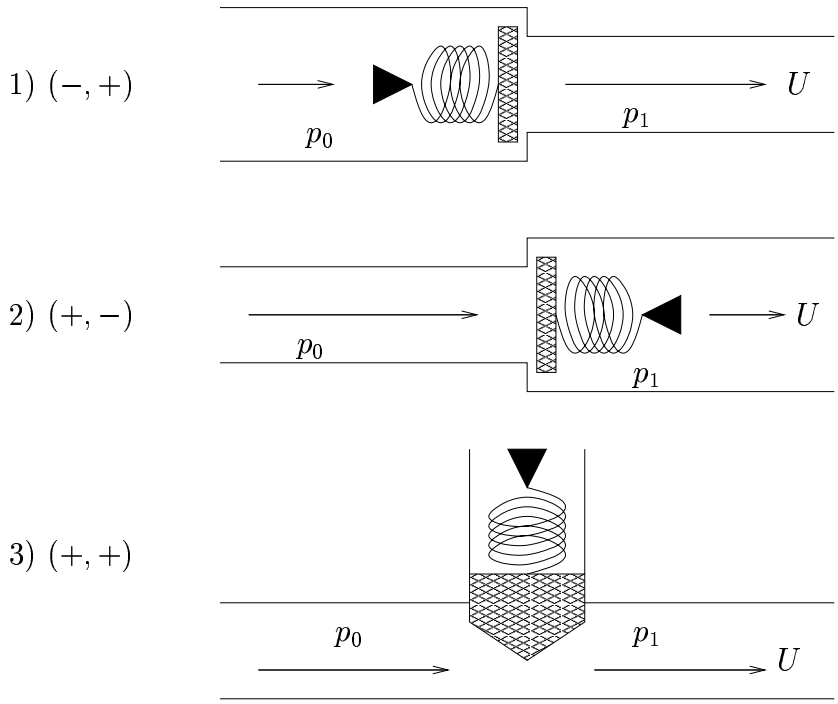


Figure 4.3: Simplified models of common configurations of the pressure-controlled valve. 1) $(-,+)$ defines a valve that is blown closed and is typical of woodwind instruments. 2) $(+,-)$ defines a valve that is blown open and is exemplified by brass and other lip-reed instruments as well as the human larynx. 3) $(+,+)$ is the principle configuration of the avian syrinx where an overpressure applied to either side of the valve will cause it to open since the motion of the valve is perpendicular to the direction of air flow [31].

The model has the following four variables which evolve during sound production:

$$\begin{aligned} p_0 &\triangleq \text{pressure on the bronchial side of the constriction} \\ U &\triangleq \text{air volume flow through the syrinx} \\ x &\triangleq \text{displacement of the membrane} \\ p_1 &\triangleq \text{pressure on the tracheal side of the constriction} \end{aligned} \tag{4.1}$$

The four model variables are simulated by discretizing their corresponding differential equations, each one very much dependent on the others.

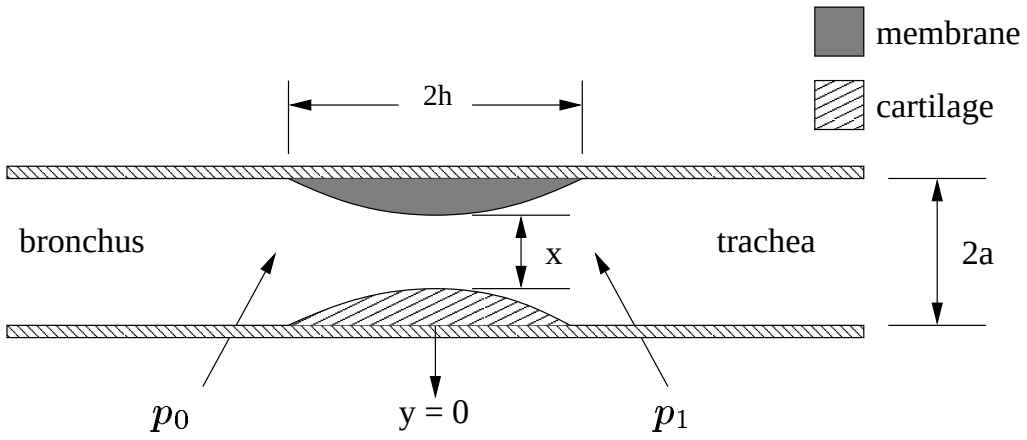


Figure 4.4: The pressure-controlled valve in the syrinx.

Bronchial pressure (p_0)

The lower bronchi are extremely short and are sufficiently modeled as having a volume, V, and an acoustic stiffness $\rho c^2/V$. The rate at which pressure builds up in the lower bronchus, dp_0/dt, is proportional to the difference in volume velocity flowing *in*, $(p_G - p_0)/Z_G$, and flowing *out*, U, of the constriction [46] and is given by

$$\frac{dp_0}{dt} = \left(\frac{\rho c^2}{V}\right)\left(\frac{p_G - p_0}{Z_G} - U\right), \tag{4.2}$$

where Z_G is defined as the impedance of the air sacs, p_G is some initial pressure gain from the lungs, ρ is the density of air, V is the volume of the bronchus and c is the speed of sound (in air).

Volume flow (U)

Since the motion of the membrane causes varying heights in the constriction, Bernoulli's equation (4.3) is used to determine the pressure at a given point y along the valve channel. Bernoulli's equation, in its more familiar form, is given by

$$p(y) = p_0 + \frac{\rho}{2}\left[v_0^2 - v(y)^2\right] \tag{4.3}$$

where v is the particle velocity.

If the membrane has a quadratic shape as shown in Figure 4.4, the height of the channel opening, $z(y)$, can be approximated by

$$z(y) \approx x + (a - x)(y/h)^2, \tag{4.4}$$

where a is the radius of the bronchus and h is half the length of the membrane (as shown in Figure

4.4) [46]. Given that particle velocity at a given point is equal to volume velocity divided by the cross-section area at that point, the following substitution can be made:

$$v(y) = \frac{U}{2a\left[x + (a - x)(y/h)^2\right]}. \tag{4.5}$$

At the point $y = 0$, in the center of the valve channel as labeled in Figure 4.4, the flow separates from the surface of the membrane and forms a jet which doesn't dissipate till beyond the base of the upper bronchus (the exhaust tube connected to the valve). The result is that the pressure acting on the entire tracheal half of the membrane is equal to the pressure at the base of the upper bronchus [46] and the pressure at point $y = 0$ is effectively p_1.

Given that the area of the opening of the valve is $2\pi a^2$ and the area at the point $y = 0$ can be approximated by $2ax$ using (4.4), equation (4.3) becomes

$$p_1 = p_0 + \frac{\rho}{2}\left[\left(\frac{U}{2\pi a^2}\right)^2 - \left(\frac{U}{2ax}\right)^2\right] \tag{4.6}$$

which can be approximated by

$$p_1 = p_0 - \frac{\rho}{2}\left(\frac{U}{2ax}\right)^2 \tag{4.7}$$

when $x \ll \pi a$ [46]. Equation (4.7) applies in a steady state. There is, however, a temporal change in flow resulting in a force which accelerates air through the syrinx [46]. To incorporating this force, a pressure drop of $\rho(dU/dt)/(2\sqrt{ax})$ must be added to (4.7). The final differential equation governing airflow for the initial syrinx model is then obtained by isolating dU/dt and is given by

$$\frac{dU}{dt} = \frac{2\sqrt{ax}}{\rho}\left(p_0 - p_1 - \frac{\rho}{8a^2x^2}U^2\right) \tag{4.8}$$

Given (4.8), it can be seen that there is a difficulty in computing the volume flow between an open and closed valve. Since x approaches zero as the valve closes, there is a singularity in this equation which makes it unstable. In the original model, this is handled by solving (4.8) only when x is greater than zero. If x is zero or less, as is the case with a closed valve, the flow was simply set to zero—since theoretically there should be no flow through a closed valve. Switching based on some level threshold in discrete time signals is a known cause of aliasing however, so the solution to this problem is later addressed in Chapter 5 to improve the model output.

Membrane motion (x)

The equation describing the displacement of the syrinx valve, like that of the brass instrument player's lips or the human larynx, is rather complicated because of the complexity of motion in biological valves [29]. Though in reality it is unlikely the valve can be accurately described by the

simplified scheme outlined in Section 4.2.1, the geometric parameters provide a satisfactory method and a good starting point for implementing a simulation.

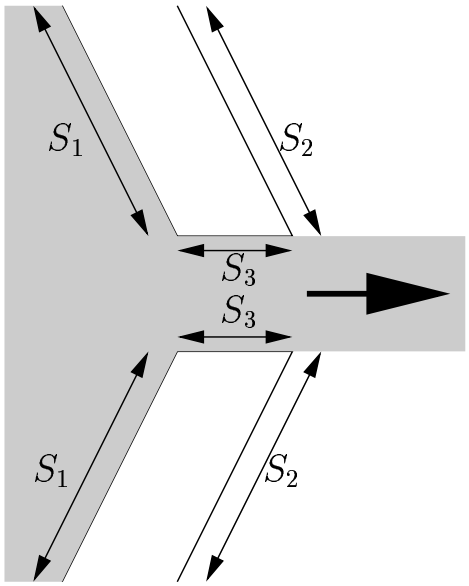


Figure 4.5: Geometry of a generalized pressure-controlled valve showing effective areas S_1, S_2, S_3 [29].

Considering the geometry seen in Figure 4.5, the surface area of the valve flap that sees the pressure p_0 (the pressure before the narrowing of the valve channel) is defined by S_1. The surface area of the valve that sees the flow at the interior of the valve channel is defined by S_3. The surface area at the downstream part of the valve—which no longer sees the volume flow because of the flow separation from the valve—sees the pressure p_1 and is defined by S_2. With these areas defined and with the geometric couplets defined in Section 4.2.1, the generalized equation describing the motion of the valve can be represented by

$$\frac{d^2x}{dt^2} + 2\kappa\frac{dx}{dt} + \omega_0(x - x_0) = \frac{1}{m}(\sigma_1 p_0 S_1 + \sigma_2 p_1 S_2) + \frac{1}{m}p_1 S_3, \tag{4.9}$$

where x_0 is the equilibrium position of the valve opening in the absence of flow [29]. The final term in (4.9) represents the influence of the Bernoulli pressure produced by flow through the valve channel [29].

The syrinx valve does not have surface areas so clearly defined as Figure 4.5. The division between S_1, S_2 and S_3 is not so angular nor is the point at which the flow separates from the surface of the membrane so ascertainable.

The key to finding the displacement equation is to first consider the force acting on the membrane, and then the equation for displacement can be described by the more familiar relationship

$$m_n \left[\frac{d^2 x_n}{dt^2} + 2\kappa \frac{dx_n}{dt} + \omega_n^2 (x_n - x_0) \right] = \epsilon_n F \tag{4.10}$$

where

$$\begin{aligned} \omega_n &\triangleq \text{the radian frequency of mode } n \\ m_n &\triangleq \text{the effective mass associated with mode } n \\ \kappa &\triangleq \text{the damping coefficient} \\ \epsilon_n &\triangleq \text{the coupling coefficient between } F \text{ and mode } n \end{aligned}$$

and F is the force driving the fundamental mode of the membrane. Given that the force on an area ,S_n, is equal to that area multiplied by the pressure across that area, we can consider the force acting on each section of the valve in order to determine the overall force acting on the membrane. Since the flow separates midway through the valve channel and forms a jet (see Figure 4.4), the pressure acting on the downstream part of the valve is simply p_1, and the force $p_1 2ah$. The force acting on the upstream part of the valve is found by integrating the pressure along y, given by

$$p(y) = p_0 + \frac{\rho}{2} \left[\left(\frac{U}{\pi a^2} \right)^2 - \left(\frac{U}{2az(y)} \right)^2 \right], \tag{4.11}$$

where $z(y)$ is the valve opening over this part of the area and is given by (4.4). Fletcher writes the equation for force driving the fundamental mode of the membrane's motion as [46]:

$$F = ah(p_0 + p_1) - \frac{2\rho U^2 h}{7(ax)^{1.5}} \quad \text{for } x > 0. \tag{4.12}$$

The overall displacement is obtained by summing the displacements for each individual mode (as many as one wishes to model) using (4.10). The frequency of the second and third modes for an isometric membrane is approximately $1.6\omega_0$ and $2\omega_0$ respectively. The higher modes become increasingly more closely spaced [46].

When the displacement of the membrane, x, is equal to zero, the membrane is touching the opposite wall and the valve channel is closed. In order to account for the damping that would occur should the membrane actually touch the opposite wall, a factor E is introduced into (4.10), that is

$$\kappa \to E\kappa \quad 10 \leq E \leq 100 \quad \text{if } x \leq 0. \tag{4.13}$$

The actual value of E depends on the stickiness of the contact between the membrane and the wall [46].

When the membrane's displacement becomes very large, surrounding tissue will likely take part

in the vibration [46]. To account for this, the mass m from (4.10) is replaced with the term

$$m \to m\left(1+\eta\left(\frac{x-x_0}{h}\right)^2\right) \tag{4.14}$$

so that the mass changes nonlinearly depending on the membrane's position (x).

The second derivative of x_n is isolated by rearranging (4.10)

$$\frac{d^2x_n}{dt^2} = \frac{\epsilon F}{m_n} - 2\kappa\frac{dx_n}{dt} - \omega^2(x_n - x_0), \tag{4.15}$$

and dx_n/dt and x_n are calculated using methods described in Section 4.2.3.

Upper bronchus pressure ($\mathbf{p_1}$)

A pressure-controlled valve must have some form of pipe connected to at least one of its ports in order for the valve to oscillate [29]. In the case of the syrinx, the upstream part of the valve is attached to the lower bronchus and the downstream part of the valve jets out into the upper bronchus (just below the junction with the trachea).

The pressure leaving the constriction, p_1, is proportional to the volume velocity, U, scaled by the characteristic impedance of the bronchus, $Z_0 = \rho c/(\pi a^2)$. However, in order to obtain the true pressure at the base of the upper bronchus, the pressure due to all previous reflections must also be considered [46]. The pressure p_1 is therefore not calculated by finding its derivative—as with the previous variables—but rather is obtained by summing the left and right traveling pressure waves at the modeled base of the upper bronchus.

4.2.3 Solving the syrinx valve model numerically

Once the derivatives of the variables p_0, x, U have been isolated, they must be discretized to digitally simulate the valve.

Backwards difference

Taking the derivative of a continuous time variable such as $x(t)$ using dx/dt is a continuous time operation. The differentiation operator can be approximated in discrete time using the substitution

$$\frac{dx}{dt} \approx \left(x_n - x_{n-1}\right)/T, \tag{4.16}$$

where T is the sampling period. This is called the *backwards difference* and it can be applied multiple times to obtain higher order derivatives. For example, the second derivative of x is given by

$$\frac{d^2x}{dt^2} \approx \left(x_n - 2x_{n-1} - x_{n-2}\right)/T. \tag{4.17}$$

The accuracy of the backwards difference approximation is dependent on the sampling period T, or alternatively, the sampling rate to which it is inversely proportionate. The smaller the sampling period T, or the larger the sampling rate, the more closely (4.16) and (4.17) will approximate the continuous time operator. The backwards difference is said to be first-order accurate in T [13].

The first implementation of the syrinx model used this approximation, but became very unstable when sampling rates were lowered to typical audio rates of 44100 or 48000 Hz. Given that the purpose of this model is to create a real-time musical instrument, it is clearly important that the model be stable at lower audio sampling rates. This algorithm is therefore insufficiently accurate for a sound synthesis simulation of this model.

The Trapezoid Rule for Numerical Integration

The second implementation used a more sophisticated algorithm which is second-order accurate in T, and uses the trapezoid rule for numerical integration [13, 14]. If the derivative of x is $y(t)$, that is, if

$$\frac{dx}{dt} = y(t), \tag{4.18}$$

then the value of $x(t)$ is found by integrating $y(t)$ over the interval $[0, t]$, that is

$$x(t) = \int_0^t y(t')dt' + x_0 \tag{4.19}$$

where $x_0 = x(0)$ and the value of x at the discrete time sample nT is given by

$$\begin{aligned} x(nT) &= \left(\int_0^{nT} y(t')dt'\right) + x_0 \\ &= \left(\int_0^{(n-1)T} y(t')dt'\right) + x_0 + \int_{(n-1)T}^{nT} y(t')dt' \\ &= x((n-1)T) + \int_{(n-1)T}^{nT} y(t')dt' \\ &\approx x((n-1)T) + \frac{T}{2}\left(y((n-1)T) + y(nT)\right), \end{aligned}$$

which yields the scheme

$$x_n - x_{n-1} = \frac{T}{2}(y_n + y_{n-1}) \tag{4.20}$$

known as the *trapezoid rule* for numerical integration [14]. This is equivalent to the bilinear transform which is defined by the substitution

$$s = c\frac{1 - z^{-1}}{1 + z^{-1}}, \quad c > 0 \tag{4.21}$$

where $c = 2/T$ [14]. The bilinear transform maps real frequencies to real frequencies, taking the right and left hand side of the s-plane to the outside and inside of the unit circle, respectively, in the z-plane (see figure 4.6) [14].

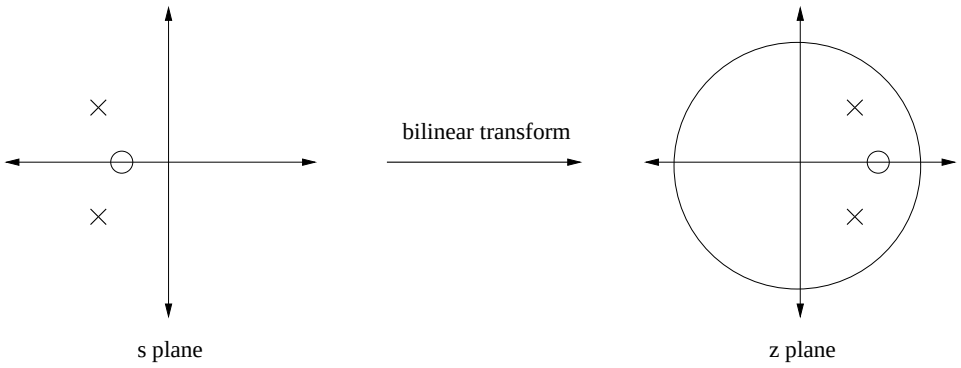


Figure 4.6: Bilinear Transform Mapping from s to z planes [14].

Applying the Trapezoid Rule to the Syrinx Model

The current model isolates the derivatives of each of the variables in (4.1)—except p_1 which is determined using the waveguide synthesis techniques outlined in Section 4.3.1—and finds a delta using the the trapezoid rule defined in (4.20). This delta is then added to the corresponding variable state in order to obtain its updated value. If $x'[n]$ and $x''[n]$ are the discrete time equivalents of dx/dt and d^2x/dt^2 respectively, then $x[n]$ and $x'[n]$ are updated using the discretized first and second-order derivatives yielding the following scheme:

$$\begin{aligned} x[n] &= x[n-1] + (x'[n] + x'[n-1])\left(\frac{T}{2}\right) \\ x'[n] &= x'[n-1] + (x''[n] + x''[n-1])\left(\frac{T}{2}\right) \end{aligned} \tag{4.22}$$

The model of the vibrating membrane can be reduced to differential equations (4.2), (4.8) and (4.15), with the value for p_1 being returned by the digital waveguide modeling the upper bronchus. When discretized using the trapezoid rule, this system of differential equations is given by

$$\begin{aligned} p_0' &= \left(p_0' + \left(\frac{\rho c^2}{V}\right)\left(\frac{p_G - p_0}{Z_G} - U\right)\right)\frac{T}{2} \\ U' &= \left(U' + D\left(p_0 - p_1 - \frac{\rho}{8a^2x^2}U^2\right)\frac{T}{2}\right) \quad 0 < x \leq a \\ x_n'' &= \left(x_n'' + \frac{\epsilon F}{m_n} - 2\kappa\frac{dx_n}{dt} - \omega^2\left(x_n - x_0\right)\right)\frac{T}{2}. \end{aligned} \tag{4.23}$$

All variables and their derivatives are initially set to zero. When a pressure from the lungs is introduced into the system by setting p_G to some initial value, the variables are free to evolve in the model.

4.3 Modeling the Upper Bronchi and Trachea

There is a pressure leaving the syrinx valve that is proportional to the volume velocity, U, scaled by the characteristic impedance of the upper bronchus, Z. The characteristic impedance describes the ratio of pressure, p, to volume velocity, U, by the relationship

$$Z = \frac{p}{U}. \tag{4.24}$$

In order to obtain the true pressure at the base of the upper bronchus however, the pressure from all previous reflections in the upper bronchi and trachea must also be considered [46]. Digital waveguide synthesis techniques easily account for such reflections and are therefore used to model the pressure waves in the bronchi and the trachea tubes. These pressure waves are ultimately used to obtain the value at the base of the upper bronchus, p_1, and in turn, used in the syrinx valve model to calculate the volume flow derivative (4.8) from Section 4.2.

4.3.1 Waveguide synthesis

The general solution to the one-dimensional wave equation was published by d'Alembert in 1747 and is given by

$$p(x,t) = p^+(x - ct) + p^-(x + ct) \tag{4.25}$$

for right (indicated by the '+' superscript) and left (indicated by the '−' superscript) longitudinal traveling pressure waves, traveling at speed c.

Bidirectional delay-lines, or digital waveguides, are computational physical models which sample the acoustic traveling-waves described by 4.25. The delay number, expressed by the '$-n$' exponent

in Figure 4.7, represents the wave variable in a medium, such as pressure in an acoustic tube or displacement along a string [12]. As seen by d'Alembert's equation (4.25), the value of the physical variable at any point along the medium can then be obtained by summing the corresponding location on the upper and lower rails of the waveguide (see Figure 4.7).

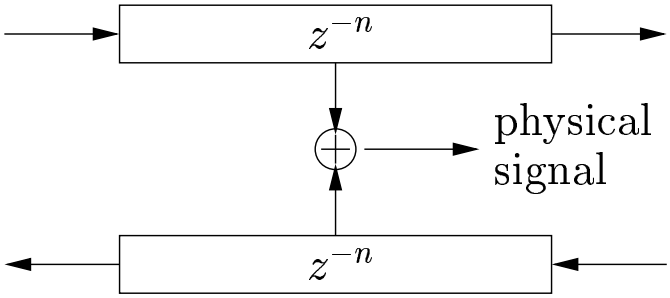


Figure 4.7: A simple digital waveguide model for one-dimensional wave propagation [12].

A digital waveguide is used to model the pressure waves traveling toward the beak and toward the syrinx along the bronchi and tracheal tubes. The pressure at the base of the upper bronchi, p_1, is obtained by summing the sampled traveling pressure waves at the input of the upper rail and at the output of the lower rail (see Figure 4.27). The input pressure to the upper rail is equal to the pressure leaving the syrinx, $p = ZU$, added to the reflected pressure from the lower rail of the upper bronchi waveguide (see Figure 4.27). The characteristic impedance is determined by the tube radius and is given by

$$Z_0 = \frac{\rho c}{A} = \frac{\rho c}{\pi a^2} \tag{4.26}$$

where ρ is the density of air and $A = \pi r^2$ is the cross-section area of the tube.

The characteristic impedance is usually constant over the length of a uniform tube and though this may not be exactly accurate in the biological case of vocal tract tubes, it is assumed close enough to permit this simplification. However given the discontinuity at the junction between the two bronchi and the trachea and the difference between the cross-sectional area of these three acoustic elements, there will certainly be a significant change in the characteristic impedance within the syrinx's airway for which the model must account. When the pressure waves reach this discontinuity in the air column, part of the wave will be transmitted and part will be reflected backward. In order to model this effect, a three-port parallel junction is used.

4.3.2 Three-Port Parallel Scattering Junction

A discontinuity in the diameter of an acoustic tube will cause traveling waves to be partially reflected, that is, only part of the wave will be transmitted while a part is reflected back. A discontinuity also occurs in the bird's vocal tract at the point where the trachea which divides into the left and right bronchi (see Figure 4.1) and results in *scattering*, a phenomenon in which the wave is scattered into

an infinity of waves propagating in different directions [82]. The model of the vocal tract therefore requires three separate digital waveguides—one for the left and right upper bronchi and one for the trachea—and a three-port parallel junction to model the point at which they connect. The three-port junction is commonly used to model tone holes in acoustic wind instruments which can be viewed as three connected tubes, with the finger hole being a very short branch off the main bore [9]. It is also use to model the bifurcation at the velum in the human vocal tract, simulating the oral and nasal airways [5].

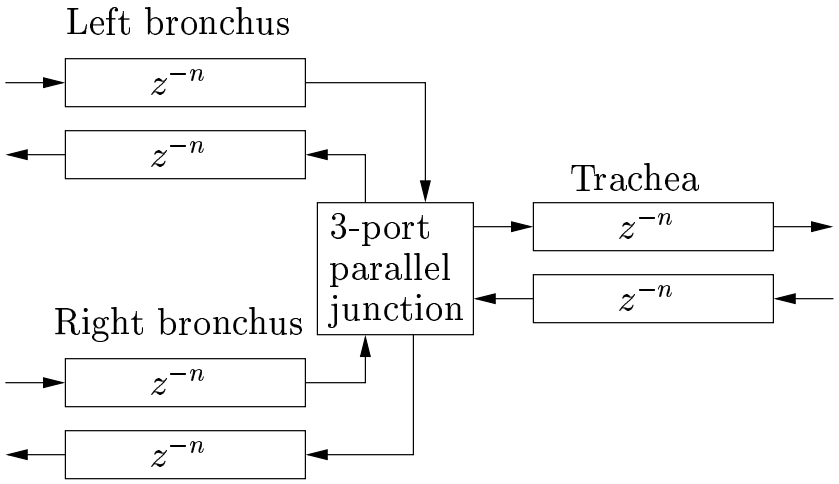


Figure 4.8: Three digital waveguides meeting at a three-port parallel junction.

The law for conservation of mass and momentum dictate that the pressure must be continuous at the junction. Therefore for a junction with N ports,

$$p_J = p_n, \qquad n = 1 \ldots N \tag{4.27}$$

where p_J is the pressure at the junction and p_n is the pressure at each of the N ports. By the same law, the volume flow at the N-port junction sums to zero,

$$\sum_{n=1}^{N} U_n = 0, \tag{4.28}$$

where U_n is the volume flow at a port (see Figure 4.9).

Each port has both an input and output (see Figure 4.9). The pressure at each port is equal to the sum of the incoming and outgoing pressures and is given by

$$p_n = p_n^+ + p_n^- \tag{4.29}$$

where the p_n^+ indicates the pressure entering the port and p_n^- indicates the pressure leaving the port

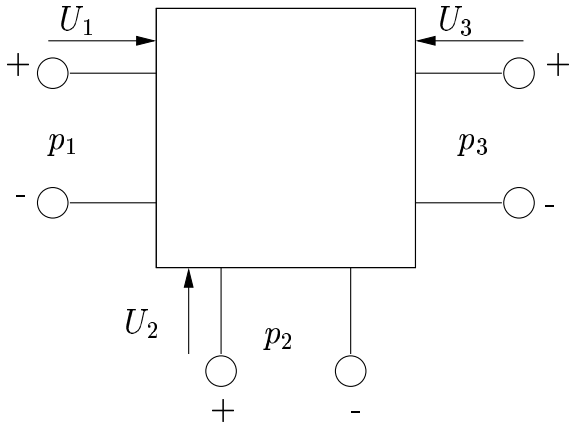


Figure 4.9: Three-Port Parallel Junction.

(see Figure 4.9). Because of pressure continuity at the junction, the pressure on a port, p_n from (4.29), can be substituted by the total junction pressure, p_J, and the pressure leaving a port is determined by

$$p_n^- = p_J - p_n^+. \tag{4.30}$$

Therefore, in order to determine the outgoing pressure on any port based on the incoming pressure, the pressure at the junction, p_J, must first be determined.

Since the characteristic impedance is the proportion of pressure to volume velocity (see 4.24) the pressure components traveling in and out of the ports are given by

$$p_n^+ = Z_n U_n^+ \tag{4.31}$$

$$p_n^- = -Z_n U_n^- \tag{4.32}$$

where the negative sign in (4.32) accounts for the fact that flow must move in the direction in which it is generating pressure [12]. The pressure at the junction can be derived beginning with the relationships given in 4.31 and 4.32 and the fact that the flow at the junction sums to zero (from 4.28). That is,

$$\sum_{n=1}^{N} \left(U_n^+ + U_n^-\right) = 0 \tag{4.33}$$

$$\sum_{n=1}^{N} \left(\frac{p_n^+}{Z_n} - \frac{p_n^-}{Z_n}\right) = 0 \tag{4.34}$$

Substituting the pressure leaving a port by the expression given by (4.30), the total pressure at the

junction is obtained by

$$\sum_{n=1}^{N} \frac{p_n^+}{Z_n} - \frac{p_J - p_n^+}{Z_n} = 0$$

$$\sum_{n=1}^{N} \frac{2p_n^+}{Z_n} - \frac{p_J}{Z_n} = 0$$

and finally,

$$p_J = \frac{2\sum \Gamma_n p_n^+}{\sum \Gamma_n}, \tag{4.35}$$

where $\Gamma_n = 1/Z_n$ is the characteristic admittance and is inversely related to the characteristic impedance.

The model of the three-port junction will receive an incoming pressure wave, p_n^+, on each of the three ports—one for each of the two bronchi and the trachea—and will return the corresponding outgoing pressure value, p_n^-.

4.4 Reflection and Transmission Filters

Radiation losses at the beak are accounted for using a lowpass filter, the output of which is negated to simulate the reflection of an open end tube.

Even the simplest waveguide models of an acoustic tube benefit from some kind of low-pass reflection filter to account for the losses that occur at the end of the tube and then a complimentary high-pass transmission filter allowing energy that wasn't reflected to radiate from the end of the tube (see Figure 4.7). Filters that account for losses due to wall attenuation are also desirable and will be discussed in detail in Section 4.5.

4.4.1 Reflection at Open End of a Cylindrical Pipe

If a pipe extends from $x = 0$ to $x = L$, the pressure at the end of the pipe will be the sum of the right and left traveling waves at the point $x = L$, that is,

$$p(L,t) = [Ae^{-jkL} + Be^{jkL}]e^{jwt}. \tag{4.36}$$

Likewise the volume velocity at the end of the pipe will be

$$U(L,t) = \frac{1}{Z_0}\left[Ae^{-jkL} - Be^{jkL}\right]e^{jwt}, \tag{4.37}$$

where the relationship of traveling volume velocity waves to traveling pressure waves is given in (4.32) and where Z_0 is the characteristic impedance as defined in (4.26). The waves are given as

complex amplitudes, Ae^{-jkx} and Be^{jkx}, following the notation of [31], to include phase factors. The flow and pressure at the point $x = L$, where L is the length of the pipe in spatial samples, are related by the impedance at the end of the pipe, Z_L, given by

$$Z_L = \frac{p(L,t)}{U(L,t)} = Z_0 \frac{[Ae^{-jkL} + Be^{jkL}]}{[Ae^{-jkL} - Be^{jkL}]}. \tag{4.38}$$

This relationship, (4.38), makes it possible to determine the ratio of the right to the left traveling pressure wave at the end of the pipe by the following:

$$\begin{aligned} Z_L(Ae^{-jkL} - Be^{jkL}) &= Z_0(Ae^{-jkL} + Be^{jkL}) \\ Z_L Ae^{-jkL} - Z_0 Ae^{-jkL} &= Z_L Be^{jkL} + Z_0 Be^{jkL} \\ Ae^{-jkL}(Z_L - Z_0) &= Be^{jkL}(Z_L + Z_0) \\ \frac{B}{A} &= e^{-2jkL}\left[\frac{Z_L - Z_0}{Z_L + Z_0}\right] \\ \frac{B}{A} &= e^{-2jkL}\left[\frac{Z_L/Z_0 - 1}{Z_L/Z_0 + 1}\right]. \end{aligned} \tag{4.39}$$

The terminating impedance, Z_L, is given by

$$Z_L = R + jX \tag{4.40}$$

where R is the acoustic resistance and X is the acoustic reactance. The values of R and X as a function of the dimensionless quantity ka, where $k = (\omega/c)$ and c is the the speed of sound in air, may be approximated by

$$R/Z_0 \approx \begin{cases} \dfrac{(ka)^2}{2}, & ka \ll 1 \\ 1.0, & ka \gg 1 \end{cases} \tag{4.41}$$

$$X/Z_0 \approx \begin{cases} \dfrac{8ka}{3\pi}, & ka \ll 1 \\ \dfrac{1}{2ka}, & ka \gg 1 \end{cases} \tag{4.42}$$

of [31, p. 200]. Further, the ratio Z_L/Z_0 may be approximated by

$$Z_L/Z_0 \approx \frac{s}{1+s} \tag{4.43}$$

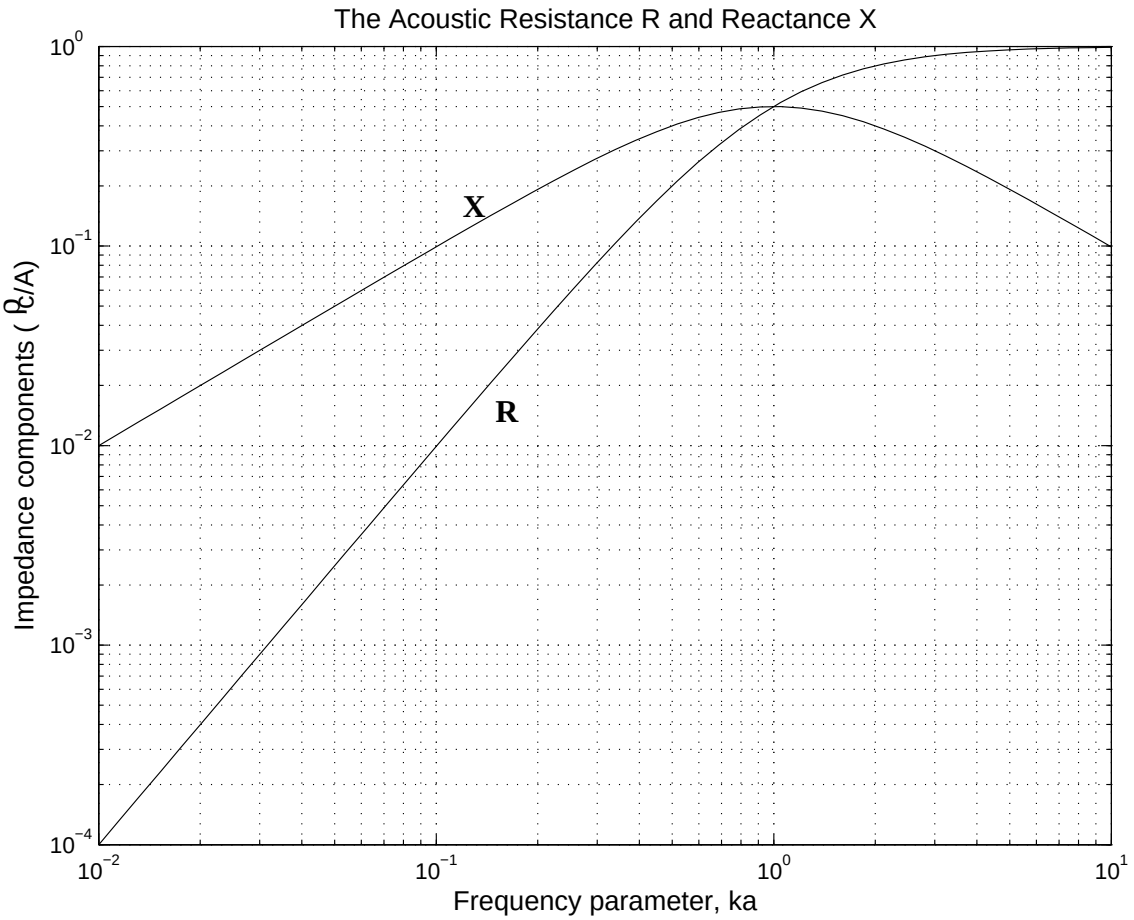


Figure 4.10: The acoustic resistance R and the acoustic reactance X as approximated by (4.45) and (4.46), both in units of Z_0 $(\rho c/\pi a^2)$ for the open end of a circular cylindrical pipe, as functions of the frequency parameter ka.

where $s = ka\sqrt{-1}$ [79]. Following (4.39), the corresponding reflection filter is given by

$$\begin{aligned} H_R(s) &= \frac{Z_L/Z_0 - 1}{Z_L/Z_0 + 1} \\ &= \frac{-1}{1+2s} \end{aligned} \tag{4.44}$$

The real part, R, in (4.40) is obtained by multiplying both the numerator and denominator by $(1+\bar{s})$ and is given by

$$R = \frac{|s|^2}{1+|s|^2} \tag{4.45}$$

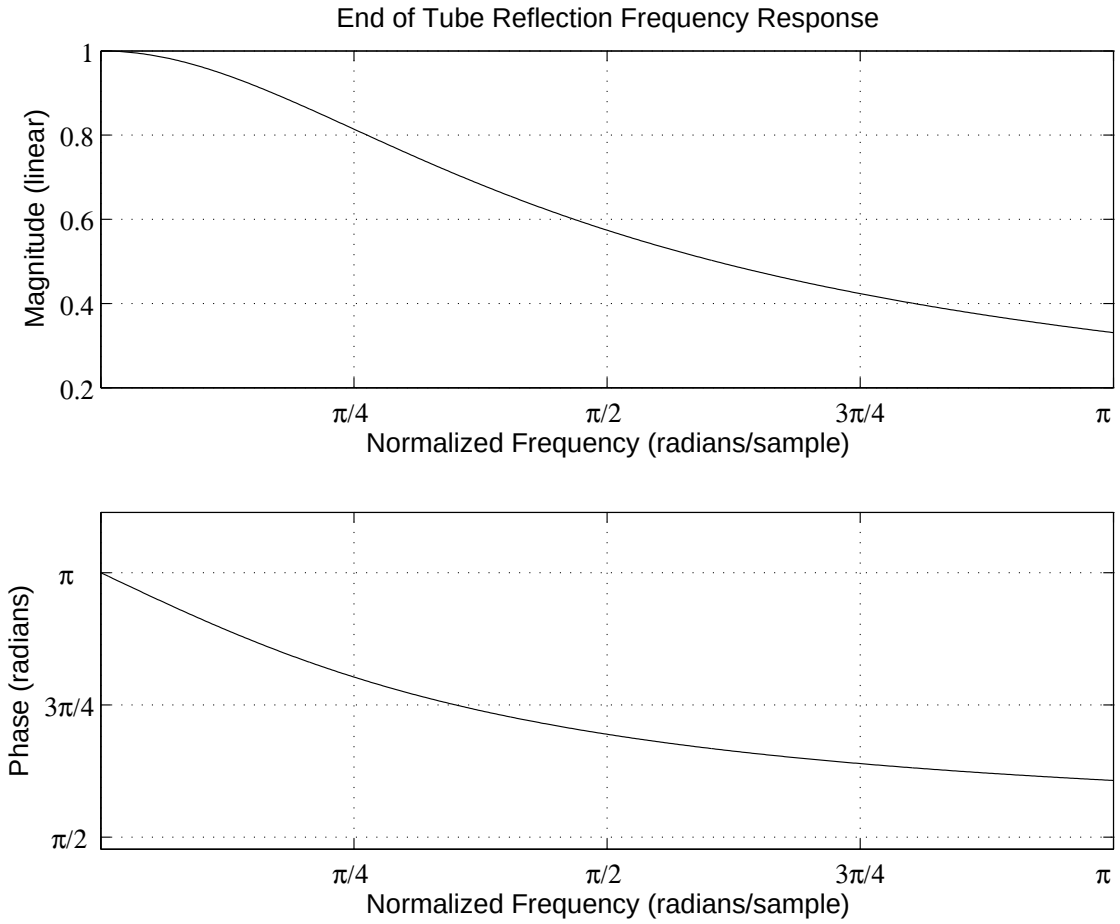


Figure 4.11: The frequency response of the reflection B/A based on the approximation (4.43) and the value of R and X given by (4.45) and (4.46) as plotted in Figure 4.10.

and likewise the imaginary part (X in (4.40)) is given by

$$X = \frac{s}{1 + |s|^2}. \tag{4.46}$$

Both (4.45) and (4.46) are plotted in Figure 4.10 showing they make satisfactory approximation to the behaviour of R and X up to the cutoff frequency $ka < 2$ as illustrated in [31, p. 200]. The corresponding frequency response given by (4.44) is plotted in Figure 4.11.

4.4.2 First-Order Low Pass Reflection Filter

Given the low-pass characteristic of the reflection frequency response as shown in Figure 4.11, it seems reasonable to assume that a first-order low-pass shelf filter (see Figure 4.12), with a DC gain

of one, a band edge gain of zero, and a transition frequency (ω_T) gain of σ_T will match the data well.

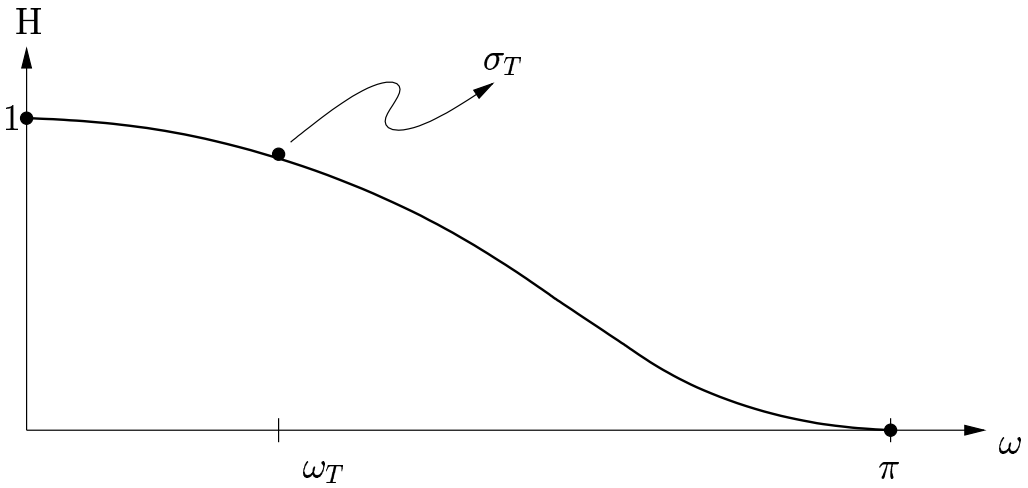


Figure 4.12: First-order shelf filter where $H(\omega) = 0$ at $\omega = \pi$ and $H(\omega) = 1$ at $\omega = 0$. The cutoff frequency is given by ω_T and $H(\omega_T)$ is given by σ_T.

The first-order transfer function for the low-pass shelf filter is given by

$$H(z) = b_0 \left(\frac{1 + z^{-1}}{1 + a_1 z^{-1}} \right). \tag{4.47}$$

It can be seen from (4.47) that the band-edge gain is zero, that is

$$\begin{aligned} H(\pi) &= b_0 \left(\frac{1 + e^{-j\pi}}{1 + a_1 e^{-j\pi}} \right) \\ &= b_0 \left(\frac{1 - 1}{1 - a_1} \right) \\ &= 0. \end{aligned}$$

If the DC gain is one, then the coefficient b_0 is solved by

$$\begin{aligned} H(0) &= b_0 \left(\frac{1 + e^{-j0}}{1 + a_1 e^{-j0}} \right) \\ &= b_0 \left(\frac{2}{1 + a_1} \right) \\ &= 1, \end{aligned}$$

and

$$b_0 = \frac{1 + a_1}{2}. \tag{4.48}$$

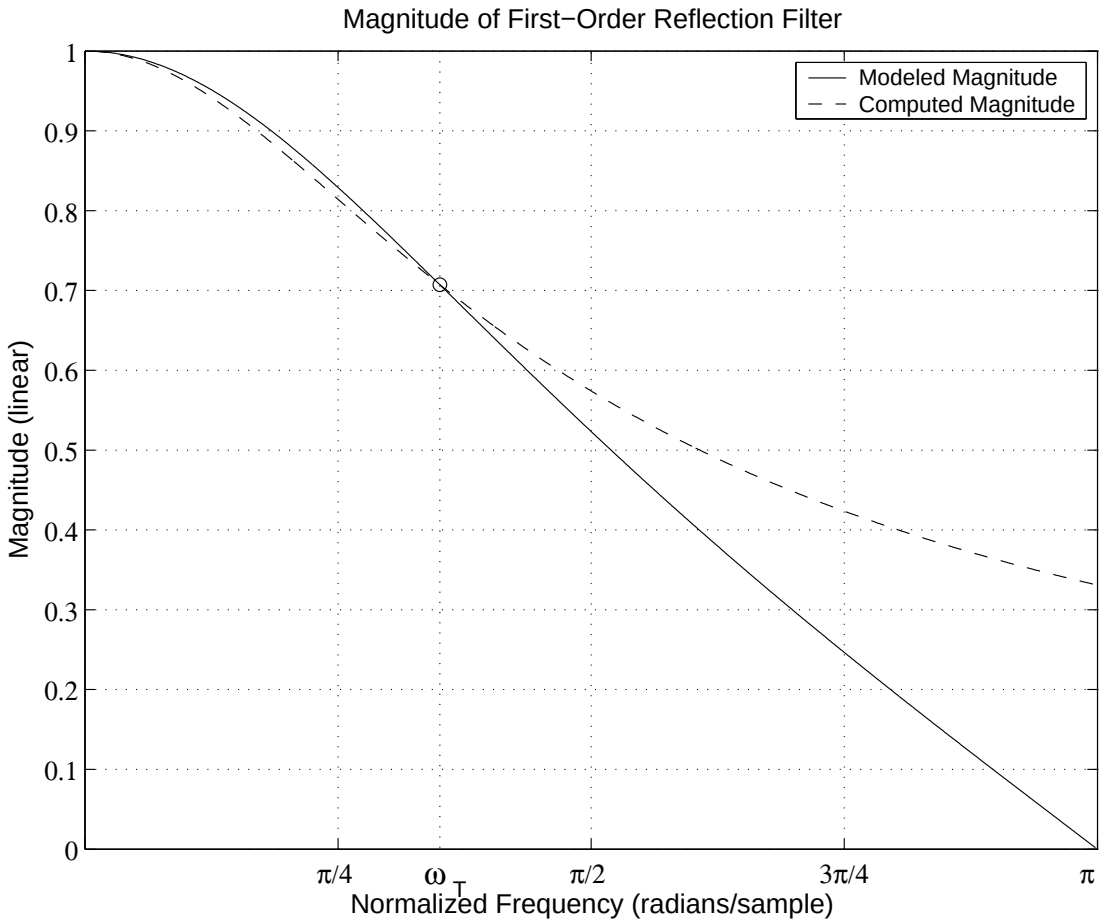


Figure 4.13: The frequency response of the reflection filter where ω_T is computed at a sampling rate of 44100, representing an actual frequency value of 13.178 kHz.

To solve for the feedback coefficient a_1, the value of $|H(\omega)|^2$ at $\omega = \omega_T$ is determined by

$$\begin{aligned} \sigma_T^2 &= |H(\omega_T)|^2 = \left| b_0 \left(\frac{1 + e^{-j\omega_T}}{1 + a_1 e^{-j\omega_T}} \right) \right|^2 \\ &= b_0^2 \left(\frac{2 + 2\cos\omega_T}{1 + a_1^2 + 2a_1 \cos\omega_T} \right) \qquad \text{since} \quad |z|^2 = z\bar{z}. \end{aligned} \tag{4.49}$$

Using the value for b_0 given in (4.48), (4.49) becomes

$$(1 + a_1^2 + 2a_1)(2 + 2\cos\omega_T) = 4\sigma_T^2(1 + a_1^2 + 2a_1\cos\omega_T) \tag{4.50}$$

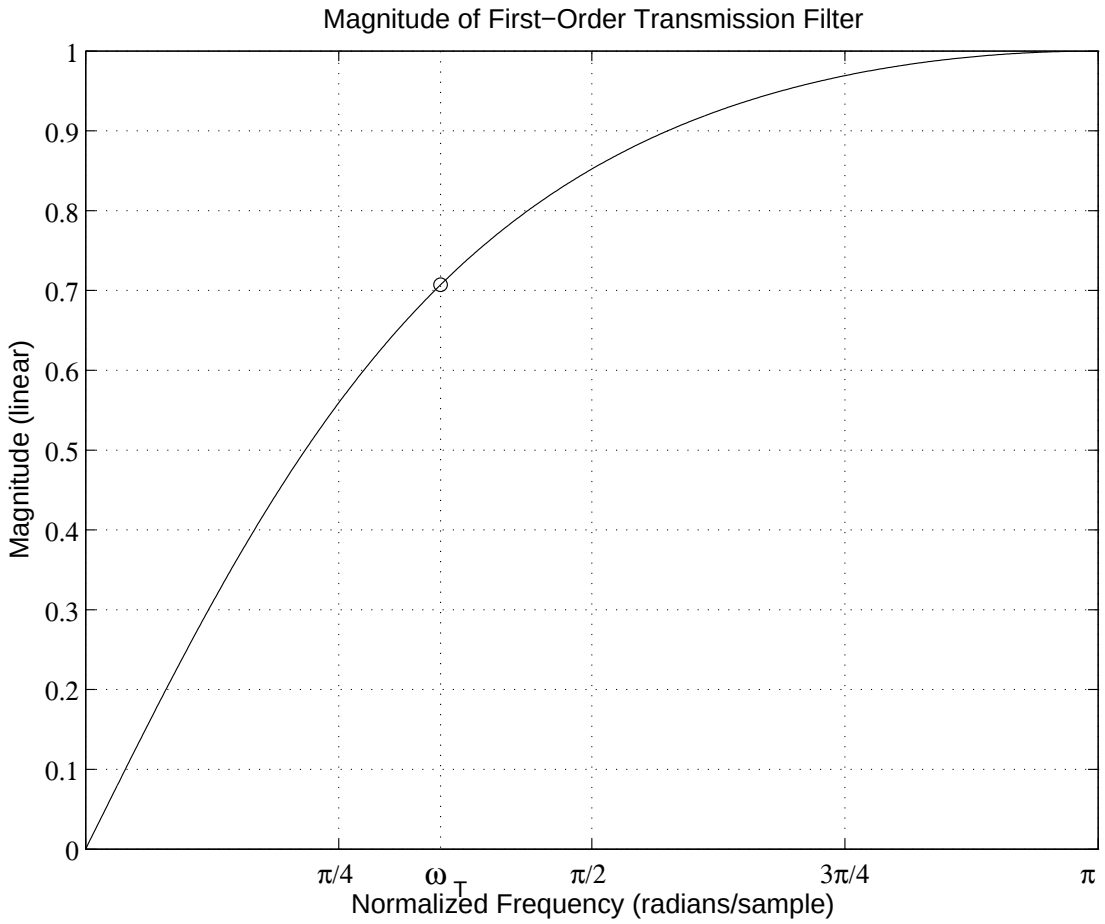


Figure 4.14: The corresponding frequency response of the transmission filter.

which can be reduced to

$$a_1^2(1+\cos\omega_T - 2\sigma_T^2) + 2a_1(1+\cos\omega_T - 2\sigma_T^2\cos\omega_T) + (1+\cos\omega_T - 2\sigma_T^2) = 0 \tag{4.51}$$

or

$$a_1^2 + 2a_1\alpha + 1 = 0 \tag{4.52}$$

where

$$\alpha = \frac{1+\cos\omega_T - 2\sigma_T^2\cos\omega_T}{1+\cos\omega_T - 2\sigma_T^2}. \tag{4.53}$$

The value of the filter coefficient a_1 (solved using the quadratic formula) is then given by

$$a_1 = -\alpha \pm \sqrt{\alpha^2 - 1}. \tag{4.54}$$

The transition frequency ω_T is determined using the quantity ka at the cutoff frequency (approximated by $ka = 0.5$) and is given by

$$w_T = 0.5 \left(\frac{c}{a}\right) T, \tag{4.55}$$

where T is the sampling period, c is the speed of sound and a is the radius of the open end of the tube. The magnitude at that frequency, σ_T, is determined using the magnitude of $H_R(s)$ given by (4.44) evaluated at $s = 0.5\sqrt{-1}$. The magnitude of the first-order shelf filter $H_R(z)$, modeling the reflection at the end of the trachea (with a radius of 3.5 mm), is given in Figure 4.13. The discrepancy between the computed and modeled reflection filters is beneficial in reducing the effects of aliasing.

It can be assumed that what is not reflected at the end of the trachea is transmitted through the beak. The corresponding transmission/radiation filter is therefore given by the following transfer function for a high pass filter:

$$\begin{aligned} H_T(z) &= 1 - \frac{b_0(1+z^{-1})}{1+a_1 z^{-1}} \\ &= \frac{1+a_1 z^{-1} - b_0(1+z^{-1})}{1+a_1 z^{-1}} \\ &= \frac{1-b_0+(a_1-b_0)z^{-1}}{1+a_1 z^{-1}} \end{aligned}$$

where the coefficients b_0 and a_1 are defined in (4.48) and (4.54) respectively. The frequency responce of the transmission filter is plotted in Figure 4.14.

4.5 Modeling Wall Attenuation

In practical implementations of acoustic tubes using digital waveguide synthesis, it is necessary to account for the losses associated with viscous drag and thermal conduction which take place primarily within a thin boundary layer along the bore walls [9, pp. 127–129]. Though these losses are *distributed* over the length of the tube, in a digital waveguide implementation it is more efficient to lump these effects by *commuting* a characteristic digital filter to each end of the waveguide (and any intervening observation points) [11, 15], as illustrated in Figure 4.15.

In [25] Benade gives formulas for calculating the propagation constant in a cylindrical tube of arbitrary radius, and presents simple approximations, valid for either low frequencies (and/or small tubes) or high frequencies (and/or large tubes). Later Keefe [33] presents separate higher-order approximations for large and small tubes. The propagation constant gives the phase velocity and the attenuation coefficient necessary for designing the loss filter $\lambda(\omega)$ which is commonly derived using the large tube approximation [9, pp. 27–29].

In this section, material previously published by Abel, Smyth and Smith in [1] is presented,

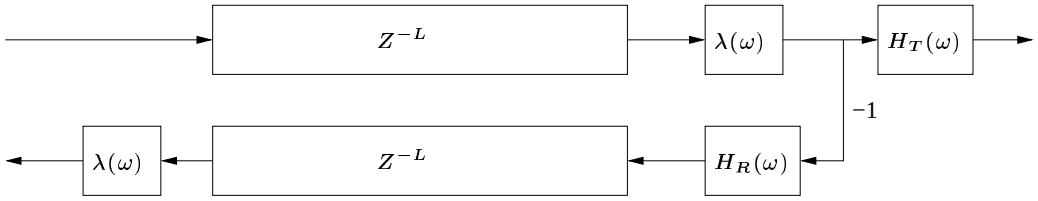


Figure 4.15: A waveguide model of a cylindrical tube with commuted wall loss filters, $\lambda(\omega)$, at upper and lower delay line observation points, a reflection filter $H_R(\omega)$ and a transmission filter $H_T(\omega)$.

which offers a uniform approximation to the attenuation and phase velocity models of Benade and Keefe. Based on this uniform approximation it was observed that the loss filter is minimum phase and several minimum phase filter design methods are evaluated, including a cascade of first-order shelf filters with parameters specified by simple functions of tube radius, tube length, and sampling rate. All filter design methods have comparable accuracy for a given order of digital filter, while the shelf filter cascade has the advantage of design simplicity and is nicely parameterized for time varying tube geometries.

4.5.1 Thermoviscous Losses in Cylindrical Conduits

In a cylindrical conduit, the walls contribute a viscous drag dependent on the ratio of the pipe radius a to the thickness of the viscous boundary layer given by the parameter

$$r_v = a\left(\frac{\eta}{\omega\rho}\right)^{-\frac{1}{2}}, \tag{4.56}$$

where ρ is the density of air, η is the viscosity and ω the radian frequency [31, pp. 193–196].

Further losses may occur due to a thermal exchange between the air and the walls with a magnitude dependent on the ratio of the tube radius a to the thermal boundary layer thickness given by the parameter

$$r_t = a\left(\frac{\kappa}{\omega\rho C_p}\right)^{-\frac{1}{2}}, \tag{4.57}$$

where κ is the thermal conductivity and C_p is the specific heat of air at constant pressure [31, pp. 193–196]. The two ratios are related by the square root of the Prandtl number ν^2 [25], that is,

$$r_t = \nu r_v, \qquad \nu = [C_p\eta/\kappa]^{\frac{1}{2}}. \tag{4.58}$$

Values for gas constants for air are given in Table 4.1.

The effects of the viscous and thermal losses lead to an attenuation in the waves propagating

down the length of the pipe. The propagation constant is given by

$$\Gamma(\omega) = \alpha(\omega) + j\frac{\omega}{v(\omega)}, \tag{4.59}$$

where $\alpha(\omega)$ is the attenuation coefficient, ω is the radian frequency and $v(\omega)$ is the phase velocity of wave disturbances [25].

Benade gives the following approximations for both the phase velocity $v(\omega)$ and the attenuation coefficient $\alpha(\omega)$ in the limit of small and large tube radius boundary-layer thickness ratios [25]:

$$v(\omega) = \begin{cases} c\dfrac{r_v}{2\sqrt{\gamma}}, & r_v \ll 1 \\ c\left[1 - \dfrac{1}{r_v\sqrt{2}}\left(1 + \dfrac{(\gamma-1)}{\nu}\right)\right], & r_v \gg 1 \end{cases} \tag{4.60}$$

$$\alpha(\omega) = \begin{cases} \left(\dfrac{\omega}{c}\right)\dfrac{2\sqrt{\gamma}}{r_v}, & r_v \ll 1 \\ \left(\dfrac{\omega}{c}\right)\dfrac{1}{r_v\sqrt{2}}\left(1 + \dfrac{(\gamma-1)}{\nu}\right), & r_v \gg 1. \end{cases} \tag{4.61}$$

Quantity	Symbol	Value	Units
air density	ρ	1.18×10^{-3}	g/cm^3
viscosity	η	1.85×10^{-4}	g/(s $\times$ cm)
Prandtl	ν^2	0.841	
ratio of specific heats	$\gamma = C_p/C_v$	1.40	
free space sound speed	c	3.47×10^4	cm/s

Table 4.1: Gas constants for air evaluated at 300°K (26.86°C) [25].

It is possible to account for attenuation due to wall losses simply by multiplying delay line outputs by a single coefficient β [46, page 466]. For a tube of length L and radius a, with observation points at the ends of the upper and lower delay lines, the constant β has the following approximate value:

$$\beta \approx 1 - 2\alpha L, \qquad \alpha \approx 2 \times 10^{-5}\omega^{\frac{1}{2}}/a, \tag{4.62}$$

where the approximation is valid for tubes sufficiently short that β is near one.

Since losses are frequency dependent, this coefficient is only a rough estimate. In practice, the simple large r_v expressions are commonly used to design wall loss filters for musical wind instrument

physical models (see, e.g. [9, page 129]) and given the relatively large bores of most instruments, this is likely an acceptable approximation. These expressions may not, however, be sufficiently accurate for smaller tubes like those found in some biological sound production mechanisms such as the bronchi and trachea in birds [19]. Therefore, in order to have a more versatile acoustic tube model, it is preferable for the wall loss filter to make use of a single less restrictive expression for phase velocity and attenuation—one that is compatible with all tube sizes and frequencies.

4.5.2 A Uniform Approximation for Phase Velocity and Attenuation

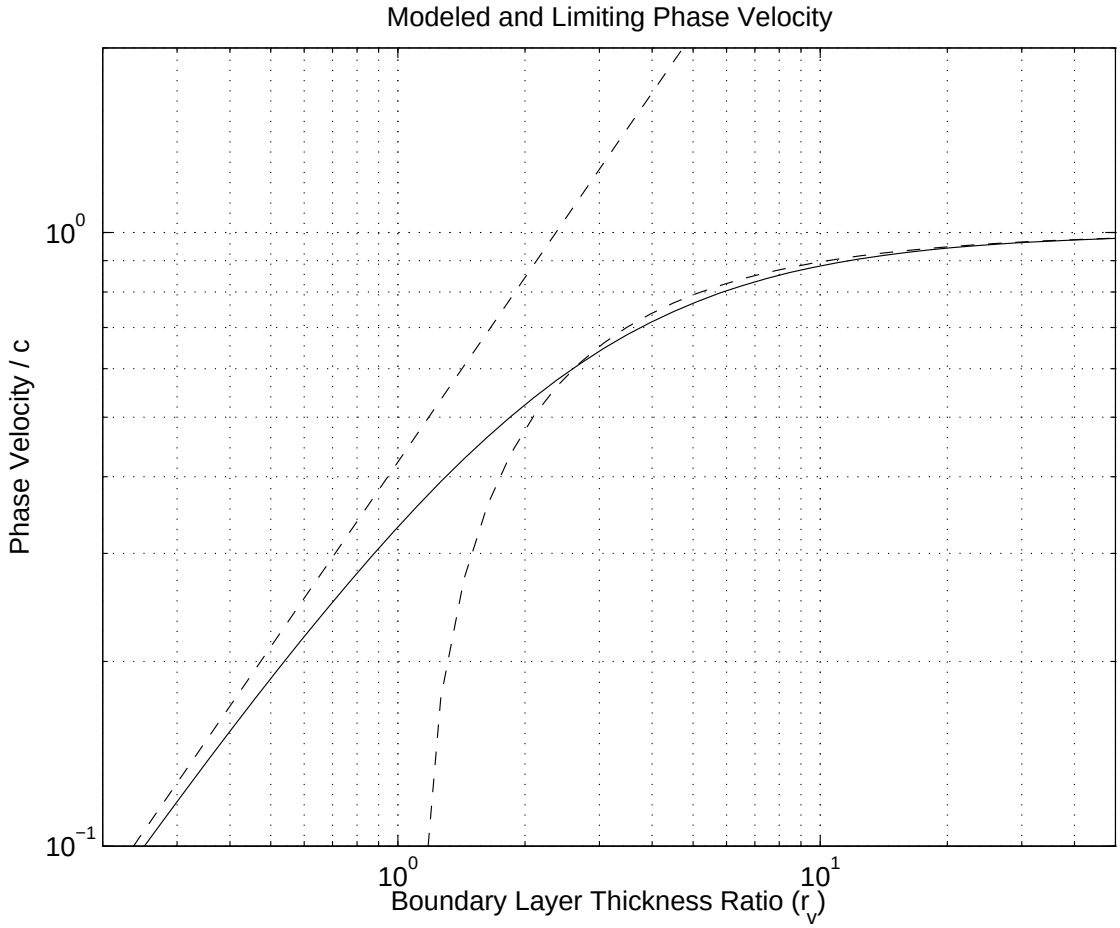


Figure 4.16: Modeled uniform phase velocity (solid) and the limiting behavior (dashed).

The asymptotic behavior of phase velocity and attenuation for large and small r_v is given by (4.60) and (4.61). The following single expression is a uniform approximation for the phase velocity,

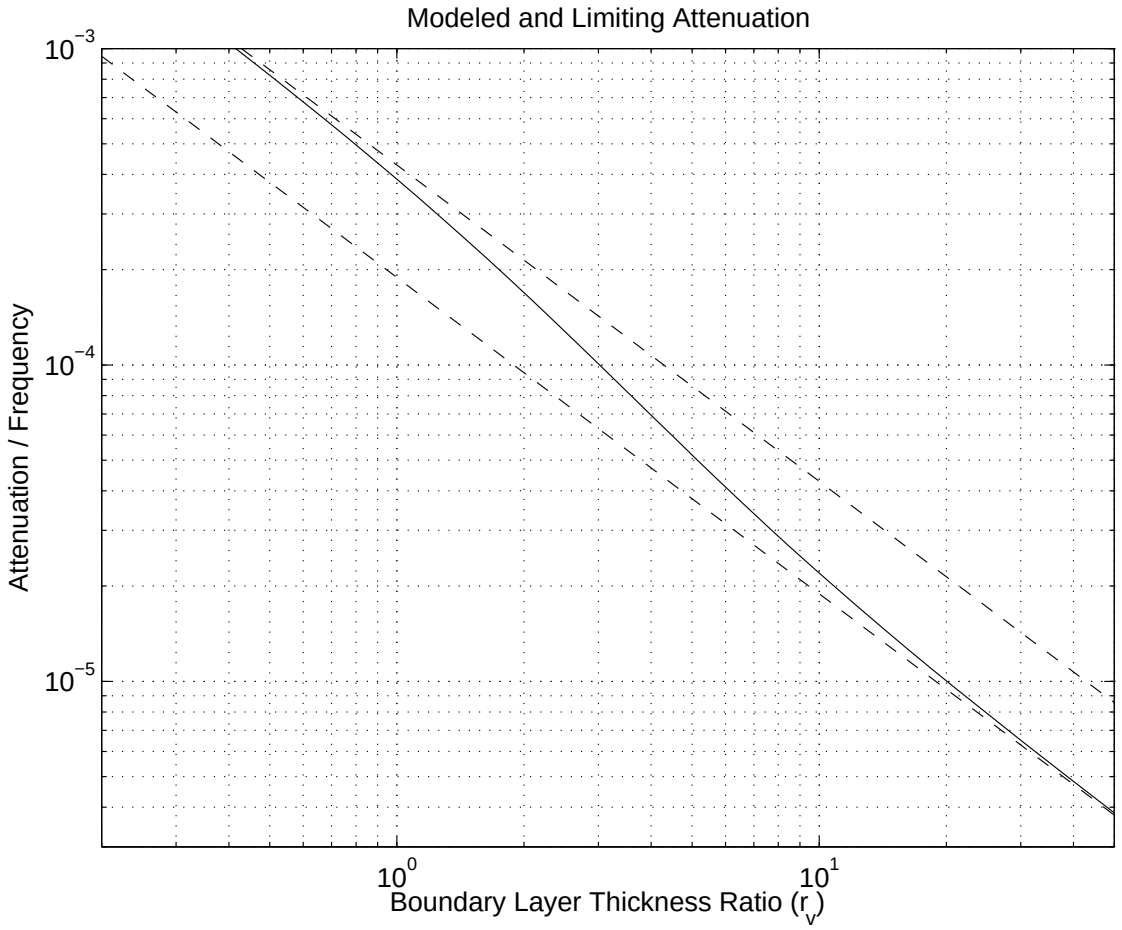


Figure 4.17: Modeled uniform attenuation (solid) and the limiting behavior (dashed).

valid for any r_v:

$$v(\omega) = c\,\frac{A_v r_v (1 + r_v/k)}{1 + A_v r_v (B_v + r_v/k)}, \tag{4.63}$$

where the parameters A_v and B_v, given by

$$A_v = \frac{1}{2\sqrt{\gamma}}, \qquad B_v = 1 + \frac{[1 + (\gamma - 1)/\nu)]}{\sqrt{2}k}, \tag{4.64}$$

were selected to match the limiting phase velocity behavior and the parameter

$$k = 2\gamma \tag{4.65}$$

marks the ratio r_v delimiting the small and large tube regions.

Similarly a uniform approximation for the attenuation coefficient is

$$\alpha(\omega) = \left(\frac{\omega}{cr_v}\right) \frac{A_\alpha + B_\alpha\,(r_v/k)}{1 + r_v/k}, \tag{4.66}$$

where the parameters A_α and B_α were again selected to match the limiting attenuation coefficient behavior and are given by

$$A_\alpha = 2\sqrt{\gamma}, \qquad B_\alpha = \frac{[1 + (\gamma - 1)/\nu]}{\sqrt{2}}. \tag{4.67}$$

Figures 4.16 and 4.17 show the uniform approximations for phase velocity and attenuation along with their limiting behaviors.

4.5.3 Wall Loss Filter

Filter Characteristic

The propagation constant $\Gamma(\omega)$ in (4.59) gives the per unit length attenuation of waves propagating along an infinitely long tube acting as a transmission line. Accordingly, the following frequency response approximates the attenuation and phase delay over a tube of length L:

$$e^{-\Gamma(\omega)L} = e^{-\alpha(\omega)L - j(\omega/v(\omega))L}. \tag{4.68}$$

Removing a pure delay of duration L/c, the desired wall loss filter frequency response, $\lambda(\omega)$, is obtained and given by

$$\lambda(\omega) = e^{-\alpha(\omega)L - j(\omega/v(\omega) - \omega/c)L}. \tag{4.69}$$

Noting that the attenuation $\alpha(\omega)$ increases with the square root of frequency, the wall loss filter $\lambda(\omega)$ is seen to have a gentle low-pass characteristic. This low-pass characteristic is more pronounced with decreasing tube radius a or increasing tube length L, as illustrated in Figures 4.18 and 4.19 respectively.

While not yet confirmed analytically, numerical results indicate that the wall loss filter $\lambda(\omega)$ is minimum phase. In the sequel, it is assumed this property holds, and consider only those design methods which produce minimum phase filters.

Filter Design

Given the desired wall loss filter transfer function $\lambda(\omega)$ expressed in terms of the tube parameters and physical constants, what remains is to compute the coefficients of a digital filter approximating the desired transfer function for frequencies up to the Nyquist limit. The wall loss magnitude is a relatively smooth function of frequency, and computationally efficient low-order rational filters are expected to provide a good fit.

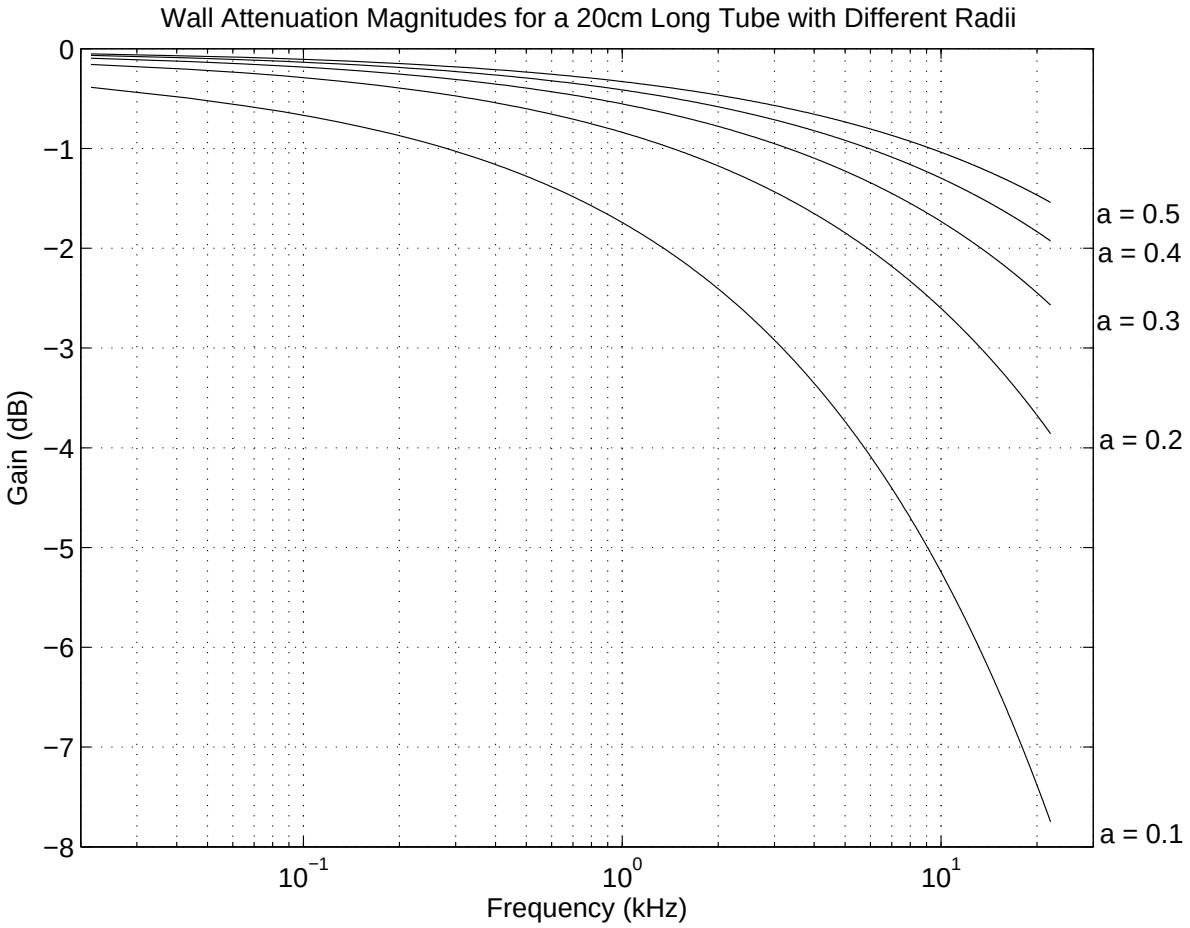


Figure 4.18: Wall loss filter magnitude $|\lambda(\omega)|$, showing a low-pass characteristic that is more pronounced with a decreasing tube radius (radius values, a, given in cm).

Warped Prony [75] and Hankel norm methods [73] were used to fit low-order rational filters to $\lambda(\omega)$. While the fit to the attenuation filter is good, as shown in the examples of Figures 4.20 and 4.21, these methods require considerable computation and are not convenient for a filter based on tube parameters likely to change in real-time. Additionally, in settings where tube parameters are time varying, the fit may produce real poles for some parameter values and complex poles for others, making it difficult to smoothly transition between wall loss filters.

To circumvent problems with standard modeling techniques, consider a cascade of minimum-phase first-order shelf filters $\sigma_i(z)$ [78],

$$\hat{\lambda}(z) = \prod_{i=1}^{N} \sigma_i(z; f_t(i), g_\pi(i)), \tag{4.70}$$

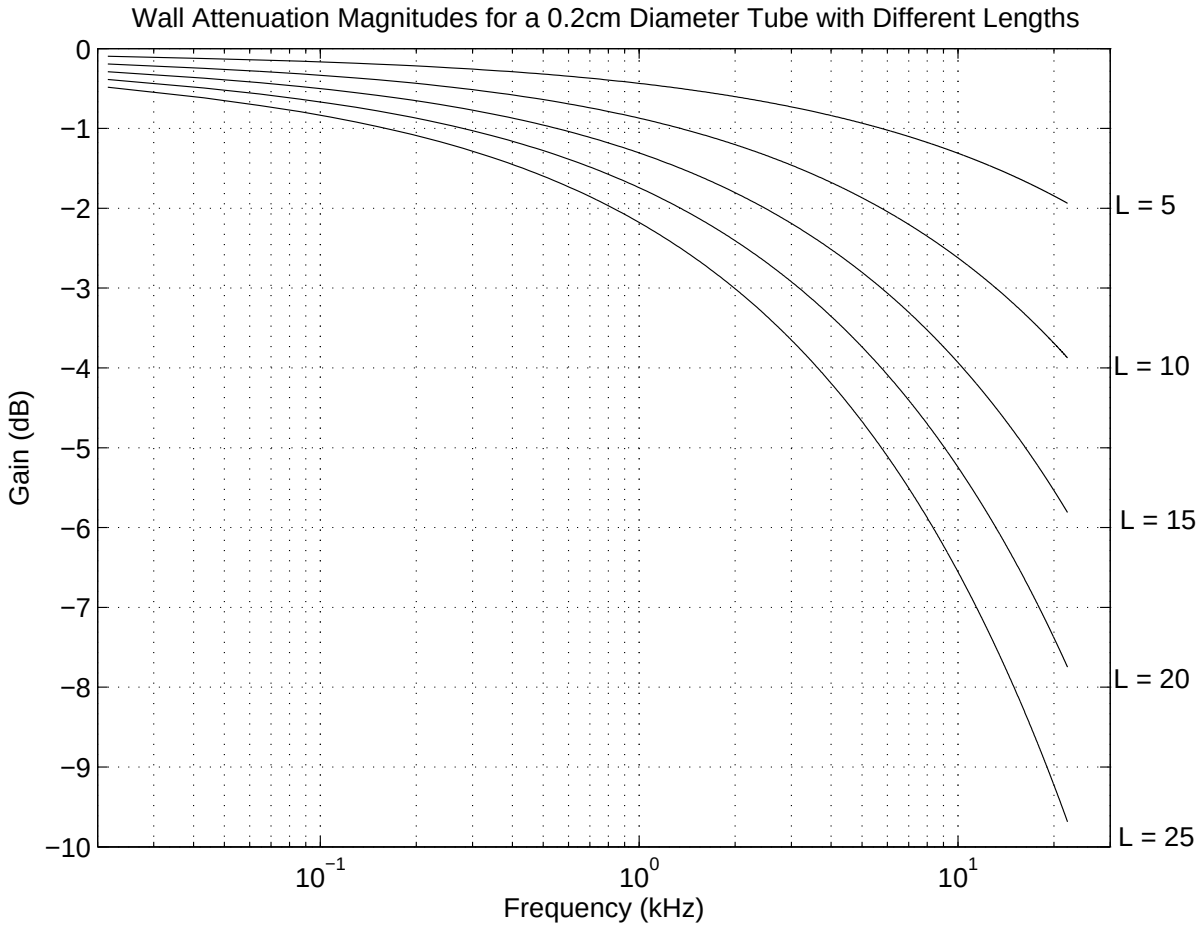


Figure 4.19: Wall loss filter magnitude $|\lambda(\omega)|$, showing a low-pass characteristic that is more pronounced with an increasing tube length (length values, L, given in cm).

with each shelf filter having a DC gain of one, a band edge gain $g_\pi(i)$, and a gain $\sqrt{g_\pi(i)}$ at its transition frequency $f_t(i)$. As shown in Appendix B, the shelf filter coefficients are easily computed in real-time. Also, since the shelf filters are first-order, they have real poles and zeros, and are relatively free of artifacts when made time varying.

First-order shelf filters have transfer functions which slowly transition between their DC and band edge values. It therefore seems reasonable that a cascade of shelf filters with judiciously chosen transition frequencies and gains could approximate the desired wall loss transfer function.

As seen in Figures 4.18 and 4.19, the desired wall loss filters as a function of tube length, radii, and sampling rate are strikingly self similar under the appropriate stretching of the log magnitude or frequency axes. Initially, the idea was to design via brute-force optimization a low-order shelf filter cascade which matched the desired transfer function at a particular sampling rate and tube

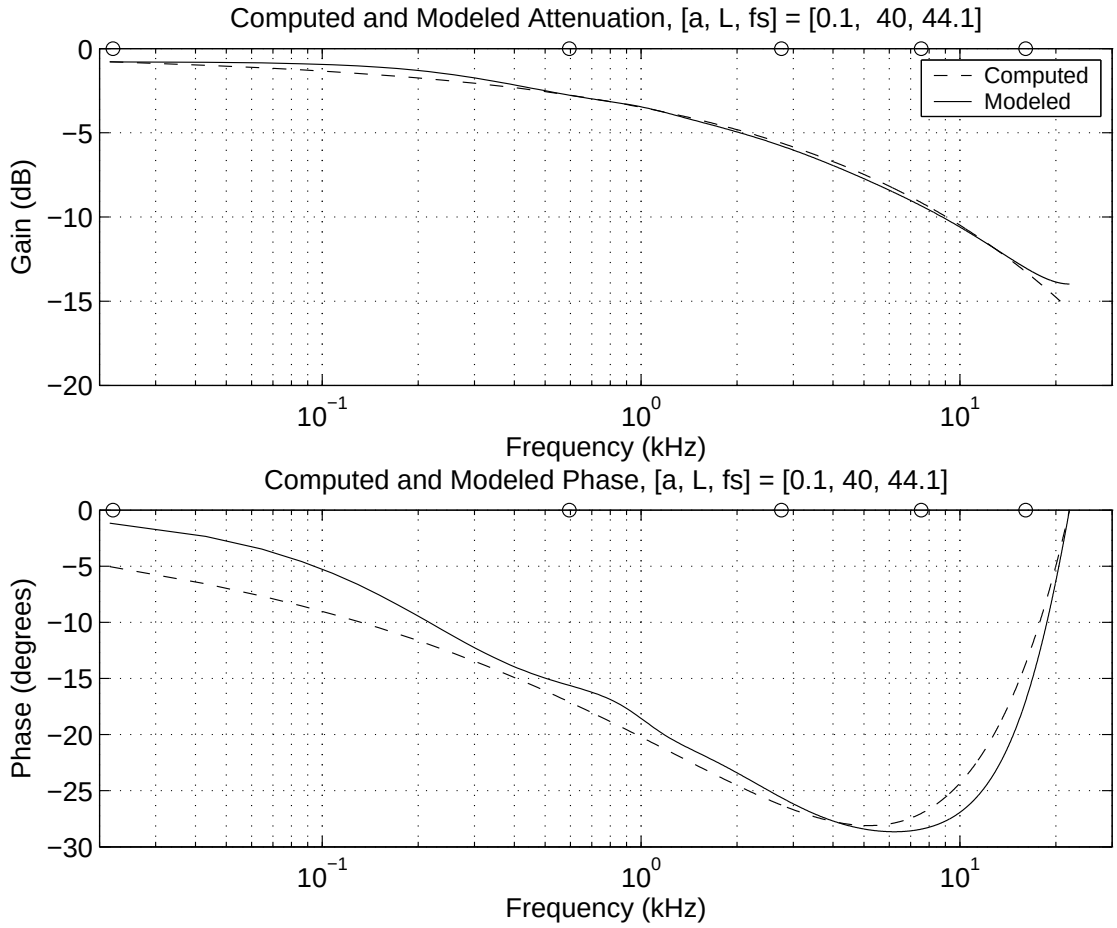


Figure 4.20: Warped Prony fifth-order model and computed wall loss filter transfer functions illustrating an excellent fit at all frequencies.

geometry. This prototype filter could then be stretched in response to sampling rate and conduit geometry.

In designing the prototype filters for different orders however, a pattern emerged and it was discovered that the cascade of shelf filters with band-edge gains (in units of amplitude) and transition frequencies (in radians/π) given by

$$\begin{aligned} g_\pi(i) &= \exp\left\{ \frac{\left[\left(i-\frac{1}{2}\right)/N\right]^{\frac{1}{2}}}{\sum_{k=1}^{N}\left[\left(k-\frac{1}{2}\right)/N\right]^{\frac{1}{2}}} \cdot \ln\left|\lambda(2\pi f_s/2)\right| \right\}, \\ f_t(i) &= \left[\left(i-\frac{1}{2}\right)/N\right]^3, \end{aligned} \tag{4.71}$$

has a transfer function which, as illustrated in Figures 4.23–4.25, is an excellent approximation to

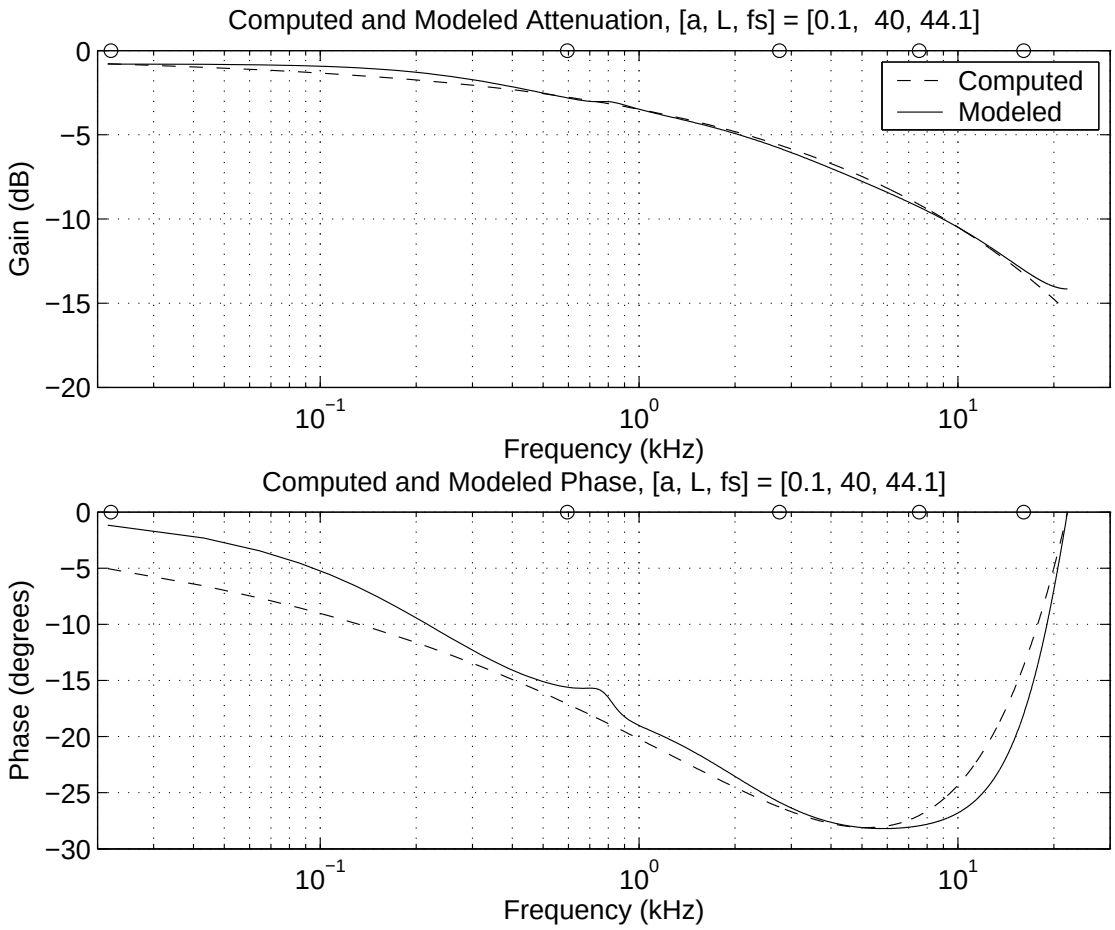


Figure 4.21: Hankel norm fifth-order model and computed wall loss filter transfer functions illustrating an excellent fit at all frequencies.

$\lambda(\omega)$ for a wide range of filter orders N, tube lengths L, tube radii a, and sampling rates f_s.

In the presence of a changing tube geometry, it is recommended that the filter $\hat{\lambda}(z)$ be implemented as a cascade of first-order sections, with coefficients linearly or exponentially interpolated between computed values. In this way, the dynamic range of the filter coefficients is minimized, and filters having interpolated coefficients are guaranteed to be stable with transfer functions roughly matching those associated with intermediate values of geometric parameters. Note that fifth-order filters are sufficient to give transfer functions accurate to withing a fraction of a dB across the audio band for typical audio sampling rates.

Wall attenuation filters are an important part of acoustic tube models (used by a large body of musical instruments) and should not be neglected for the sake of simplicity. Though their audibility may be slight for certain tube geometries, they can contribute significantly to the reduction of high

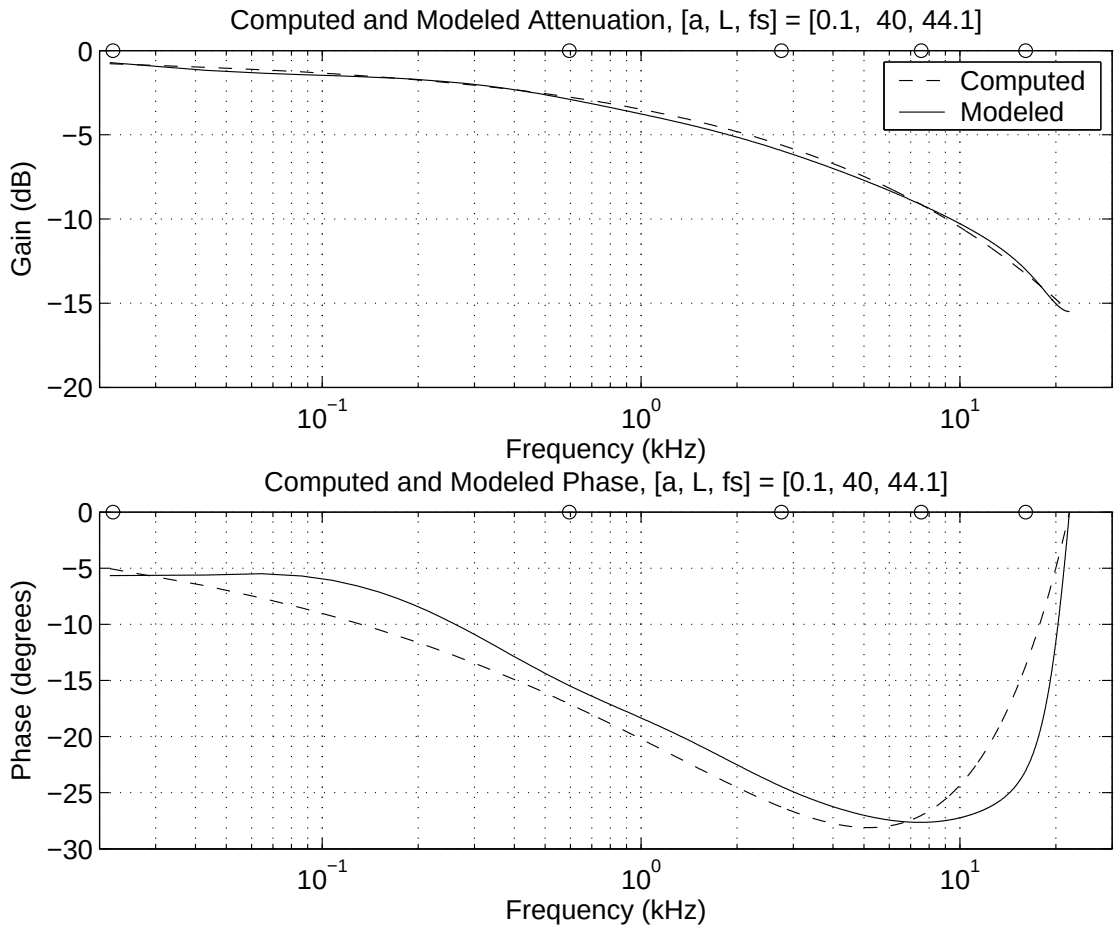


Figure 4.22: Computed and modeled wall loss filter illustrating an excellent fit at all frequencies using a cascade of first-order filters to produce a fifth-order IIR filter.

frequency components that lead to aliasing and can also eliminate some sources of instability in the model. Though approximations will often be satisfactory, it is worthwhile to aim for a more robust, scientifically sound and versatile solution.

As exemplified by Figures 4.23–4.25, a cascade of first-order shelf filters provides an excellent approximation to the cylindrical tube wall loss transfer function $\lambda(\omega)$ for a wide range of filter orders, tube lengths, tube radii, and sampling rates. It is also efficient to compute for continually changing parameters, making it an ideal choice for modeling wall losses in acoustic tubes.

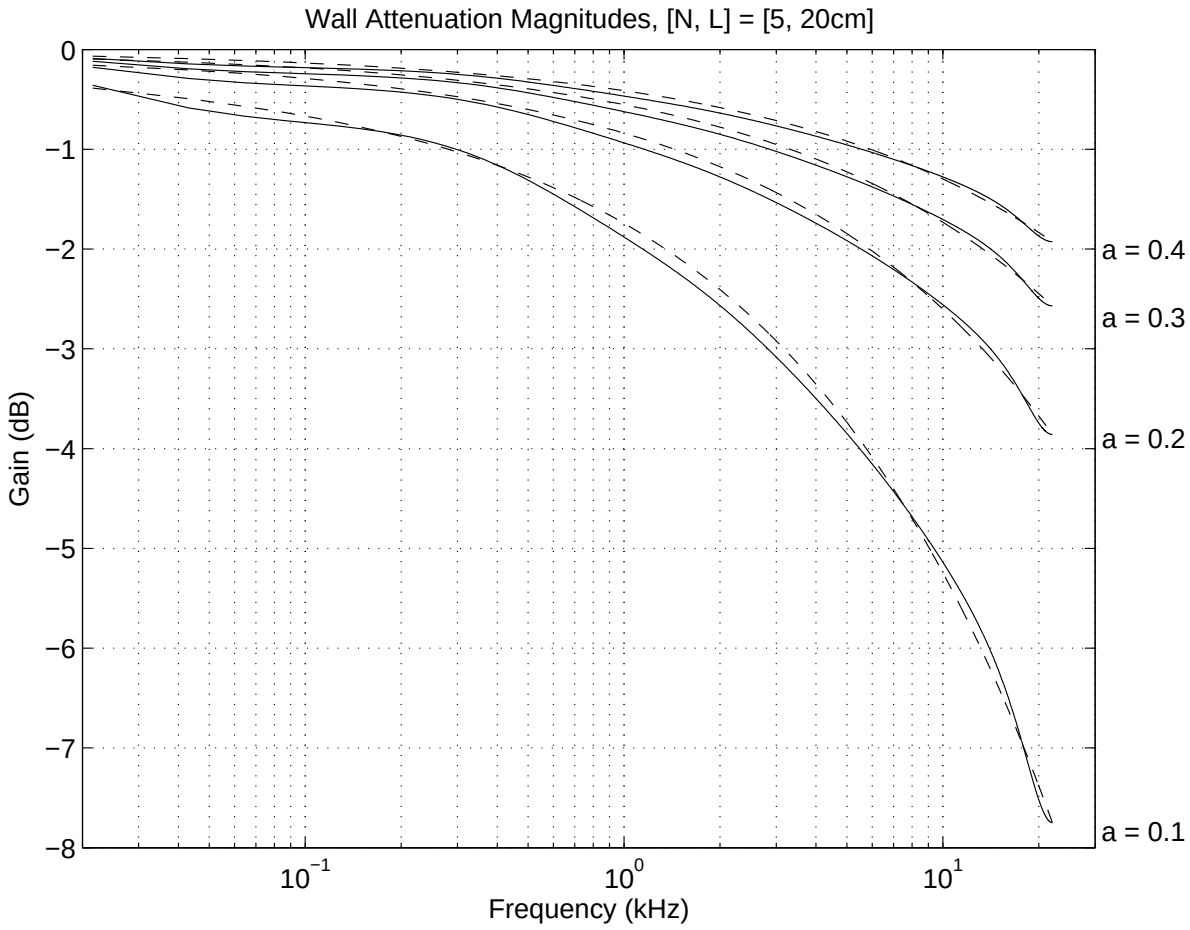


Figure 4.23: Computed and shelf filter cascade wall loss filter magnitude at various tube radii (values given in cm) with the filter order, N, and the tube length, L, constant.

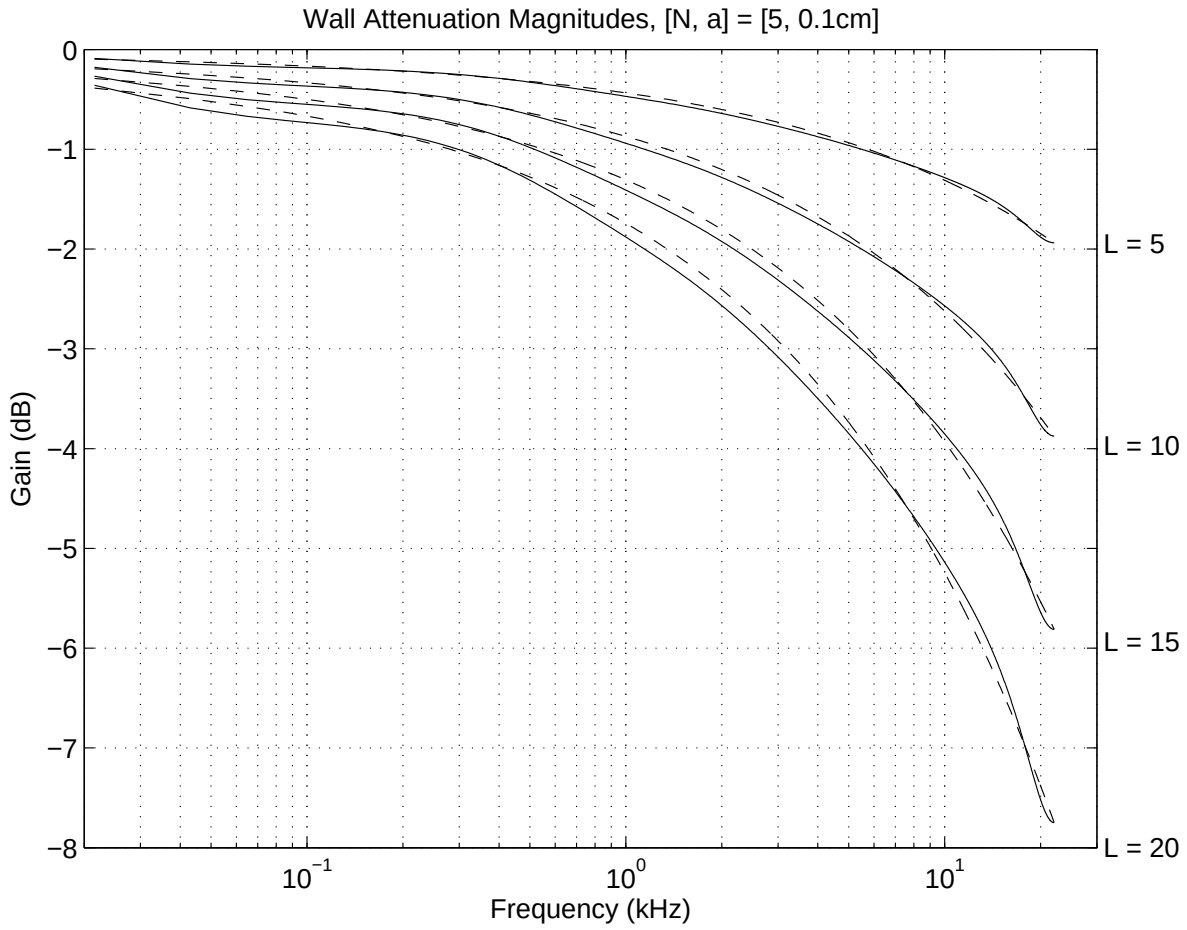


Figure 4.24: Computed and shelf filter cascade wall loss filter magnitude at various tube lengths (values given in cm) with the filter order, N, and the tube radius, a, constant.

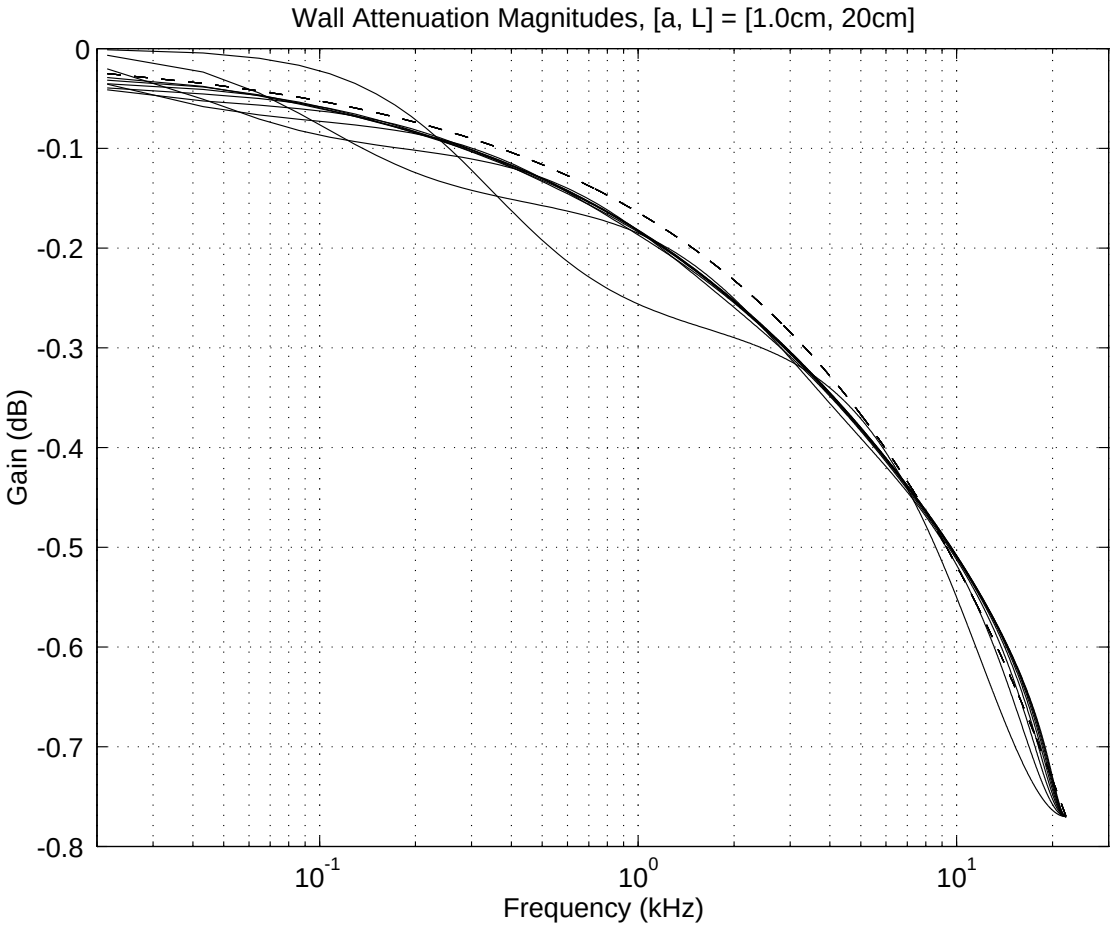


Figure 4.25: Computed and shelf filter cascade wall loss filter magnitude using $2, 3, \ldots, 10$ first-order shelf filters.

4.6 The Model

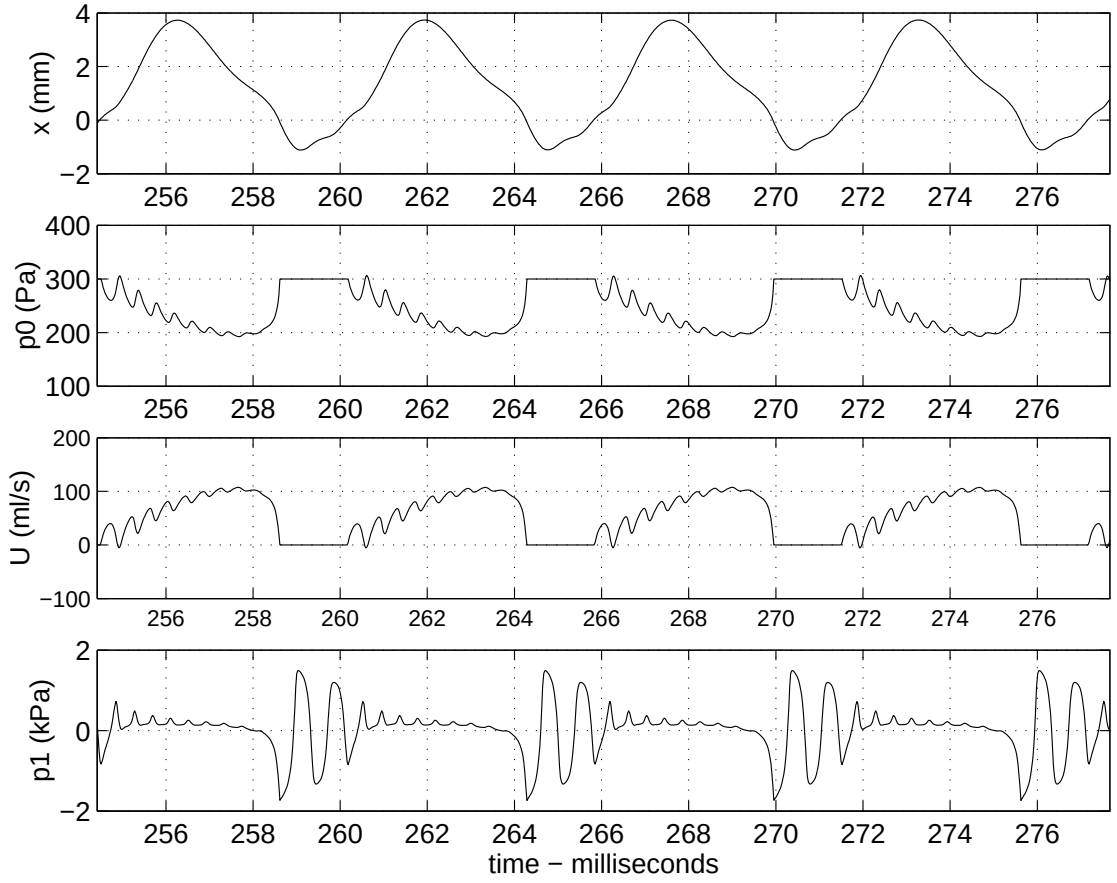


Figure 4.26: Waveforms are comparable to those published by Fletcher, but use waveguides with lumped losses and a 2nd-order-error algorithm for numerical integration to achieve stable discretization at audio sampling rates.

With all the model elements in place, Figure 4.26 shows the time evolution of the four variables in response to an example applied pressure p_G. When the valve is open, the pressure difference across the valve channel and the volume flow through the valve drive the volume flow derivative, and in turn influence the valve position and the pressure on either side of the valve. When the valve is closed, there is no volume flow through the valve, and the pressure on either side of the valve is free to evolve independently. These two behaviours are expressed every cycle as the valve opens and closes in response to an input pressure and the energy reflected from the trachea opening.

The pressure entering the valve displays a pulsation that clearly coincides with the frequency of the oscillating membrane. When the membrane closes the constriction completely (when $x = 0$) the amplitude of p_1 is high and its frequency is close to the resonance of a tube closed at one end.

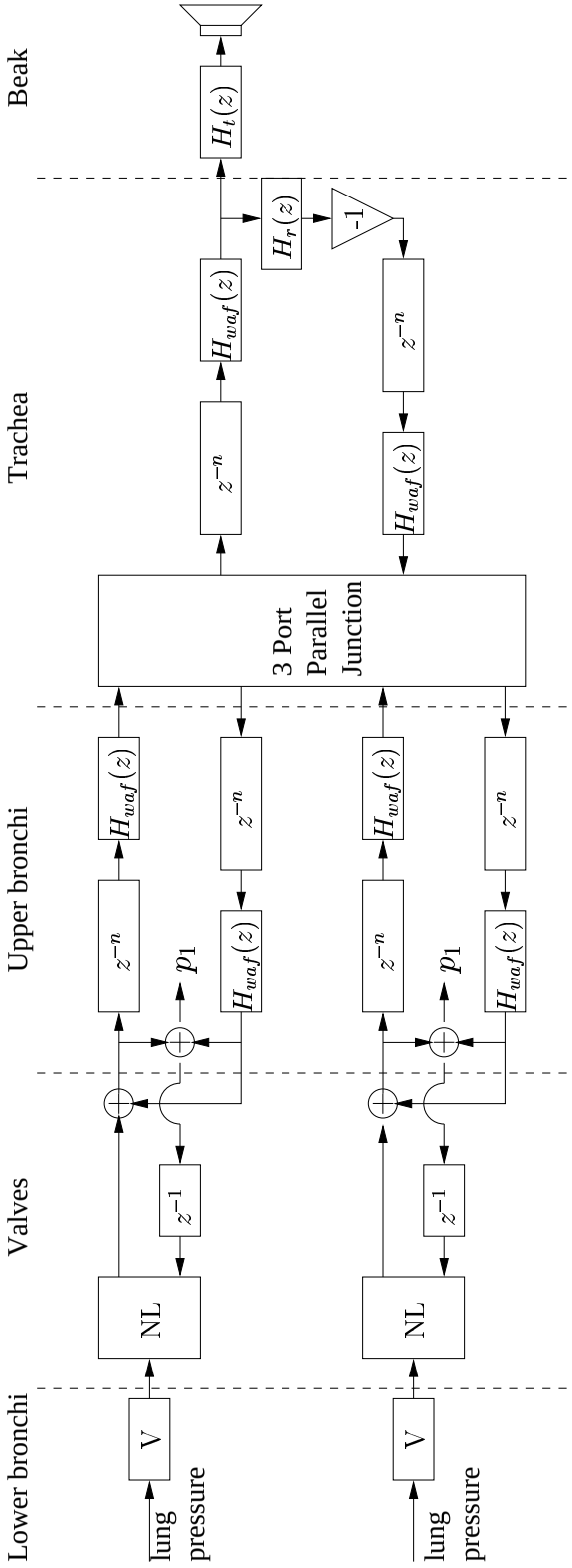


Figure 4.27: The syrinx model.

When the membrane is open, the pressure amplitude decreases and the frequency increases to the resonance of a tube open at both ends [46]. This provides satisfactory evidence that the model is behaving as it should, though in the following Chapter (Chapter 5) an improvement to volume flow (U) is discussed.

Figure 4.27 shows a signal flow diagram of all the model elements discussed in this chapter. Air pressure from the lungs enters the lower bronchi (modeled as having a volume and acoustic stiffness), the nonlinear vibrating valve (modeled using finite difference methods), the upper bronchi (modeled using a digital waveguide), the three port junction and finally the trachea (also modeled as a digital waveguide). Commuted at the waveguide endpoints are digital filters modeling wall losses and reflection/transmission filter in the trachea.

4.7 Conclusion

The output of the model is the voiced sound produced by the vibrating membranes in the syrinx. It is very rich in harmonics and is not the flute-like tone often associated with the songbird. The aim of this research is to create a computer musical instrument—one that is based on a physical model with physical parameters that can be manipulated and a sound that changes accordingly. Though it is desirable to keep the model as close as possible to scientific reality, it may be useful to extend it to ensure the production of the songbird's characteristic pure-tone.

Chapter 5

Feathering Collisions in Valve Simulation

As discussed in the previous chapter, pressure controlled valves exist in mechanically driven musical instruments such as brasses and woodwinds and are also found in many biological sound producing mechanisms such as the vocal folds in the human larynx and the vibrating membrane in the songbird's vocal organ, the syrinx.

Aliasing is a common problem in physical modeling synthesis and in particular in models of pressure controlled valves where the valve, usually a membrane or reed, has the ability to close completely. Depending on how this event is handled, an abrupt termination of air flow as a result of a closed valve can create undesirable discontinuities—among other artifacts—in the time evolution of model variables.

Relatively low audio sampling rates can create situations where, in the worst case, the model becomes completely unstable when discretized using an algorithm whose accuracy is dependent on the sampling period. In some instances this is remedied by moving to a higher-order-error algorithm such as the trapezoid rule for numeric integration as described in Chapter 4 and [13, 18, 19].

Much of the high frequency content circulating in a system involving a valve and an acoustic tube will be removed through the use of reflection and wall attenuation filters which, lowpass by nature, tend to reduce some of the effects of mild aliasing (as discussed in Chapter 4. In some cases however, a model may seem satisfactorily stable yet there will still be evidence of aliasing components in the output spectrum that can only be removed at the source. This chapter presents material previously published by Smyth, Abel and Smith in [21] that focuses on reducing the source of aliasing in a pressure-controlled valve simulation by "feathering" the collision between an open and closed valve. An improved solution to the time evolution of volume flow given in Chapter 4, provides a smooth, nearly alias free transition between these two states.

5.1 Air Flow Through a Pressure-Controlled Valve

In Chapter 4, four variables of the syrinx model (one example of a pressure controlled valve model) were given as

$p_0 \triangleq$ 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.

The time evolution of these variables is simulated by discretizing their corresponding differential equations according to the geometry of the avian syrinx and the transverse configuration of a pressure-controlled valve as defined in 4.2.1. The valve motion derivative is expressed in terms of the volume flow and pressures, the pressure derivative as a function of the flow and valve geometry, and the volume flow derivative as a function of the volume flow, pressures, and valve position.

Figure 4.26 shows the change in these variables over time in response to an applied pressure. There is a steep slope in the volume flow, U, which arises when setting the flow to zero in the case of a closed valve. When using relatively low audio sampling rates, this steep slope tends to cause aliasing in the output waveform. The derivative of volume flow is therefore re-examined in this chapter to create a more gradual *feathered* slope to zero flow when the valve closes. The volume flow is the focus of this chapter and is given by

$$\frac{dU}{dt} = \frac{2\sqrt{A(t)}}{\rho}(p_0 - p_1) - \frac{U^2}{4A(t)^{\frac{3}{2}}} \tag{5.1}$$

where $A(t) = ax(t)$ is half the cross-sectional area of the valve.

In simulating the valve, care must be taken in computing the volume flow behaviour between open and closed states. This is made more difficult by the singularity in (5.1) as the valve opening approaches zero. In the original model, this is handled by solving the equation only if the valve is open and A is greater than zero. If the valve is closed and A is zero or less, there should be no flow and (5.1) is set to zero.

Though the time evolution of the four variables in Section 4 (Figure 4.26) may seem well behaved, they are in fact heavily aliased. In Figure 5.7, an example of the model output spectrum is shown for a constant input pressure and an increasing syrinx tension. As the syrinx tension increases, the pitch increases. There is a corresponding decrease in the pitch of the aliased components, creating a clearly visible crosswork pattern.

5.2 Aliasing Caused by Switching in Discrete-Time Signals

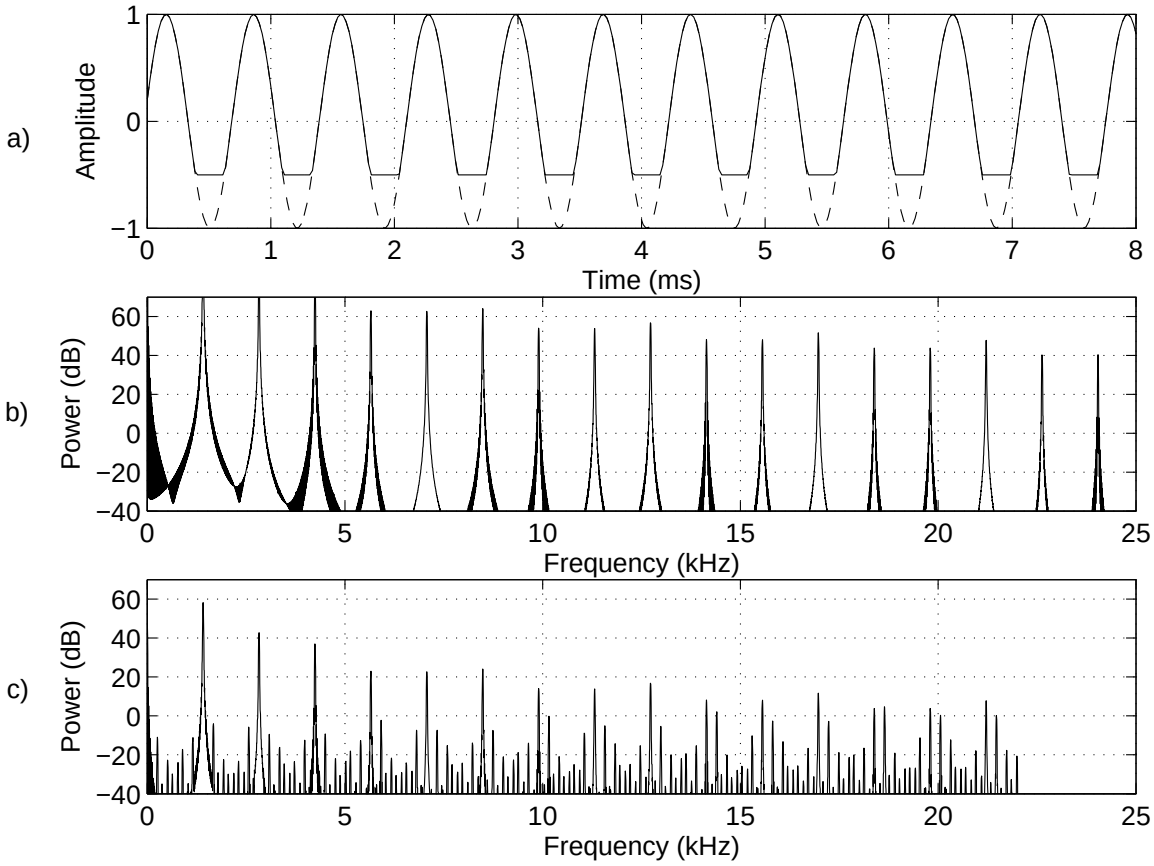


Figure 5.1: Figure (a) shows a full (dotted line) and truncated (solid line) version of a sine wave. Plot (b) shows the power spectrum of the truncated sinewave when a relatively high sampling rate of 300 kHz is used and plot (c) shows the artifacts introduced in the spectrum as a result of abruptly truncating the sinewave when a relatively low audio sampling rate of 44.1 kHz is used.

It is well known that aliasing results from switching based on a level threshold in discrete time signals. When airflow sets a valve into motion, it causes a change in the height of the valve channel and, if the motion is extreme, it can potentially close off the channel completely, creating a sudden termination in airflow. Examples of this occur when a reed beats against the lay of the mouthpiece or when the syrinx membrane touches the opposite wall of the upper bronchus. Such abrupt changes are difficult to synthesize because of the relatively low sampling rates afforded to audio. If a change occurs during a 0.023 ms interval (the time period between samples when sampled at 44100 Hz) the event may not be captured by the model at the precise moment and unwanted discontinuities and artifacts could result in the produced sound. Alternatively, if a change is calculated for a particular sample period, if the sample period is too long the change may no longer be valid by the time the

next sample is reached. Simulations that aren't concerned with producing audio do not have this problem since sampling rates can be made extremely high (and time intervals between samples very short) making the effects of time quantization negligible.

An example of this is seen in Figure 5.1 where a sine wave is abruptly truncated (a). This event (analogous to suddenly cutting the flow through a valve) produces only harmonic audible components in the spectrum when a very high sampling rate of 300kHz is used (b). However, at an audio rate of 44.1kH (c), the "jitter" in the truncation time results in aliased components which appear as peaks at nonharmonic frequencies.

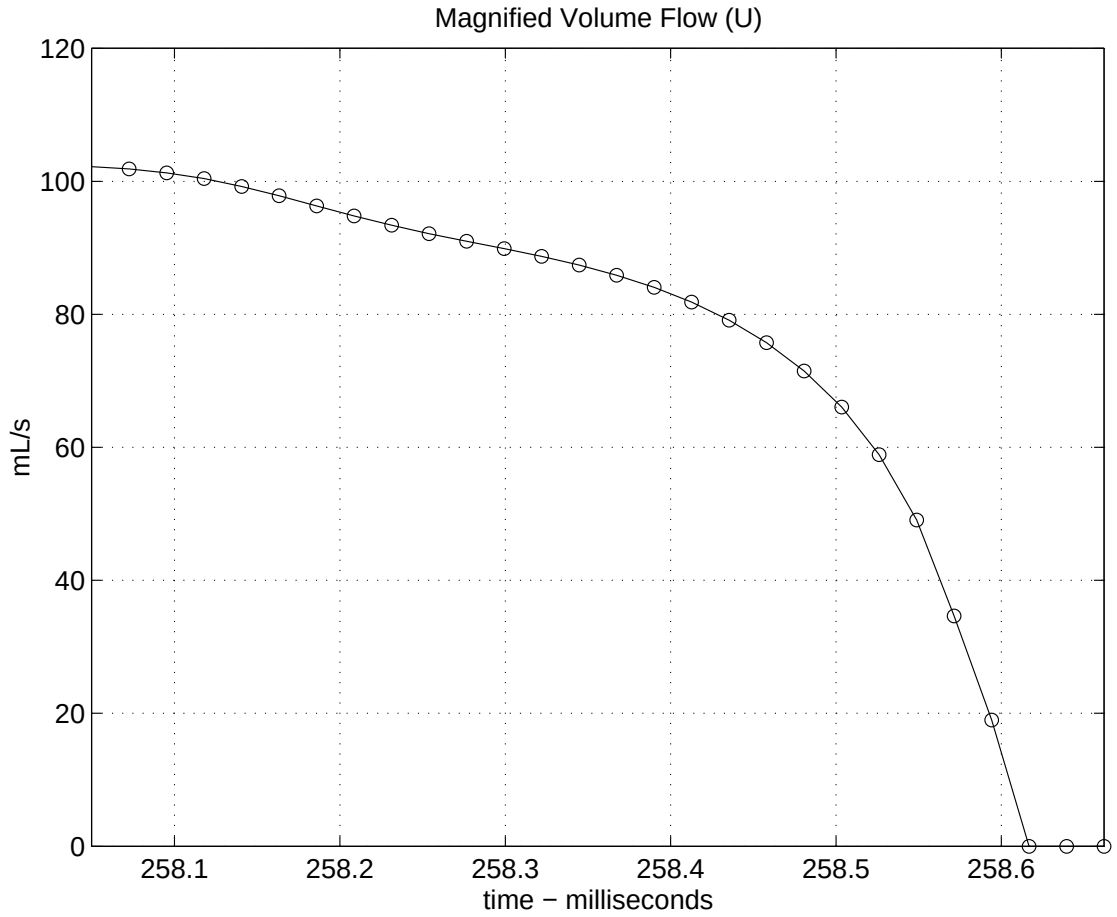


Figure 5.2: A magnified view of volume flow, U, taken from Figure 4.26. Only an excerpt from Figure 4.26 is shown to clearly illustrate the truncation on a sample boundary (in this case the sample following 258.6 ms).

Thus, as Figure 4.26 from Chapter 4 indicates, aliasing is caused by the truncation of the volume flow $U(t)$ to zero when the valve closes. As illustrated by the magnified plot of volume flow in Figure 5.2, $U(t)$ is forced to zero on a sample boundary. Regardless of the value of $U(t)$ predicted by

updating (5.1), air volume flow is still being set to zero when the valve is closed since no air should flow through a closed valve.

Truncating the volume flow on sample boundaries is therefore problematic. Depending on the period of the signal, the clipping may not happen at the same phase every period and aliased components will be generated.

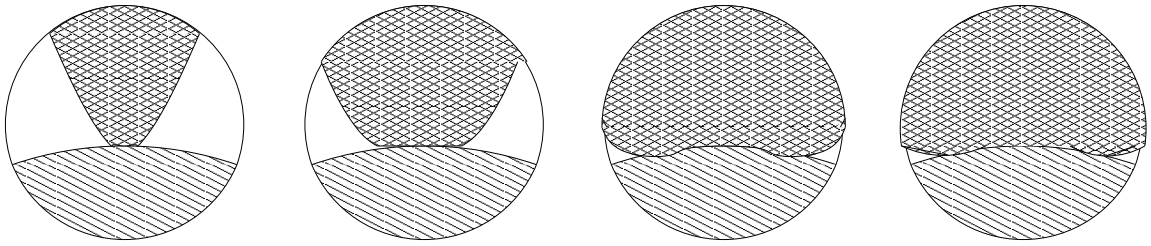


Figure 5.3: A hypothetical view of a flexible biological membrane beating on the cartilage of the opposite valve wall. As the valve closes, it likely starts with the center bulge and the remainder gently settles on either side before the channel is closed off completely.

If the vibrating membrane in the syrinx were to reach to opposite wall of the bronchus, the membrane's flexible biological material would likely cause it to touch gradually, starting with the center bulge and the remainder settling gently on either side before finally closing off the channel (as suggested by Figure 5.3). That is, instead of the channel being sealed the moment the membrane touches the opposite wall, it is more likely that flow will be able to seep through side corners and any other potential openings before the channel is closed completely.

In order to model this "leaky" characteristic of the valve the equation governing airflow through a pressure-controlled valve is re-examined. Since the aim is to have events occur less abruptly, it is more beneficial to look at how air flow changes over time rather than simply how it behaves at any one point in time. The differential equation for airflow is therefore adapted, and in particular for the case where the channel area is very small, and then the result is used to update the flow on sample boundaries.

5.3 Volume Flow Behaviour

Consider the case where the valve channel area $A(t) = ax(t)$ is sufficiently small that the first term of (5.1) can be ignored and the differential equation for $U(t)$ can be approximated by

$$\frac{dU}{dt} \approx -\frac{U^2}{4A(t)^{\frac{3}{2}}}, \qquad A(t) \ll 1, \tag{5.2}$$

which is in the form of a so-called *Bernoulli differential equation* [69]. Though this differential equation is nonlinear in $U(t)$, it may be converted to a linear form by the substitution

$$W(t) = \frac{1}{U(t)}. \tag{5.3}$$

Writing (5.2) in terms of $W(t)$ gives the following new differential equation for $U(t)$

$$\frac{dU}{dt} = -\frac{1}{W(t)^2}\frac{dW}{dt}, \tag{5.4}$$

where

$$\frac{dW}{dt} = \frac{1}{4A(t)^{3/2}}. \tag{5.5}$$

This equation is easily integrated to solve for volume flow:

$$W(t) = \int^t \frac{dt}{4(A(t))^{3/2}} + C, \tag{5.6}$$

$$U(t) = \frac{1}{\displaystyle\int^t \frac{dt}{4A(t)^{3/2}} + C}, \tag{5.7}$$

where the constant of integration C may be set given knowledge of $U(t)$ at a particular time t_0. Solving for C,

$$U(t) = \frac{U(t_0)}{1 + U(t_0)\displaystyle\int_{t_0}^t \frac{dt}{4A(t)^{3/2}}}. \tag{5.8}$$

Note that when the area $A(t)$ is small, the integral in the denominator of (5.8) is large, and any initial positive value of volume flow is quickly reduced to zero without crossing zero, as would be expected for a closing valve. This observation provides justification for having zero volume flow when the valve area is zero. The small valve area solution to (5.8) suggests a possible alternative to truncating U when the valve is closed (the original solution to the singularity in (5.1) that resulted in aliasing). If the valve were slightly leaky, e.g.,

$$\tilde{A}(t) = A(t) + \lambda, \tag{5.9}$$

for a small leakage area λ, the singularity at zero area would be avoided, and the volume flow behaviour would be relatively unchanged. However, it is not sufficient to use a leaky valve in place of one that is truncated because though this may reduce the slope of $U(t)$ it also introduces the undesirable behaviour of volume flow oscillating about zero.

5.4 Corrected Volume Flow Update

The difficulty with discretizing (5.1) in the presence of small valve areas is illustrated in Figure 5.4. Since, as Figure 5.2 indicates, the slope of $U(t)$ is decreasing with decreasing volume flow, predictions of the slope based on (5.1) tend to overshoot zero volume flow. It would therefore be preferable to use the value of $dU(t)/dt$ as predicted by the small area solution (5.8) to update the volume flow when $A(t)$ is small.

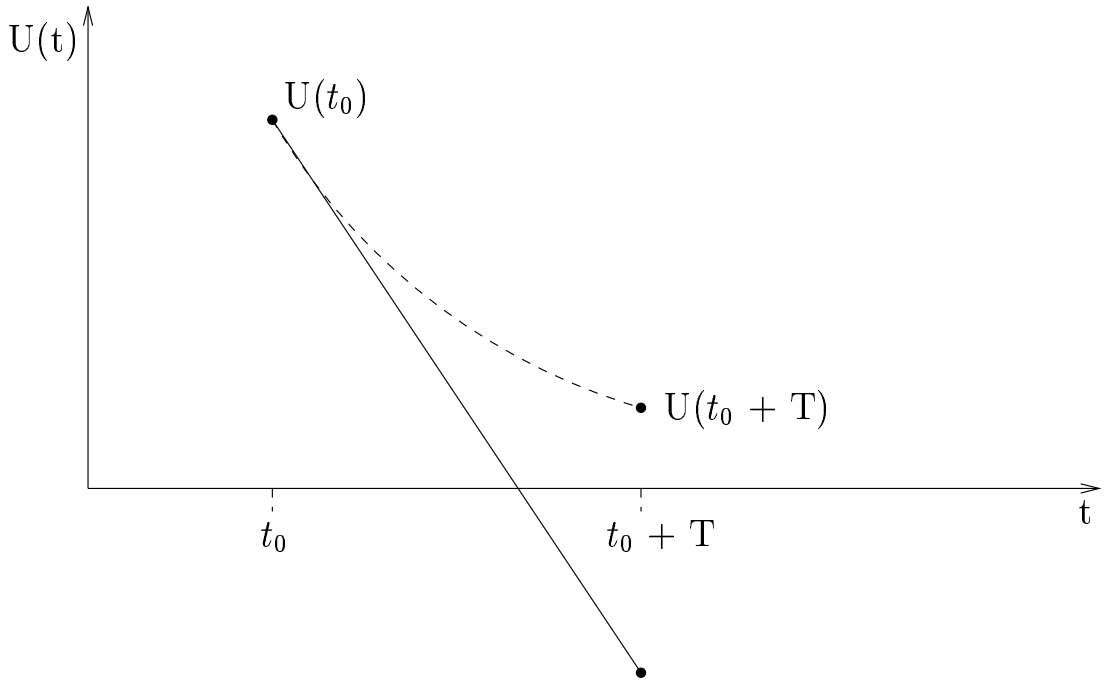


Figure 5.4: In the case of a large sampling period T, updating the volume flow using (5.1) can cause U to overshoot. The dotted line represents the actual value of U.

In order to see how this solution should be incorporated into the volume flow update, consider the value of $U(t)$ at time $t_0 + T$, T being the sampling period. Starting with the small area solution for volume flow (5.8), but in a more convenient form

$$U(t) = \left[\frac{1}{U(t_0)} + \int_{t_0}^{t} \frac{1}{4A(t)^{3/2}} dt\right]^{-1}, \tag{5.10}$$

the valve channel area is assumed to be constant at $A(t_0)$ during the time interval $[t_0, t_0 + T]$. Substituting into (5.10) gives

$$U(t) = \left[\frac{1}{U(t_0)} + \frac{1}{4A(t_0)^{3/2}}(t - t_0)\right]^{-1}, \tag{5.11}$$

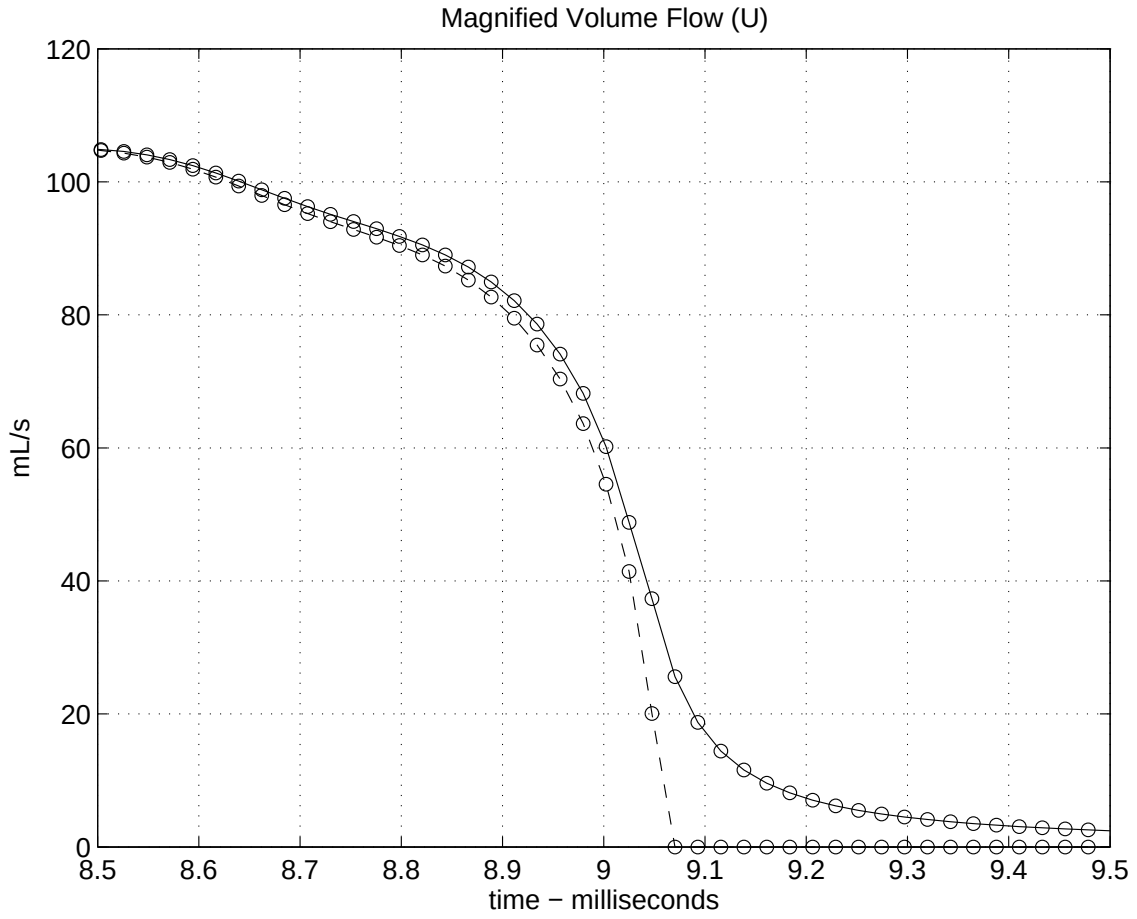


Figure 5.5: The prior model with truncated volume flow (dotted line) vs. the new flow with the "leaky" term added (solid line).

and the volume flow at $t_0 + T$ is

$$U(t_0 + T) = \left[\frac{1}{U(t_0)} + \frac{T}{4A(t_0)^{3/2}}\right]^{-1} \tag{5.12}$$

$$= U(t_0)\left[1 + \frac{U(t_0)}{4A(t_0)^{3/2}}T\right]^{-1}. \tag{5.13}$$

Using the first-order backwards difference approximation, the new differential equation for $U(t)$

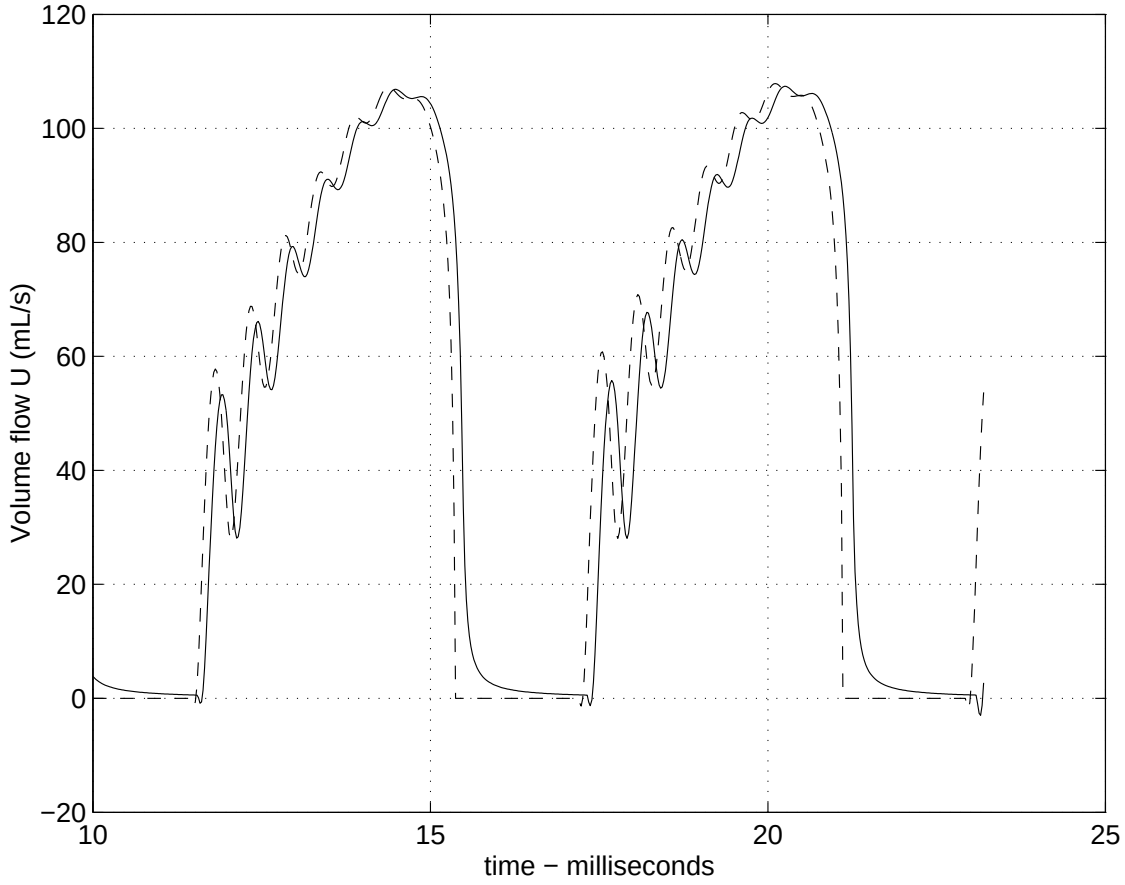


Figure 5.6: Volume flow truncated and with "leaky" term added.

becomes

$$\frac{dU}{dt} = \frac{U(t_0+T)-U(t_0)}{T} \tag{5.14}$$

$$= -\frac{U(t_0)^2}{4A(t_0)^{3/2}} \cdot \left[1+\frac{U(t_0)}{4A(t_0)^{3/2}}T\right]^{-1}. \tag{5.15}$$

Comparing the form of (5.15) to (5.1) note that the Bernoulli terms are identical, save a factor of

$$\left[1+U(t_0)/4A(t_0)^{3/2}T\right]^{-1}.$$

This factor has the effect of reducing the derivative in the presence of small channel areas or large

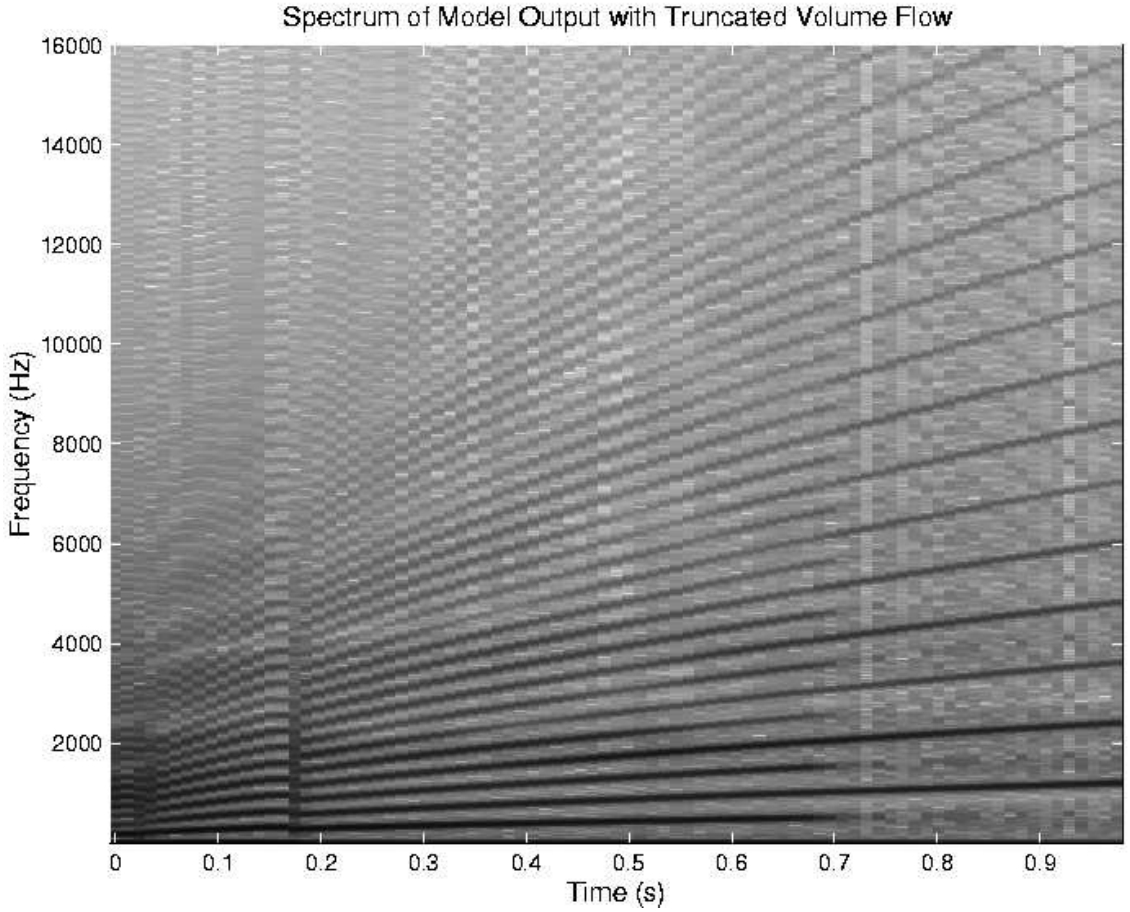


Figure 5.7: To illustrate the aliasing (indicated by the crosshatch pattern) that occurs at higher frequencies in the model output spectrum with a truncated volume flow, the pressure control parameter is held constant while the tension control parameter is ramped to increase the pitch.

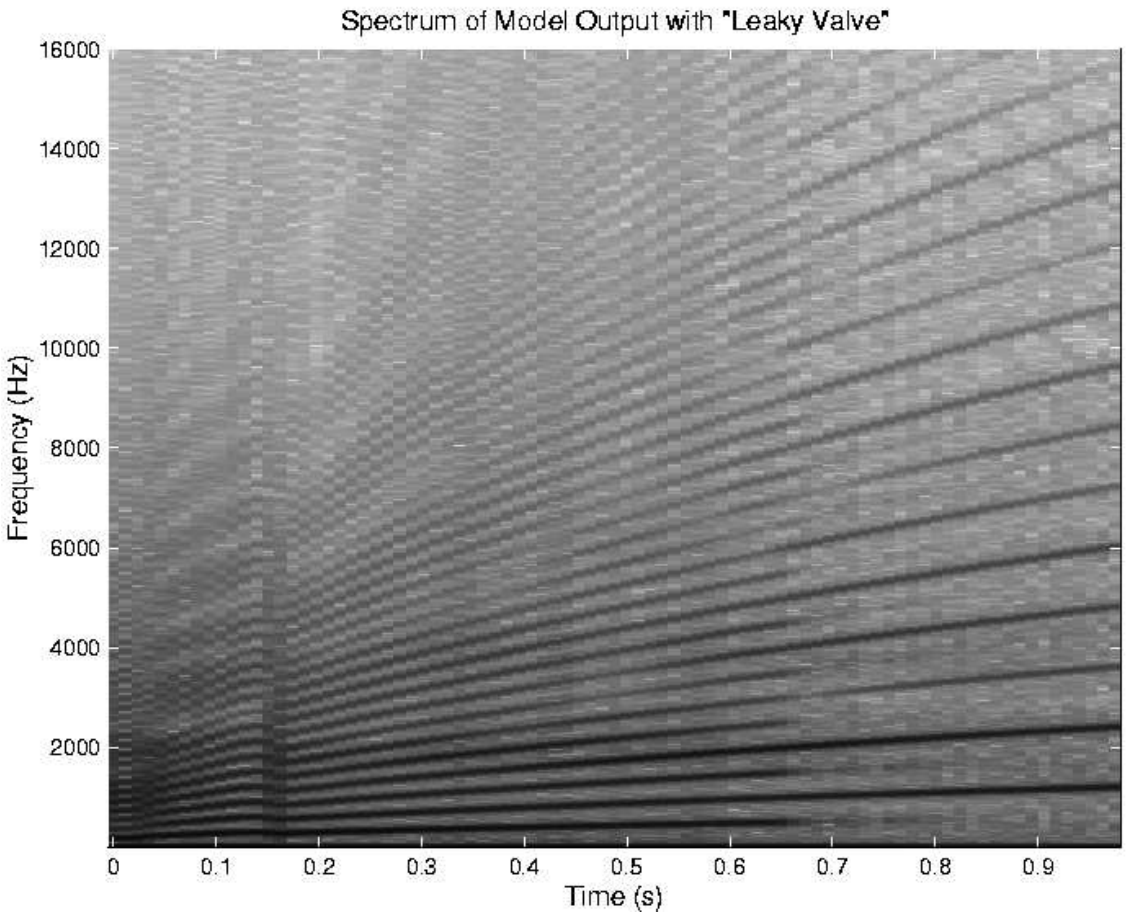


Figure 5.8: The model output spectrum with the "feathered" valve collision, shows a significant reduction in aliasing components even when using the same control parameters as Figure 5.7.

sample periods. Rewriting (5.15) gives

$$\frac{dU}{dt} = -\frac{U(t_0)^2}{4A(t_0)^{3/2} + U(t_0)T}. \tag{5.16}$$

Note that in this form the Bernoulli term is similar to that of (5.1), with a valve having increased area. In other words, it has become a leaky valve whose leakage increases with increasing volume flow.

In addition to creating the desired effect of a gentler slope, this new form for the Bernoulli term (5.16) solves the numerical instability in (5.1) by adding a non-zero term to the denominator, allowing $A(t)$ to take on a zero value. It is no longer necessary to manually halt the air flow the moment the valve is closed. Rather, as can be seen in Figure 5.6, the air flow is now being brought to zero along a more accurate and smooth trajectory, and, as illustrated in Figure 5.8, the presence of aliasing is greatly reduced when compared to Figure 5.7.

The final feathered update for volume flow when discretized using the trapezoid rule for numerical integration [13] is given by

$$\left.\frac{dU}{dt}\right|_{t_0+T} = \left.\frac{dU}{dt}\right|_{t_0} \frac{T}{2} + \left(\frac{2\sqrt{A(t_0)}}{\rho}(p_0 - p_1) - \frac{U(t_0)^2}{4A(t_0)^{\frac{3}{2}} + U(t_0)T}\right)\frac{T}{2}. \tag{5.17}$$

5.5 Applying Feathered Flow Model to a Clarinet Reed

In the bird's vocal organ, the syrinx, air pressure from the lungs controls the oscillation of a membrane (by changing the pressure across the membrane), its displacement causing a variable constriction through which air flows before reaching the upper bronchus and trachea. Similarly, blowing into the mouthpiece of a clarinet will cause the reed to vibrate, narrowing and widening the airflow aperture to the bore.

Digital simulations of many woodwind and brass musical instruments (instruments that use a pressure controlled valve) suffer from a problem similar to that of the original syrinx model when the reed beats against the lay of the mouthpiece or, more generally, when the valve has the ability to close completely. In this section, the syrinx valve improvements are applied to the clarinet reed.

Just as air from their lungs during birdsong causes motion of the syrinx valve and thus a change in the height of the valve channel, so does blowing into the mouthpiece of a clarinet cause the reed to vibrate, narrowing and widening the airflow aperture to the bore with the possibility of closing it completely. Though the cane reed is much more rigid than the bird's syrinx membrane, its simulation can benefit from the same principles of feathering—particularly when meeting the demands of relatively low audio sampling rates.

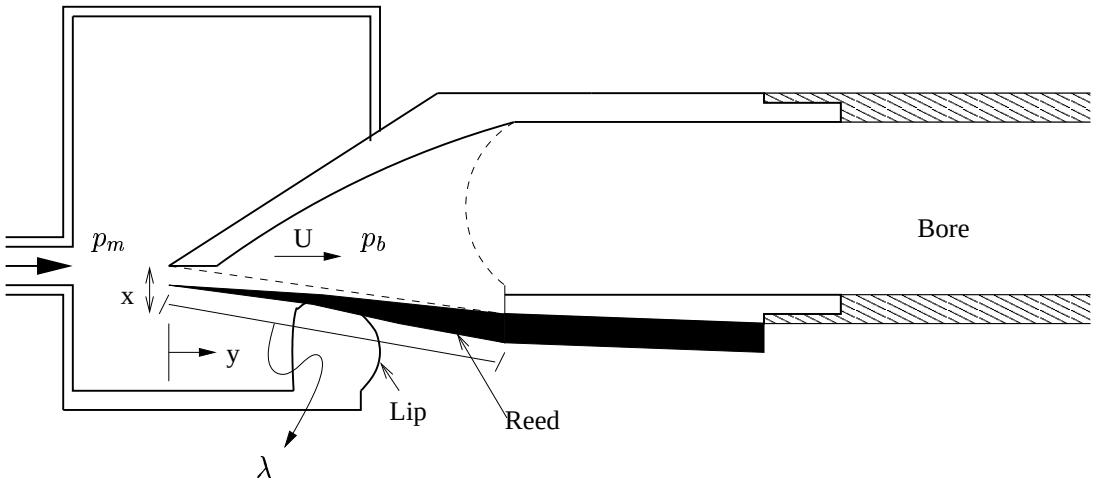


Figure 5.9: A simplified diagram of a clarinet reed. The variable p_m represents the mouth pressure, p_b is the pressure in the bore, U is the volume flow, x is the displacement of the reed, y indicates the position along the reed and λ is the length of the unclamped end of the reed.

5.5.1 The Quasi-Static Clarinet Model

Most current models of the clarinet reed are implemented using a lookup table which matches values for flow with the pressure drop across the reed valve. This is known as a quasi-static model since the value of flow, U, is established by using a lookup table (see Figure 5.10) rather than being determined by its behaviour over time as indicated by its differential equation. Though most developers are aware of the existence of a dynamic model, in practice the quasi-static is used more frequently since very satisfying results are obtained with relatively low computational cost. The results however, are not quite so impressive when the reed beats against the lay of the mouthpiece since the point of collisions between the reed and the lay is often too abrupt, causing the sound to be metallic and artificial.

An excellent description of the quasi-static clarinet reed model was published by Dalmont, Gilbert and Ollivier in [27] but is summarized here to give context to the discussion that follows. The steady flow through a valve is determined based on an input (or blowing) pressure pressure and a resulting output (or mouthpiece) pressure. The difference between these two pressured gives Δp.

Neglecting inertia, which introduces a delay in U and hence the Bernoulli pressure [31], the stationary equation for volume flow is given by

$$U = A\sqrt{\frac{2\Delta p}{\rho}}, \tag{5.18}$$

where Δp is the pressure difference across the reed, A is the cross section area of the air column (and the jet) and ρ is the density of air. This equation forms the table found in Figure 5.10 and is

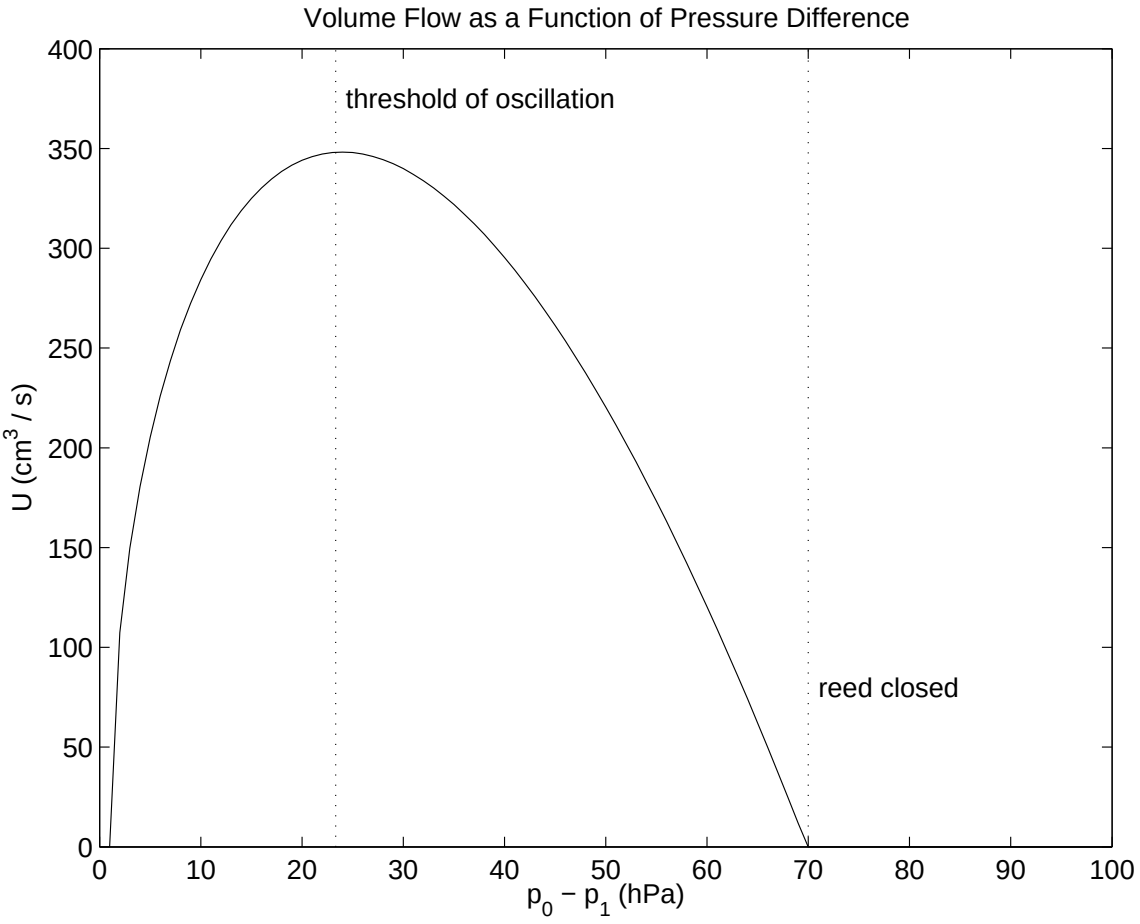


Figure 5.10: The region of oscillation is usually between the two dotted lines.

comparable to (5.1) with the derivative, dU/dt, set to zero.

The geometry of the clarinet valve is given by the width of the reed channel, w and the height of the opening (or alternatively, the distance between the reed and the lay), H. The area, A, is therefore given by

$$A = wH. \tag{5.19}$$

The motion of the reed, like that of the syrinx valve, follows the familiar equation

$$\frac{d^2x}{dt^2} + 2\gamma\frac{dx}{dt} + \omega^2(x - x_0) = \frac{A}{m}\Delta p, \tag{5.20}$$

where γ is a damping coefficient, m is the effective mass of the reed and ω is the reed's resonant

frequency. Since ω is related to the stiffness and the mass by

$$\omega = \sqrt{\frac{k}{m}}, \tag{5.21}$$

where κ is the reed stiffness (in Pa/m), equation (5.20) can be rewritten for convenience as

$$\mu \frac{d^2x}{dt^2} + \mu g \frac{dx}{dt} + \kappa x = \Delta p, \tag{5.22}$$

where μ is mass per m^2 and g is a viscous-damping coefficient (in s^{-1}) [27]. In the quasi-static model, the derivatives in (5.22) are set to zero rendering the mechanical reed effectively massless, with the stiffness being the only reactive element. The equation for the quasi-static reed therefore becomes

$$x = \frac{\Delta p}{\kappa}. \tag{5.23}$$

If H_0 is the equilibrium opening, that is, the opening of the valve in the absence of flow, the displacement of the reed determines the valve opening, H, by

$$H = H_0 - x. \tag{5.24}$$

From (5.23) and (5.24), the pressure difference corresponding to a closed reed is determined by setting the opening to zero or, alternatively, the displacement to its maximum, $x = H_0$. That is,

$$x = \frac{\Delta p}{\kappa}, \tag{5.25}$$

$$H_0 - H = \frac{\Delta p}{\kappa}, \tag{5.26}$$

$$\Delta p_{max} = \kappa H_0, \qquad \text{if } H = 0, \tag{5.27}$$

where p_{max} represents the maximum pressure difference above which the valve is closed.

By applying (5.24) and (5.25), the area of the reed opening from (5.19) becomes

$$A = w\left(H_0 - \frac{\Delta p}{k}\right) \tag{5.28}$$

which can be further reduced to

$$A = wH_0\left(1 - \frac{\Delta p}{\Delta p_{max}}\right) \tag{5.29}$$

by applying (5.27). The stationary volume flow from (5.18) can therefore be written as

$$U = wH_0\left(1 - \frac{\Delta p}{\Delta p_{max}}\right)\sqrt{\frac{2\Delta p}{\rho}}. \tag{5.30}$$

Quantity	Symbol	Value Range
Equilibrium reed opening	H_0	$0.04 - 0.1$ cm
Reed stiffness	κ	$800 - 1300$ hPa/cm
Effective width of jet	αw	$1.2 - 1.8$ cm
Maximum volume flow	U_{max}	$200 - 600$ cm^3/s (ml/s)
Flow on the reed	μ	0.9 cm
Damping coefficients	γ	1000/s
Density of air	ρ	0.0000012 kg/cm^3
Vibration frequency of reed (mode 1)	f_1	1045 Hz
Reed length	λ	3.4 cm

Table 5.1: Example value ranges for variables of the quasi-static clarinet model (some values are taken from [27]).

By substituting

$$\Delta p = \Delta p_{max}/3, \tag{5.31}$$

which is the value of Δp just below the threshold of oscillation [35], the maximum value for flow, U_{max}, is obtained using the equation

$$U = \frac{2}{3} w H_0 \sqrt{\frac{2\Delta p_m ax}{3\rho}}. \tag{5.32}$$

If the pressure difference is greater than p_{max}, it is assumed there is no flow through the valve channel (see Figure 5.10) and U is set to zero. This handling of the flow between open and closed valves is reminiscent of the first iteration of the syrinx model. It can also be improved by feathering the collisions between the open and closed states.

5.5.2 Improving the Model by Applying the Feathered Beating Reed

Before applying the *feathered* beating reed, the equations for flow and displacement in the quasi-static model must first be replaced with their corresponding differential equations, incorporating the appropriate valve geometry for the clarinet.

Volume Flow

The strategy for determining the volume flow derivative in the clarinet reed model is similar to that for the syrinx, discussed in Section 4.2.2, where Bernoulli's equation (4.3) is again employed.

The shape of the clarinet reed is quite different from the syrinx membrane and only a thin slice

dy along the y axis (see Figure 5.9) which is in contact with the flow (before the flow separates and forms a jet) need be considered. The force on this part of the reed is given by

$$F = A(y;x)\Delta p(y). \tag{5.33}$$

where $A(y;x)$ is the area of the valve channel at this position and $\Delta p(y)$ is the pressure drop across this section of the reed. This force is applied to a mass of

$$m = \rho A(y;x)dy, \tag{5.34}$$

where ρ is the air density and $A(y;x)dy$ is the volume to which the force is applied. Since pressure is proportional to force, Newton's second law, $F = ma$, can be applied to (5.33) to obtain

$$A(y;x)\Delta p(y) = \rho A(y;x)dy\frac{dv}{dt}, \tag{5.35}$$

where acceleration is given by the time derivative of the particle velocity, dv/dt, assumed constant over this section of the reed.

Since volume flow is equal to particle velocity scaled by area, the expression for differential pressure as a function of position y along the reed channel is given by

$$\Delta p(y) = \rho\frac{dU}{dt}dy/A(y;x). \tag{5.36}$$

To obtain the differential equation for air flow, equation (5.36) is integrated over the length of the channel to obtain

$$p(0) - p(\mu) = \rho\frac{dU}{dt}\int_{y=0}^{y=\mu} dy/A(y;x). \tag{5.37}$$

where $y = 0$ is the channel entrance and $y = \mu$ is the point at which the flow separates from the surface of the reed and forms a jet. The pressure at the channel entrance is determined using Bernoulli's equation, (4.3), and is given by

$$p(0) = p_m - \frac{\rho}{2}\left(\frac{U}{A(0;x)}\right)^2, \tag{5.38}$$

where p_m is the mouth pressure. Since the pressure at the point of flow separation, $p(\mu)$, is equal to the bore pressure p_b, (5.37) becomes

$$p_m - p_b - \frac{\rho}{2}\left(\frac{U}{A(0;x)}\right)^2 = \rho\frac{dU}{dt}\int_{y=0}^{y=\mu} dy/A(y;x). \tag{5.39}$$

The differential equation governing volume flow is then given by

$$\frac{dU}{dt} = (p_m - p_b)\frac{A(x)}{\mu\rho} - \frac{U^2}{2\mu A(x)}, \tag{5.40}$$

where the flow is assumed to be in contact with the reed for a distance of μ at a constant area equal to

$$A(x) = A(0;x) = w(H_0 - x) \tag{5.41}$$

where w is the width of the reed (or jet) and H_0 is the opening of the valve channel in the absence of flow.

Displacement

The equation for displacement is determined much as described in Section 4.2.2 by considering the force acting on the clarinet reed. If the reed is rigid and hinged with spring constant k at a point far from the mouth side of the lip, let λ be the length of the reed roughly along the y axis which sees the mouth pressure, where μ is defined as before, and a force of

$$F_m = w\lambda p_m \tag{5.42}$$

which forces the reed closed [79]. In contrast, the force on the bore side of the reed away from the jet, given by

$$F_b = -w(\lambda - \mu)p_b, \tag{5.43}$$

forces the reed open. The force applied by the flow (which also forces the reed open) is found by integrating the pressure along the flow and is given by

$$F_U = -w\int_{y=0}^{y=\mu}\left(p_m - \frac{\rho}{2}\left(\frac{U}{A(y;x)}\right)^2\right)dy. \tag{5.44}$$

The overall force acting of the reed is obtained by summing (5.42), (5.43) and (5.44) and is given by

$$F = w(\lambda - \mu)(p_m - p_b) - w\mu\frac{\rho}{2}\left(\frac{U}{A(x)}\right)^2. \tag{5.45}$$

Once the force is known, the displacement of the reed is obtained using the familiar differential equation

$$\frac{d^2x}{dt^2} + 2\gamma\frac{dx}{dt} + \frac{k}{m}(x - x_0) = \frac{F}{m}, \tag{5.46}$$

where F is defined by (5.45).

By feathering the collisions of the beating reed in the clarinet simulation, the sound is improved and, in particular, the aliasing associated with this event is reduced as shown in Figure 5.12 when

compared with Figure 5.11.

5.6 Conclusions

The differential equation (5.1) describing the behaviour of volume flow (5.1) can be numerically unstable because of the singularity in the Bernoulli term when the valve closes. Moreover, aliasing is caused by abruptly cutting off flow when the computed flow passes through zero. Both problems are addressed by incorporating the new small-area solution for $U(t)$. The volume flow is now updated in a way which produces smoother transitions between open and close valves. This more accurate and numerically robust solution eliminates the original instability and reduces aliasing, as shown in the syrinx model spectrograms of Figures 5.7 and 5.8 and the clarinet model spectrograms of Figures 5.11 and 5.12. In each case, the control parameter most responsible for pitch is increased over time so there is a greater potential for aliasing. In both cases, the "feathered" valve gives a much cleaner output.

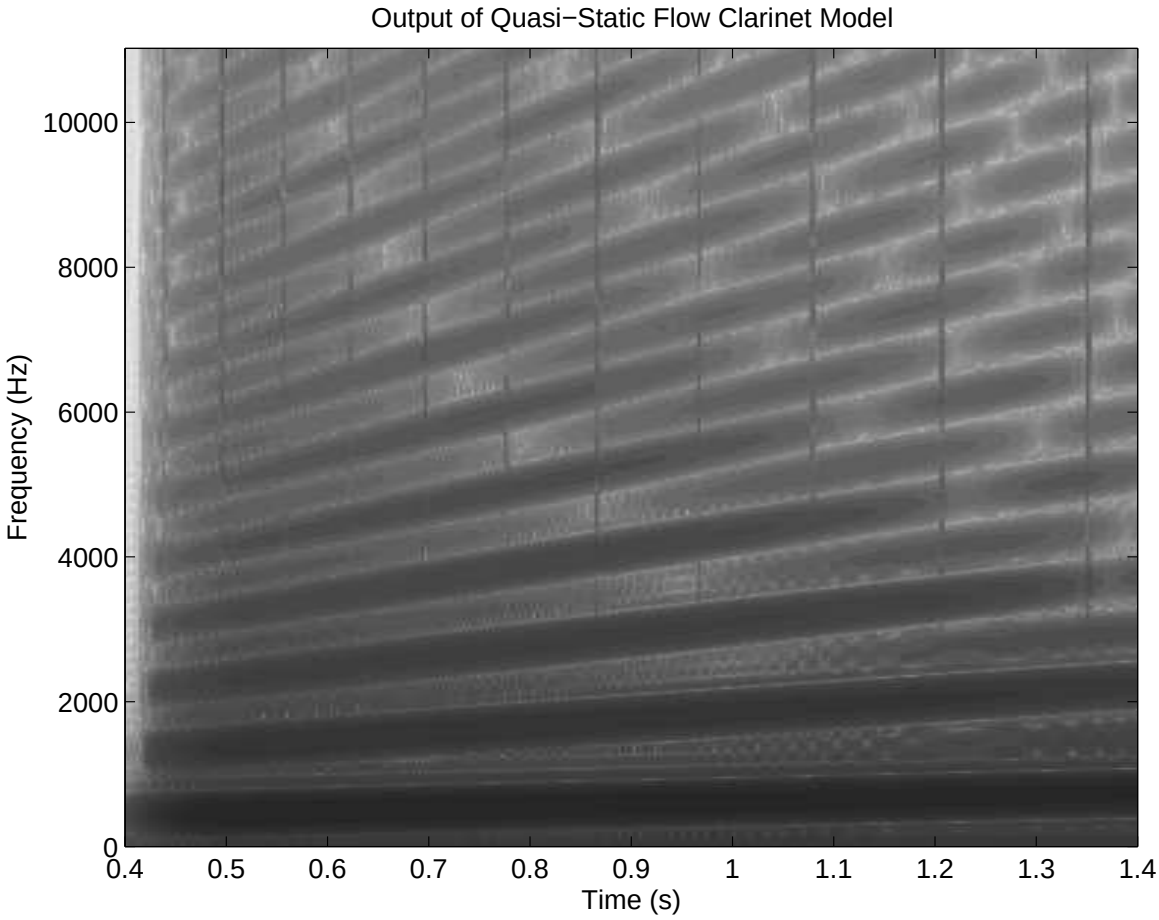


Figure 5.11: Output of the quasi-static model of the clarinet. Control parameter values were mouth pressure, 70 - 10 hPa, and frequency, 300-600 Hz. Lines through the spectrum illustrate undesirable artifacts.

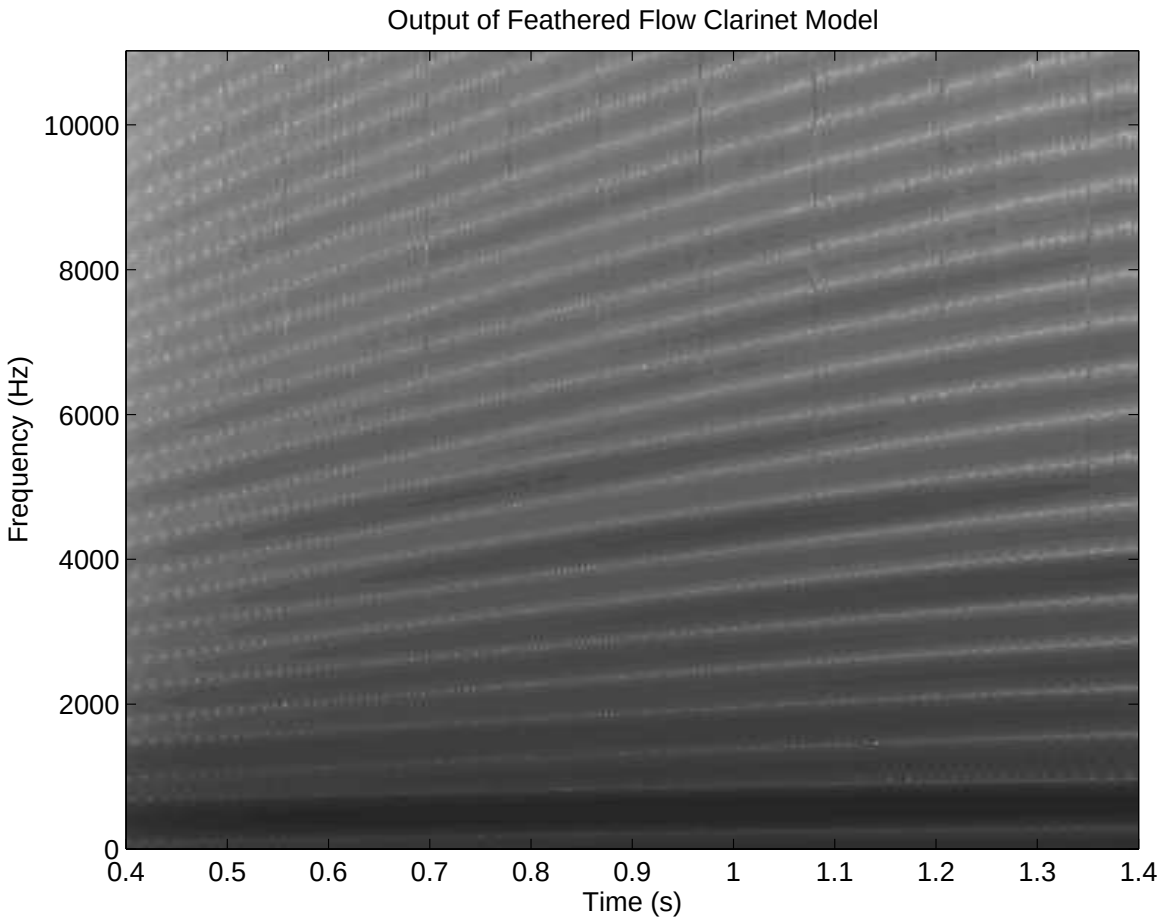


Figure 5.12: Output of the "feathered" dynamic model of the clarinet with the same control parameters as Figure 5.11. The output is improved overall and almost free of artifacts.

Chapter 6

Estimating the Control Parameters of the Syrinx

This chapter addresses the final stage in developing computer music instruments as outline in Chapter 1. In previous chapters the development of sound synthesis engines (and in particular, synthesis models of both the cicada and the song bird) and the development of a mechanical, haptic controller based on the cicada's buckling mechanism were discussed. The final stage is mapping, where intuitive relationships between the user's input (as obtained through the controller), the synthesis control parameters and the intended produced sound are determined—usually by way of software.

In a model as complex as the syrinx, there are numerous parameters that can be modified during performance, with some values for these parameters giving far better results than others. By restricting and scaling the control parameters of the model to something that would likely be used by a real bird, it is hoped that a more manageable range of likely parameter values can be isolated and made available to the user, improving his/her ability to interact with the model while still obtaining satisfactory results. Once this range of values is fully explored and mastered, it is up to the curious and adventurous player to extend the possibilities.

In this chapter, a method is presented for extracting the two primary control parameters of the model, lung pressure and membrane tension, from recorded birdsong. Much of the material included here has been previously published by Smyth, Abel and Smith in [20]. The method employs a *maximum likelihood* approach illustrated in Figure 6.1 to extract pressure and tension trajectories. The recording is first segmented to produce a sequence of power spectra frames over time which can be compared to the tabulated model power spectra using a likelihood function to indicate goodness of fit. Those control parameter sets having large likelihood at each frame are then aligned to form trajectories over time. Finally, if multiple parameter sets match the data well, the trajectories are adjusted to account for the ability of the bird to change pressure and tension using *minimum action*.

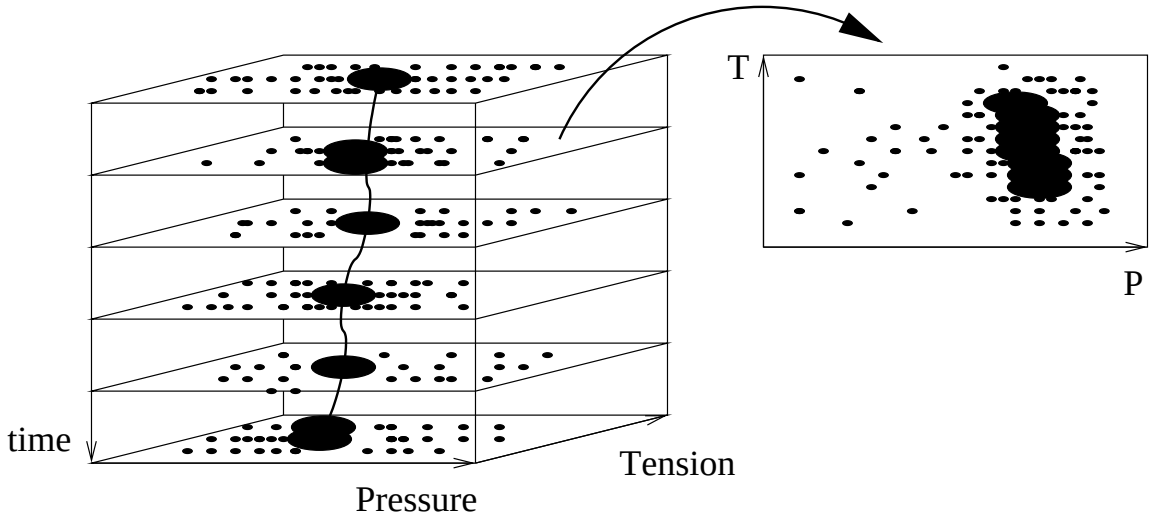


Figure 6.1: Values for pressure and tension are determined over time by threading trajectories through a stack of likelihood images.

Those aspects of the avian syrinx model (first introduced in Chapter 4) important to the parameter estimation algorithm will first be reviewed to give context to the discussion that follows. The maximum likelihood and minimum action approach is then described and the likelihood function detailed. Finally, the results obtained by applying the method to a field recording of a zebra finch are illustrated.

6.1 The Synthesis Model Revisited

Recall that the bird's airway consists of a trachea and a left and right bronchus. Though the bird is able to control the pressure and tension of the left and right side separately, for the purpose of simplification and to reduce the required amount of stored data, only one bronchus and one valve is considered for parameter estimation. For this reason, less successful results are obtained for extremely virtuosic birdsong, where the bird is likely making use of the left and right side of the syrinx to rapidly jump from high to low registers and in some cases to play two notes simultaneously [56].

The syrinx shows tremendous variation in structure between different bird species—example values of some anatomical parameters used in the model for various bird sizes are given in Table 6.1.

The model of the syrinx valve has the following four variables which evolve during sound production (also given in (4.1): the pressure on the bronchial side of the valve (p_0), air volume flow through the valve channel (U), displacement of the membrane (x) and the pressure on the tracheal side of

Quantity	Unit	Small	Medium	Large
Membrane density	kg/m^3	1000	1000	1000
Membrane width	mm	1.9	3.5	4.1
Membrane thickness	μm	100	100	100
Left bronchus length	mm	5	14	30
Right bronchus length	mm	5	14	30
trachea length	mm	17.1	23	35.6
Left bronchus radius	mm	1.47	2.5	3.5
Right bronchus radius	mm	1.47	2.5	3.5
Trachea radius	mm	1.9	3.5	4.1

Table 6.1: Examples of fixed anatomical parameters for different bird sizes. Though it is the case here, it is not necessary that the left and right bronchus be symmetrical.

the valve (p_1). Values for these variables are determined over time according to their corresponding differential equations as outlined in Chapter 4. These variables are illustrated on a simplified diagram of the valve in Figure 4.4 and a signal flow diagram of the complete model is shown in Figure 4.27.

The model of the valve displacement and the resulting pressure through the constriction is based on the mechanical properties of the membrane and the Bernoulli equation for the air flow (introduced in Chapter 4 and later developed in Chapter 5). The primary control parameters are therefore, the air pressure from the lungs and the tension in the membrane, the latter of which the bird controls by contracting the syringeal muscles and by raising the pressure in the interclavicular air sac which encases the syrinx.

There is a definite relationship between the pitch of the sound produced by the bird and the tension of the syringeal membrane. Recall that the displacement of the membrane as a function of time $x(t)$ for mode n is given by [46]

$$m_n \left[\frac{d^2 x_n}{dt^2} + 2\kappa \frac{dx_n}{dt} + \omega_n^2 (x_n - x_0) \right] = \epsilon_n F, \tag{6.1}$$

where the frequency of the first mode ω_1 is given by

$$\omega_1 = \left(\frac{5T}{\rho_M a H d} \right)^{\frac{1}{2}}, \tag{6.2}$$

with T being tension, ρ_M the membrane density, a the radius of the bronchus, H the membrane width and d the membrane thickness (with example values given in Table 6.1). It can be seen

therefore, that when the bird increases the tension of the membrane the pitch of the song will also increase (though (6.2) shows this relationship is not linear). Likewise, an increase in air pressure from the lungs will cause a general increase in the amplitude of the produced sound.

The mapping of pressure and tension to loudness and pitch respectively is not straightforward however, since nonlinearities intrinsic to the dynamics of the syrinx cause less predictable behaviour [45]. A slight change in one parameter, for instance, can cause effects such as period doubling, mode-locking and transitions from periodic to chaotic behaviour making it extremely difficult and unintuitive to control (though of course, a bird possessing this instrument since hatching would have a much better time of it). It would also be somewhat unreasonable to ask a player of the model—already facing these inherent difficulties—to control the parameters with the speed and virtuosity of a songbird unless s/he is equipped with a special controller making this possible. It is clear therefore, that another mapping strategy besides a one-to-one relationship is necessary.

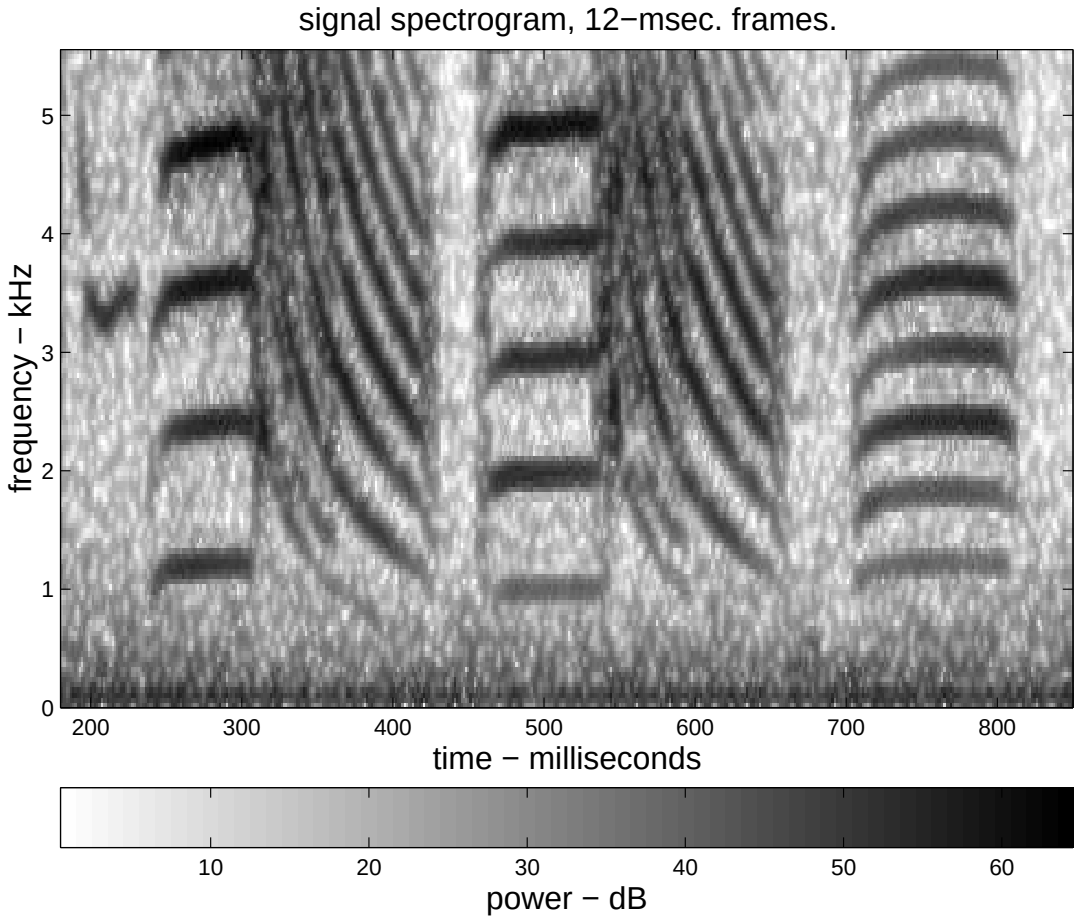


Figure 6.2: Power spectra for field recording of a zebra finch.

6.2 Maximum Likelihood Model

First, the fixed anatomical parameters of the model are specified according to the bird species whose song is to be simulated and of which there is a recording. In this case the technique is illustrated making reference to a recording of a zebra finch (the spectrogram for which is shown in Figure 6.2) and makes use of the small bird anatomical parameters from Table 6.1.

A look-up table is generated by pairing combinations of the control parameters pressure and tension with the model's corresponding output power spectra. Figure 6.3 shows an excerpt of the spectrum table for tension ranging from zero to eight at five pressure values. The actual table has many more entries and since it only need be computed once (for each bird species), it can be made as large and dense as is permitted by computer memory.

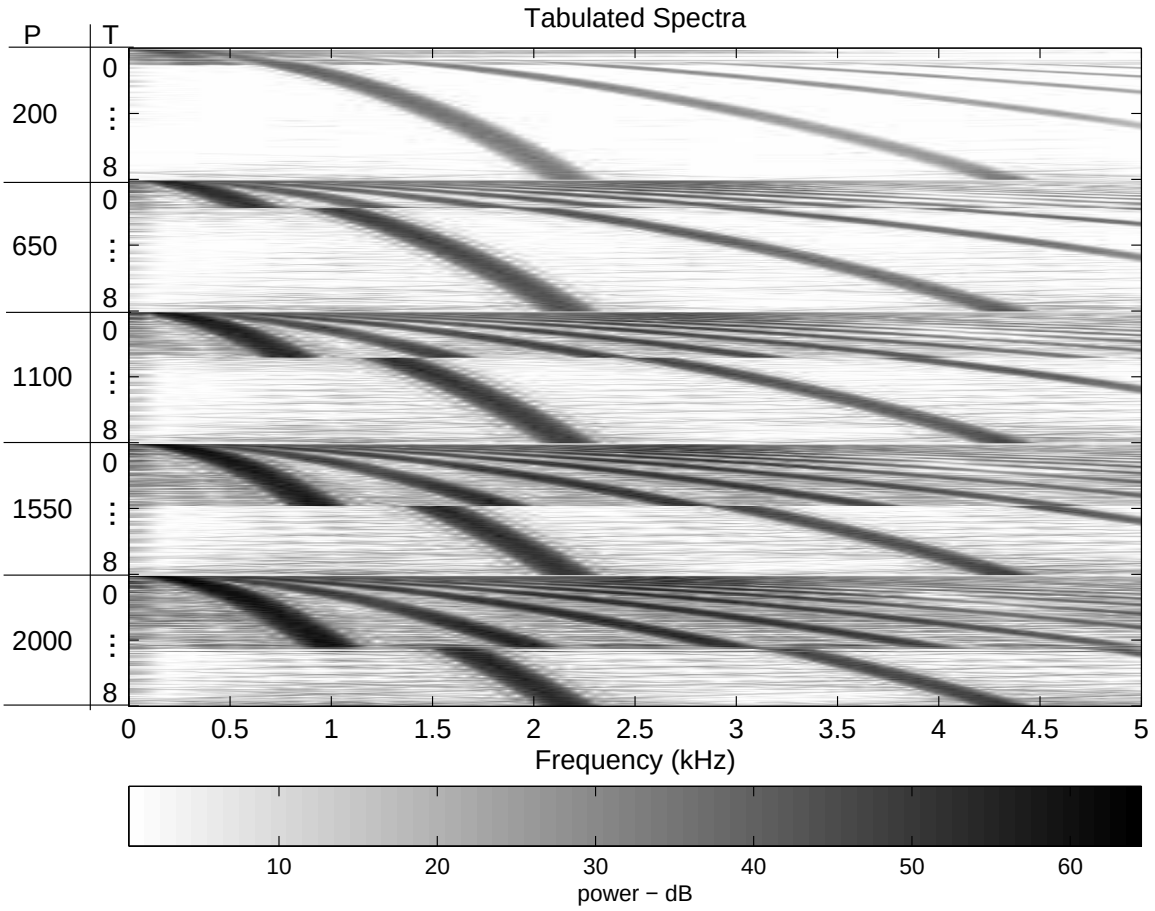


Figure 6.3: An excerpt of a look-up table that pairs combinations of the two primary control parameters, pressure from the lungs 'P' (in pascals (Pa)) and tension in the membrane 'T' (in newtons (N)), with the model's corresponding output power spectra.

An extension of this table would allow for combinations of pressure and tension values for both the left and right side of the syrinx. However, this would contribute unnecessary complexity for the purpose of illustrating the maximum likelihood model (and, of course, would increase the size of the table substantially) so for now, this is left to future work.

With the tabulated model spectra in place, the recorded birdsong is processed via a short-time Fourier transform [68] to form a sequence of power spectra over time. An example spectrogram in Figure 6.2, taken from a field recording, shows a sequence of gestures for the song of a zebra finch. As is typical of field recordings, it contains a significant amount of background noise concentrated in the low frequencies.

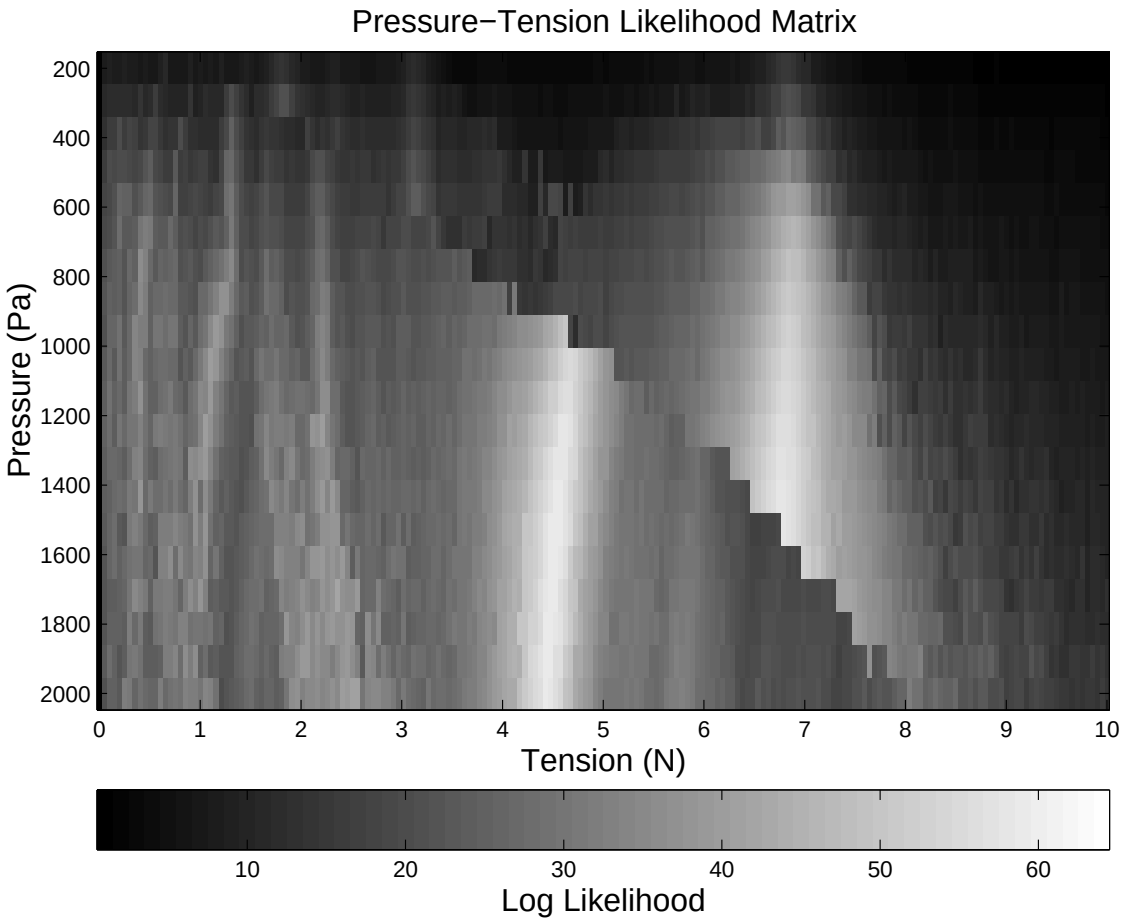


Figure 6.4: A pressure-tension matrix for one time frame. The lighter pixels indicate pressure-tension settings producing spectra very similar to the measured frame, while the darker pixels indicate pressure-tension settings producing very different spectra. In the case where there are two regions showing strong likelihood, the values for the parameters are selected by incorporating minimum action.

At each time frame, a normalized *log likelihood function* [72] is used to indicate the similarity between the birdsong power spectrum and each of the entries in the spectrum table, filling a pressure-tension matrix similar to the one shown in Figure 6.4. The likelihood function used here is normalized so that entries having a value close to one—indicated by the lighter pixels in Figure 6.4—correspond to pressure-tension settings producing spectra very similar to the measured frame spectrum. Those producing entries close to zero—indicated by the darker pixels—have very different spectra from the measured frame under consideration. Occasionally there will be well separated likelihood maxima corresponding to different registers or modes of oscillation with a division between the two (this is very apparent in Figure 6.4). This is handled by incorporating *minimum action* discussed below in Section 6.3.

The likelihood function was chosen for its statistical properties: In the limit of small estimation errors, its peak location is known to be unbiased with minimum variance [72]. In other words, the likelihood function will, on average, peak at the correct control parameter values, and the peak location will be as insensitive as possible to measurement noise.

It is assumed that the recorded signal $s(t)$ consists of the model output $\mu_\theta(t)$ with measurement and/or modeling error $\nu(t)$ which is additive and Gaussian-distributed, that is,

$$s(t) = \alpha\mu_\theta(t) + \nu(t), \tag{6.3}$$

where α is an unknown scaling factor and $\theta = [P, T]$ is the parameter vector containing pressure P and tension T.

Assuming the measured signal is stationary over the duration of an analysis frame, the power spectrum of the measured signal is given by

$$S(\omega) = \alpha^2 M_\theta(\omega) + N(\omega) \tag{6.4}$$

where M_θ represents the model output power spectrum for parameters θ and $N(\omega)$ represents the noise power spectrum at frequency ω. In the case that the noise power $N(\omega)$ is much smaller than the signal, the probability of observing a power spectrum $\mathbf{S}$ given control parameters θ may be approximated by

$$\begin{aligned} \wp(\mathbf{S};\theta) &\approx \mathcal{N}(\alpha^2\mathbf{M}_\theta, \Sigma) \\ &= \frac{\exp\{-\frac{1}{2}(\mathbf{S}-\alpha^2\mathbf{M}_\theta)^\top\Sigma^{-1}(\mathbf{S}-\alpha^2\mathbf{M}_\theta)\}}{\det(2\pi\Sigma)^{\frac{1}{2}}} \end{aligned} \tag{6.5}$$

where $\mathbf{S}$ is the stack containing measured power spectrum bins $S(\omega_i)$ and similarly $\mathbf{M}_\theta$ is the stack of model power spectrum bins.

Under the assumption that the additive spectral noise is independent and identically distributed,

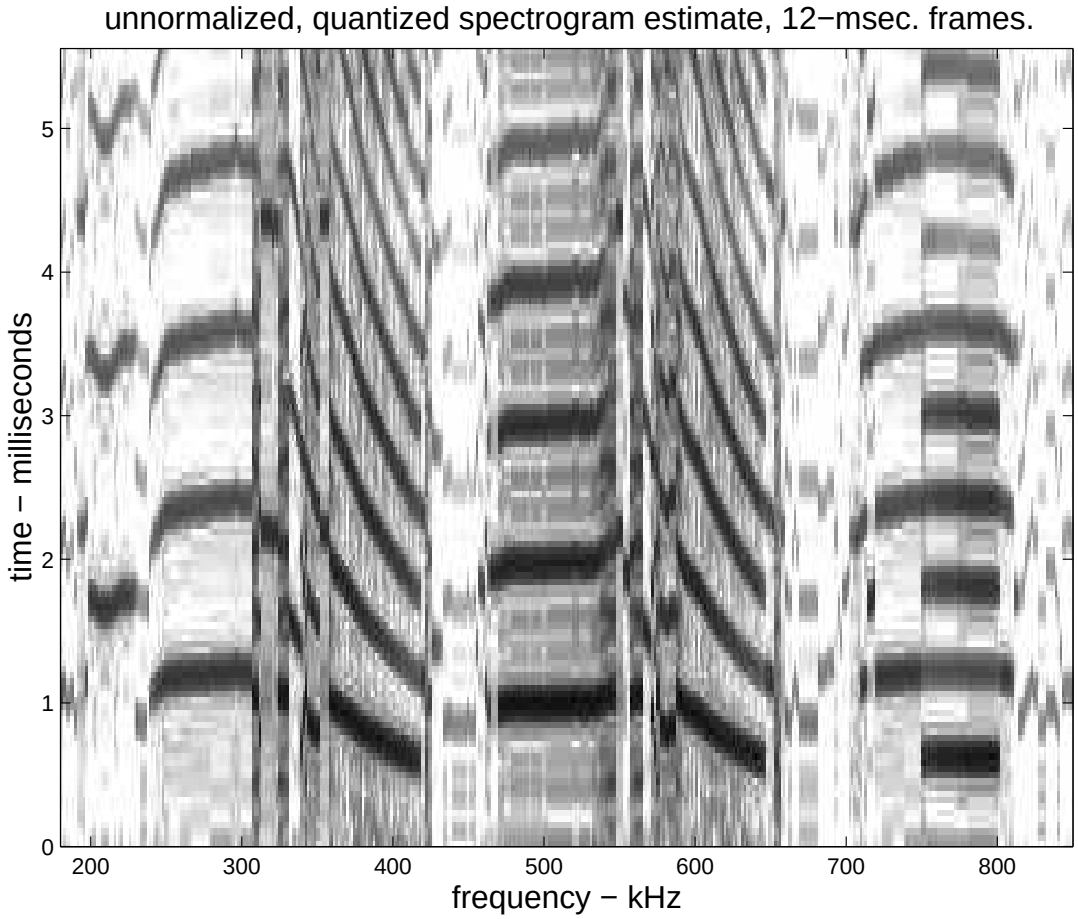


Figure 6.5: A sequence of power spectra from the look-up table based on maximum likelihood control estimates.

the log likelihood for a parameter set θ given measurements $\mathbf{S}$ is

$$l(\theta;\mathbf{S}) = -\frac{1}{2\sigma^2}(\mathbf{S}-\alpha^2\mathbf{M}_\theta)^\top(\mathbf{S}-\alpha^2\mathbf{M}_\theta). \tag{6.6}$$

For convenience, the likelihood is normalized so that it ranges from zero, indicating an unlikely match, to one, indicating a good fit:

$$\bar{l}(\theta;\mathbf{S}) = 1 - \frac{(\mathbf{S}-\alpha^2\mathbf{M}_\theta)^\top(\mathbf{S}-\alpha^2\mathbf{M}_\theta)}{(\mathbf{S}+\alpha^2\mathbf{M}_\theta)^\top(\mathbf{S}+\alpha^2\mathbf{M}_\theta)}. \tag{6.7}$$

The second term from (6.7) may be interpreted as the squared distance between the scaled measured data and the table entry, normalized by the squared length of their sum. When the parameters are such that the measured and model spectra coincide, the numerator above will be small and the

likelihood close to one. When the measured and model spectra are very different, the numerator is large—it can never be larger than the denominator as both spectra are positive—and the likelihood is close to zero.

6.3 Minimum Action

Typically there will be a range of likelihood ratios that match the data well. The pressure-tension likelihood functions tend to have maxima that are wide in pressure and narrow in tension. On occasion there will be well separated likelihood maxima corresponding to different registers or modes of oscillation with a division between the two (this is very apparent in Figure 6.4). It is assumed that the bird will not use any unnecessary energy during song performance and in the event of multiple well separated matching maxima the one requiring the least effort on the part of the bird is chosen.

Minimum action is introduced into the likelihood function in two ways: one takes into account the effort involved to move from one parameter value to another over time while the other simply incorporates the instantaneous effort where higher values of tension and pressure are considered to require more effort.

For the instantaneous effort it is assumed that it is more difficult to produce higher values of pressure and tension and represent this added difficulty with a penalty function added to the likelihood function. On the assumption that it is difficult for the bird to rapidly slew control parameters, the sequence of likelihood function maximizers are median filtered to eliminate sporadic, short lived, jumps in the trajectories. A further slew limiting filter is applied to ensure rates of parameter change that are physically plausible.

6.4 Conclusions

Estimating the two primary control parameters of the syrinx, lung pressure and membrane tension, serves two purposes: 1) to judge the model's ability to produce realistic birdsong by calibrating it to recorded birdsong and 2) to restrict and scale the control parameter space in an effort to improve the user's ability to interact with the model by, for instance, having a controller that follows predetermined trajectories through the matrix stack.

The parameter estimation method was applied to the zebra finch song (the spectrogram of which is shown in Figure 6.2). Tabulated power spectra from the look-up table corresponding to the most likely parameters at each time frame are sequenced to produce the spectrogram shown in Figure 6.5. The similarity between the original zebra finch spectrogram (Figure 6.2) and the constructed sequence (Figure 6.5 shows very good potential in the model's ability to estimate likely parameter values and ultimately to produce accurate bird-like song.

The final control trajectories extracted from the zebra finch recording are shown in Figure 6.6.

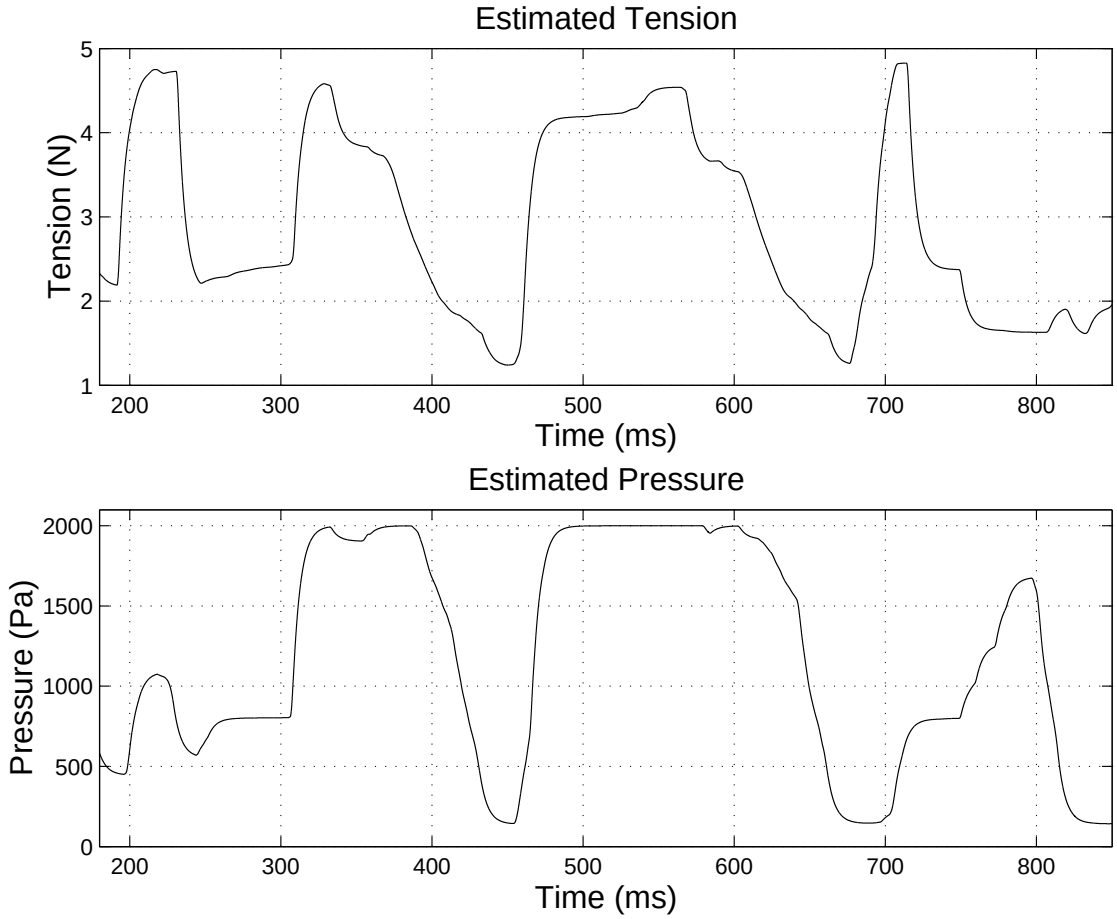


Figure 6.6: Zebra finch control trajectories for pressure and tension over time as determined by connecting most likely values from a stack of pressure-tension matrices (one of which is shown in Figure 6.4).

The spectrogram of the corresponding model output is shown in Figure 6.7. Though not identical to either Figure 6.2 or Figure 6.5 the sound output is perceptually similar and the song gesture is well captured. These results would likely be greatly improved if the look-up table were expanded to include data from both sides of the syrinx—since independent control of left and right sides of the syrinx is what, in most cases, enables the bird to sing virtuosic songs [56]. The results would also be much improved if more accurate measurements of the syrinx anatomy were available for different bird species. In this research, anatomical parameter values were only approximated for small, medium and large birds. Given the variety of bird species, this is clearly an oversimplification. Finally, if original bird recordings were less noisy, the parameter estimation algorithm wouldn't attempt to have the model produce this noise—something which is likely causing some error—and the output of the model would likely be a closer match to the original recording

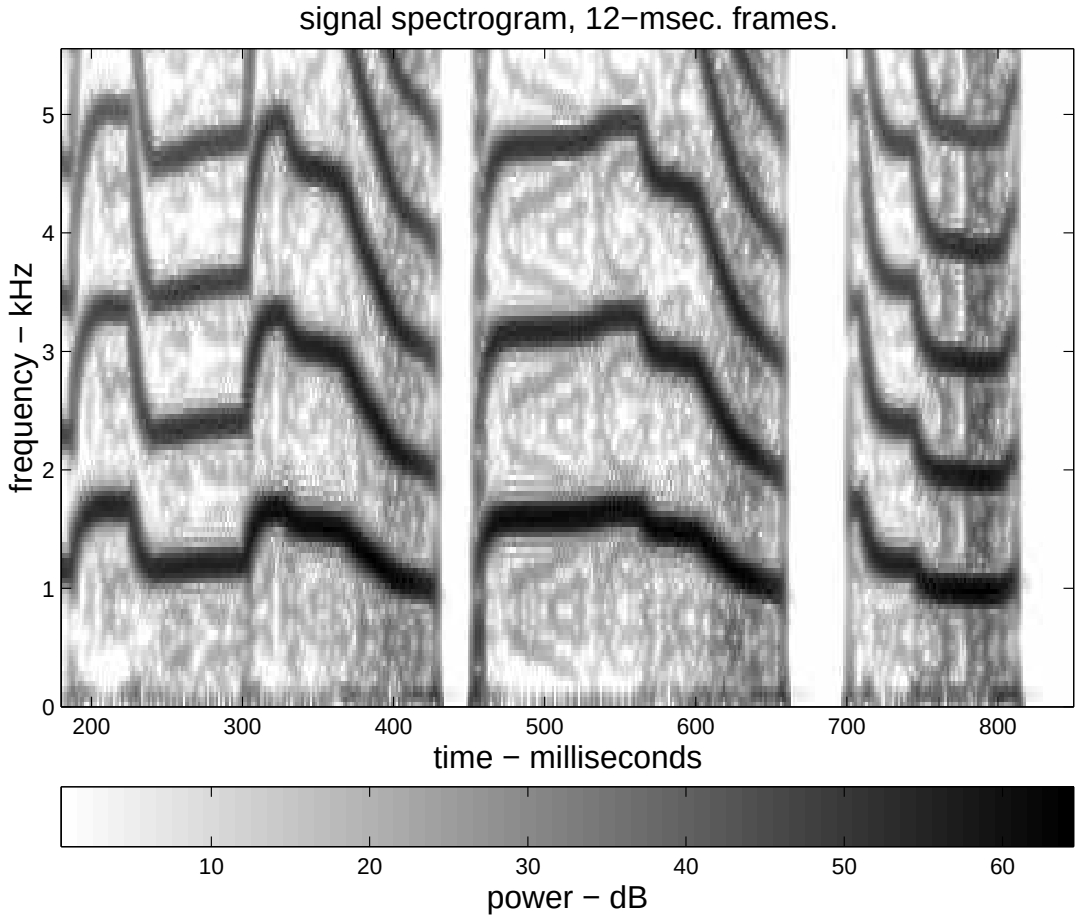


Figure 6.7: Power spectra of model output using control trajectories from Figure 6.6.

By mapping the pressure and tension trajectories of Figure 6.6 to an input device with two individual continuous controls, the player of the model is able to reproduce bird-like song without requiring the bird's expert technique. Figure 6.8 shows how the trajectories were also used in a non real-time Matlab graphical user interface to the model. A drop down menu allows different trajectories to be selected and other menu boxes permit the user to set the anatomical parameters as s/he wishes.

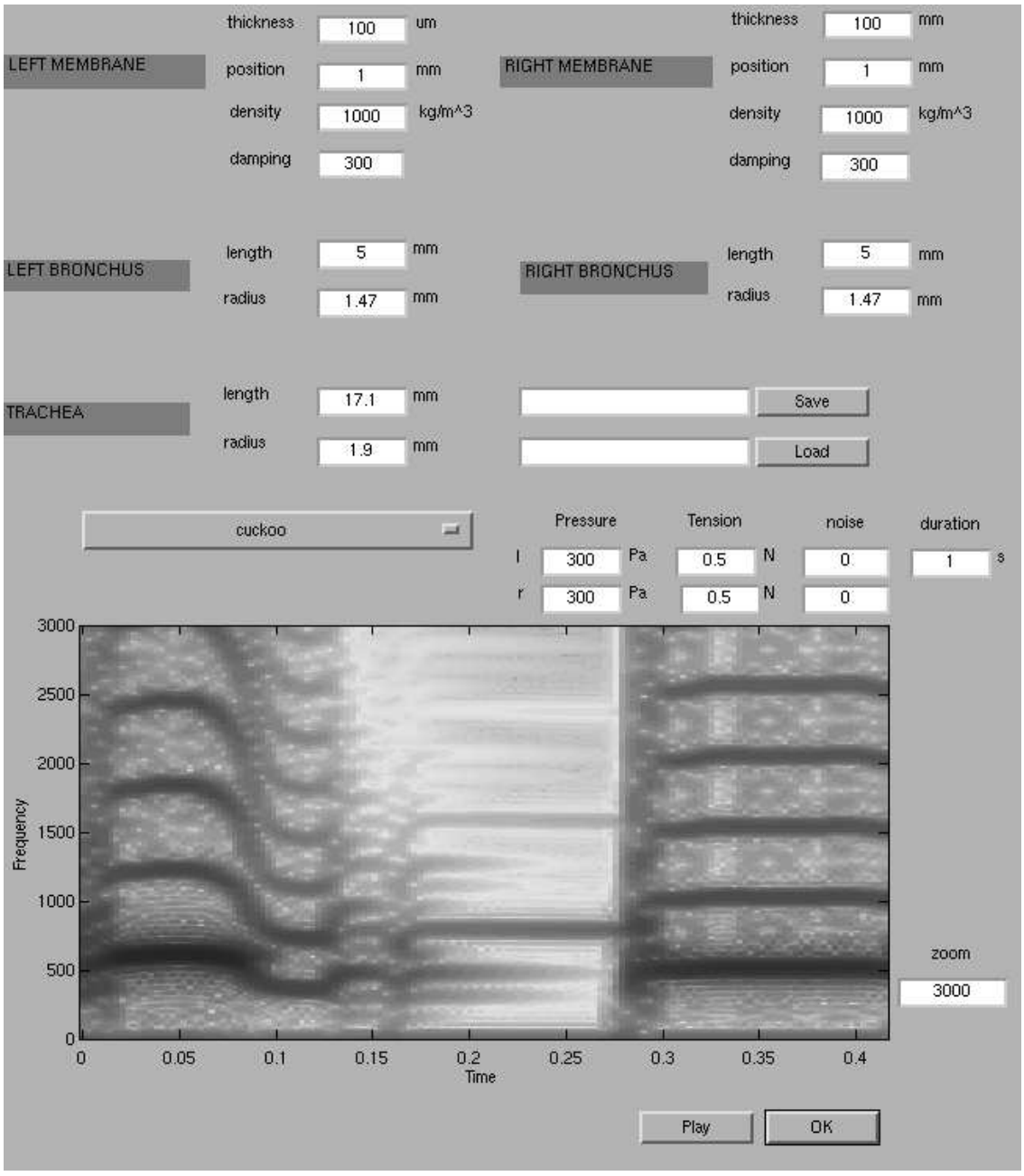


Figure 6.8: Graphic user interface in Matlab shows an example use of the pressure-tension control trajectories for the cuckoo bird. A drop down menu allows this and other other possible control trajectories (for other bird species) to be loaded. Other menu buttons allow the user alter the anatomical parameters. The 'OK' button computes the spectrogram and the 'Play' button allows the user to hear the model's produced sound.

Chapter 7

Conclusions and Future Research

7.1 Conclusions

Researchers are increasingly looking at biological structures for new approaches to design problems, and many advances in technology are the result of taking a closer look at natural processes. Researchers have found, for instance, that iron is the key ingredient for a *super glue* allowing the common blue mussel to cling to rocks by forming an extremely strong binding agent under seawater. This will provide a new adhesive particularly suited to surgery since it works well in wet conditions [86]. The gecko, a lizard seemingly capable of defying gravity with its ability to cling to ceilings and polished glass, has inspired the creation of the *mecho-gecko*, a robot with sticky feet and a tail for balance that can scuttle across dangerous terrain [90]. Computer models of insect swarms have allowed researchers to determine the best route for London's annual Nottingham Hill carnival, decreasing the likelihood of crowding and making the event safer for the public [94]. The foraging ant which, though almost blind, is able to track the shortest distance from anthill to food, has also helped researchers in information technology develop networks which flow more efficiently [89]. Insects have also inspired solutions for various problems in robotics [87, 80].

This research focused on animal sound production mechanisms applied to various aspects of computer instrument technology: sound synthesis, music controllers, and the mapping between the two. Though this work focused on the sound production of the cicada and the bird, if larger in scope, many other animal sound production mechanisms could have been included.

The cicada and the songbird both possess elements in their vocal organs that are not typically found in musical instruments. The cicada uses a rapid sequence of buckling ribs to initiate and efficiently sustain vibrations of the tymbal plate. The coupling of the vibrating plate to the abdominal air sac tuned to the same frequency and the large tympana, or eardrums, which improve impedance matching between insect and air, allow the sound to propagate freely, resulting in the loud sustained tone for which the cicada is renowned. The song bird has a unique topography of acoustic elements,

a trachea and a bifid bronchi with two pressure controlled valve in the transverse configuration (not typical of musical instruments whose valves are more accurately described as being blown open or close). Having these two separate valves which can be operated independently, allows the bird to sing extremely virtuosic song, reserving one side of the syrinx for the high registers and the other side for the low registers. Because the bird is capable of switching the left and right sides so rapidly, the bird's song contains rapid interval leaps that are not even perceptible by the human ear (unless slowed down in a recording) and in many cases, can play two tones simultaneously.

By creating a physical model of the cicada's sound mechanism, and in particular, by creating a real-time, interactive means of controlling it, a mechanical device, with natural haptic feedback was developed which is very responsive to the player and allows him/her to repeat notes extremely rapidly. This device, which allows for a sort of plucking mechanism with after-touch [92], does not necessarily have to control the cicada model as it has also been shown to work very well with a string model (among other models).

From the research on the syrinx came a very accurate model of the transverse configuration of a pressure controlled valve—a configuration not yet modeled in computer music since it is not typically found in existing musical instruments. In addition to offering this new configuration, the syrinx model also inspired a solution for *beating*, a term referring to a valve which collides with the opposite wall of the valve channel, thus closing the channel completely. This event is known to cause aliasing in discrete time signals, since it often terminates the flow too abruptly to be accurately captured at relatively low audio sampling rates. The softer biological tissue in the syrinx gave rise to the idea that this event could be slowed down over time. The *feathered* valve was later used to improve an existing model of the clarinet.

7.2 Future Work

Though the classic valve/reed model was significantly improved in this research, there is still room for enhancements. More attention to mapping (predetermined relationships between the user input and the synthesis parameters) would heighten the stability of the model and make it even more suitable for real-time performance. It may also be useful to apply different modeling methods for the valve/reed (such as a wave digital filter) so that individual modes would not have to be expressly modeled but rather would fall out of the model naturally and also to ensure stability of the model. Finally, incorporating turbulence, something not yet addressed but for which there is a definite need, would likely make a tremendous improvement to the overall quality of the produced sound.

Different models of the pressure-controlled valve are used in woodwind and brass musical instruments. These instruments (along with the human voice) generally use only one valve with a single connected bore. Historically however, there existed instruments with multiple valve/bore

topologies—much like in the avian syrinx. It would be very worthwhile to determine what musical purpose these valve configurations served, as compared to the song bird's use of a multi-valve topology, in order to expand the current use of the valve in musical instruments (both acoustic and computer/electronic). It is known, for example, that placing two valves in series, such as when a didjeridu player vocalizes into the instrument and places the vocal folds in series with the lips, produces extremely low frequencies equal to the difference in the fundamental frequency of both vibrating valves [30]. What musical effects can be obtained by placing two valves in parallel?

Appendix A

Building the *Tymbalimba*

As much of the work in this project involved designing and building the mechanical controller, a section is included here describing the methods used.

Machining Methods

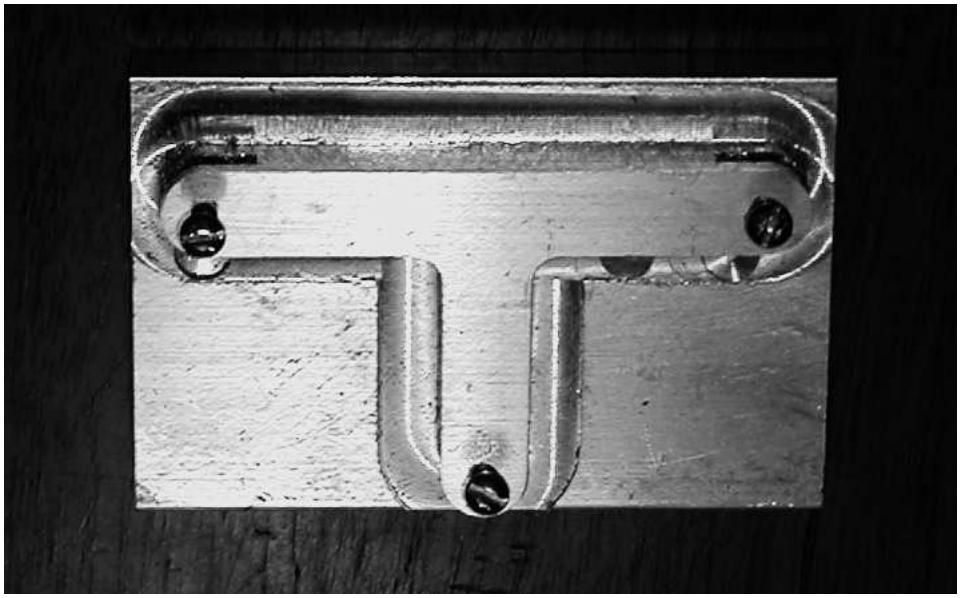


Figure A.1: Fixturing the T-part for CNC machining.

The parts of the model were machined from aluminum (6061-T6) on a CNC mill at Stanford University's Product Realization Lab (PRL). Fixturing of the ribs (T-parts) was done with each part attached to a block of aluminum via socket-cap screws. Each screw was placed at each of the

3 tips of the T-parts, taking advantage of the need for holes in these locations to place the axial shafts when assembling the device.

The blocks which suspend the ribs, referred to as the L and C-parts, were also machined on a CNC mill. They are attached to the base plate through a long channel cut into the plate, using socket-head screws. The user need only adjust the screws to loosen the L and C-parts so they can slide easily through the channel, thereby adjusting their relative position to one another.

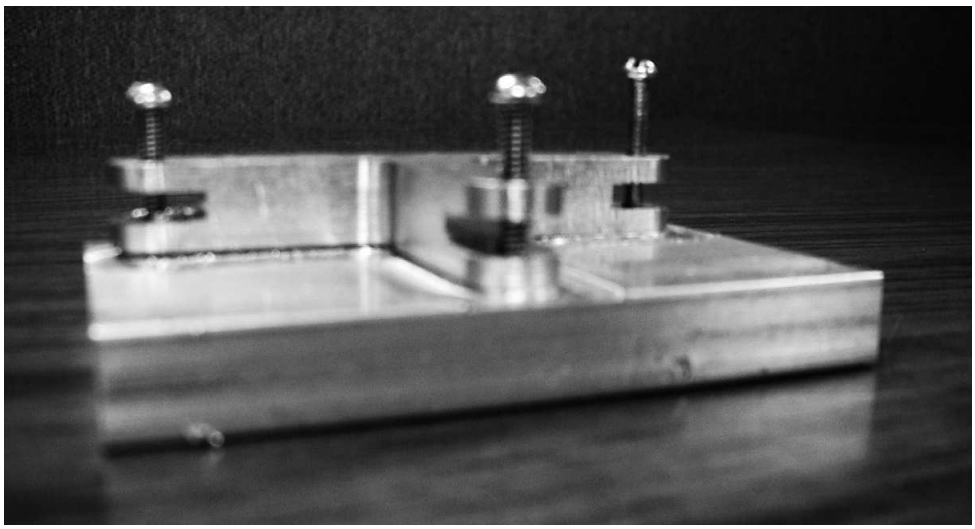


Figure A.2: A side view of the T-part when fixtured for CNC machining.

Code for Basic Stamp (BS2sx)

```
'{$STAMP BS2sx}

clockpin con 15
inputpin con 12
outputpin  con 13
lsb con 0
msb con 1
behindclock con 2   'DATA is msb first;
                        'sample bits after clock pulse (MSBPOST)
                        'first data bit ready after first clock pulse
'baudmode con 16407  'in((2,500,000/57,600)-20 + 16384 (16384 = inverted)
baudmode con 16385
baudmode96 con 16624
'baudmodeFast   con 16385  '115200 baud
```

```
MSBflag con %10000000
LSBflag con %00000000
nchannels con 4

MSBdata var nib
LSBdata var byte
channel var byte
controlbyte  var byte
outbyte1 var byte
outbyte2 var byte

DIRS = %1111111101111111

channel = 0

''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
'CONTROL BYTE
''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
' bit 7 | bit 6 | bit 5 | bit 4 | bit 3 | bit 2 | bit 1 | bit 0
' start | sel2  | sel1  | sel0  | rng   | bip | pd1   | pd0
'
' start: defines beginning of controlbyte
' sel2 sel1 sel0: 3 bits select the desired on channel
' rng: selects the voltage range
' bip: selects unipolar or bipolar conversion mode
' pd1 pd2: select clock and power-down modes
'
' input range | rng | bip
' ------------------------
'   0-5V      | 0   | 0
'   0-10V     | 1   | 0
'   +-5V      | 0   | 1
'   +-5V      | 1   | 1

controlbyte = %10001001

outbyte1 = 0
```

```
outbyte2 = 0

''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
'PROTOCOL
''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
'
'16 bit message sent in 2 bytes, outbyte1 and outbyte2 (10 bits for data)
'1           0000        000         0        0000000
'MSBflag | Chan (0-15) | MSBdata | LSBflag |    LSBdata
'       outbyte1                 |         outbyte2
'
''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
top:
FOR channel = 0 TO (nchannels - 1)
  shiftout outputpin, clockpin, msb, [controlbyte \ 8]
  shiftout outputpin, clockpin, msb, [0 \ 4]
  shiftin inputpin, clockpin, behindclock, [MSBdata\3, LSBdata\7]

  outbyte1 = MSBflag | (channel << 3) | MSBdata
  outbyte2 = LSBflag | LSBdata
  serout 16, baudmode, [outbyte1, outbyte2]
  controlbyte = controlbyte + 16
NEXT

controlbyte = %10001001

goto top:

printdata:
 outbyte1 = ((MSBdata << 7) + LSBdata) >> 4
 debug dec ? outbyte1
return
```

Appendix B

Computing Shelf Filter Coefficients

This appendix presents formulas for computing the coefficients of the first-order shelf filters used in Section 4.5.

The transfer function for the first-order shelf filter is given by

$$\sigma(z; f_t, g_\pi) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1}}. \tag{B.1}$$

Having unity gain at DC, a band edge gain $g_\pi \geq 0$ (the gain at the Nyquist limit) and a gain $\sqrt{g_\pi}$ at the transition frequency f_t measured in radians/π, the coefficients in (B.1) are given by

$$b_0 = \frac{\beta_0 + \rho\beta_1}{1 + \rho\alpha_1}, \tag{B.2}$$

$$b_1 = \frac{\beta_1 + \rho\beta_0}{1 + \rho\alpha_1}, \tag{B.3}$$

$$a_1 = \frac{\rho + \alpha_1}{1 + \rho\alpha_1}, \tag{B.4}$$

where

$$\beta_0 = (1 + g_\pi)/2 + (1 - g_\pi)\alpha_1/2, \tag{B.5}$$

$$\beta_1 = (1 - g_\pi)/2 + (1 + g_\pi)\alpha_1/2, \tag{B.6}$$

and

$$\alpha_1 = \begin{cases} 0, & g_\pi = 1 \\ \eta - \mathrm{sign}(\eta)(\eta^2 - 1)^{\frac{1}{2}}, & g_\pi \neq 1. \end{cases} \tag{B.7}$$

To form the coefficients of a prototype shelf filter with a transition frequency of $\pi/2$, η from (B.7)

is set to the following:

$$\eta = (g_\pi + 1)/(g_\pi - 1), \quad g_\pi \neq 1. \tag{B.8}$$

The parameter

$$\rho = \sin(\pi f_t/2 - \pi/4)/\sin(\pi f_t/2 + \pi/4) \tag{B.9}$$

is the coefficient of the first-order allpass transformation warping the prototype filter to the proper transition frequency.

Bibliography

Physical Modeling for Music and Audio

[1] Jonathan Abel, Tamara Smyth, and Julius O. Smith, "A simple, accurate wall loss filter for acoustic tubes," in *DAFX 2003 Proceedings*, London, England, September 2003, International Conference on Digital Audio Effects.

[2] Seiji Adachi and Masa aki Sato, "Trumpet sound simulation using a two-dimensional lip vibration model," *Journal of the Acoustical Society of America*, vol. 99, no. 2, pp. 1200–1209, February 1996.

[3] Julien Bensa, Stefan D. Bilbao, Richard Kronland-Martinet, and Julius O. Smith, "From the physics of piano strings to digital waveguides," in *Proceedings of ICMC 2002*, G'oteborg, Sweden, September 2002, International Computer Music Conference, pp. 45–48.

[4] Julien Bensa, *Analysis and Synthesis of piano sounds using physical and signal models*, Ph.D. thesis, Laboratoire de Méchanique et d'Acoustique (LMA), Marseille, France, 2003.

[5] Perry R. Cook, *Identification of Control Parameters in an Articulatory Vocal Tract Model, with Applications to the Synthesis of Singing*, Ph.D. thesis, Stanford University, Stanford, California, December 1990.

[6] Perry R. Cook, "A meta-wind-instrument physical model, and a meta- controller for real time performance control," in *International Computer Music Conference*, San Jose, California, October 1992.

[7] Patricio De la Cuadra, Tamara Smyth, Chris Chafe, and Han Baoqiang, "Waveguide simulation of neolithic chinese flutes," in *Proceedings of ISMA 2001*, Perugia, Italy, September 2001, International Symposium on Musical Acoustics.

[8] Mark Kahrs and Frederico Avanzini, "Computer synthesis of bird songs and calls," in *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFX-01)*, Limerick, Ireland, December 2001.

[9] Gary Paul Scavone, *An Acoustic Analysis of Single-Reed Woodwind Instruments with an Emphasis on Design and Performance Issues and Digital Waveguide Modeling Techniques*, Ph.D. thesis, CCRMA, Music Dept., Stanford University, Stanford, California, March 1997, available as CCRMA Technical Report No. STAN–M–100 or from ftp://ccrma-ftp.stanford.edu/pub/Publications/Theses/GaryScavoneThesis/.

[10] Julius O. Smith, "A new approach to digital reverberation using closed waveguide networks," in *Proceedings of the 1985 International Computer Music Conference, Vancouver*, Vancouver, British Columbia, 1985, Computer Music Association, pp. 47–53.

[11] Julius O. Smith, "Physical modeling using digital waveguides," *Computer Music Journal*, vol. 16, no. 4, pp. 74–91, 1992, Special issue: Physical Modeling of Musical Instruments, Part I.

[12] Julius O. Smith, "A basic introduction to digital waveguide synthesis," www-ccrma.stanford.edu/~jos/swgt/, October 2003.

[13] Julius O. Smith, "Discrete time lumped models," www-ccrma.stanford.edu/~jos/NumericalInt/, March 2002, Course notes for MUS421/EE367B, Stanford University.

[14] Julius O. Smith, "Wave digital filters," www-ccrma.stanford.edu/~jos/WaveDigitalFilters/, March 2002, Course notes for MUS421/EE367B, Stanford University.

[15] Julius O. Smith, *Digital Waveguide Modeling of Musical Instruments*, www-ccrma.stanford.edu/~jos/waveguide/, 2003.

[16] Tamara Smyth and Julius O. Smith, "A musical instrument based on a bioacoustic model of a cicada," in *Proceeding of ICMC 2001*, Havana, Cuba, Septemper 2001, International Computer Music Conference.

[17] Tamara Smyth and Julius O. Smith, "Applications of bioacoustics in physical modeling and the creation of new musical instruments," in *Proceedings of ISMA 2001*, Perugia, Italy, September 2001, International Symposium on Musical Acoustics.

[18] Tamara Smyth and Julius O. Smith, "The sounds of the avian syrinx—are they really flute-like?," in *DAFX 2002 Proceedings*, Hamburg, Germany, September 2002, International Conference on Digital Audio Effects.

[19] Tamara Smyth and Julius O. Smith, "The syrinx: Nature's hybrid wind instrument," in *CD-ROM Paper Collection*, Cancun, Mexico, September 2002, Pan-America/Iberian Meeting on Acoustics.

[20] Tamara Smyth, Jonathan Abel, and Julius O. Smith, "The estimation of birdsong control parameters using maximum likelihood and minimum action," in *Proceedings of SMAC 03*, Stockholm, Sweden, August 2003, Stockholm Music Acoustics Conference.

[21] Tamara Smyth, Jonathan Abel, and Julius O. Smith, "Discrete-time simulation of air-flow cut-off in pressure-controlled valves," New Paltz, New York, October 2003, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics.

[22] Tamara Smyth, Jonathan Abel, and Julius O. Smith, "Feathered collisions in beating reed simulation," Austin, Texas, November 2003, Acoustical Society of America, by invitation.

[23] Vesa Välimäki, *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*, Ph.D. thesis, Helsinki University of Technology, Faculty of Electrical Engineering, Laboratory of Acoustic and Audio Signal Processing, Espoo, Finland, 1995, Report no. 37.

[24] Scott Van Duyne and Julius O. Smith, "The 2-d digital waveguide mesh," in *Proceedings of the 1993 IEEE Workshop of Applied Signal Processing to Audio and Acoustics*, New Paltz, New York, October 1993, IEEE Press.

Acoustics

[25] A. H. Benade, "On the propagation of sound waves in a cylindrical conduit," *Joural of the Acoustical Society of America*, vol. 44, no. 2, pp. 616–623, 1968.

[26] Arthur H. Benade, *Fundamentals of Musical Acoustics*, Dover Publications, Inc, Mineola, New York, 2nd, revised edition edition, 1990.

[27] Jean-Pierre Dalmont, Joë Gilbert, and Sébastien Oliver, "Nonlinear characteristics of single-reed instruments: Quasistatic vlume flow and reed opening measurements," *Journal of the Acoustical Society of America*, vol. 114, no. 4, pp. 2253–2262, October 2003.

[28] Neville H. Fletcher, "Mode locking in nonlinearly excited inharmonic musical oscillators," *Journal of the Acoustical Society of America*, , no. 6, pp. 1566–1569, December 1978.

[29] Neville H. Fletcher, "Autonomous vibration of simple pressure-controlled valves in gas flows," *Jounal of the Acoustical Society of America*, vol. 93, no. 4, pp. 2172–2180, April 1993.

[30] Neville H. Fletcher, "The didjeridu (didgeridoo)," *Acoustics Australia*, vol. 24, pp. 11–15, 1996.

[31] Neville H. Fletcher and Thomas D. Rossing, *The Physics of Musical Instruments*, Springer-Verlag, 1995.

[32] Hermann Helmholtz, *On the Sensations of Tone*, Dover Publications, Inc, New York, New York, 2nd english edition edition, 1954.

[33] D. H. Keefe, "Acoustical wave propagation in cylindrical ducts: Transmission line approximation for isothermal and nonisothermal boundary conditions," *Joural of the Acoustical Society of America*, vol. 75, no. 1, pp. 58–62, Januaray 1984.

[34] D. H. Keefe, "Woodwind air column models," *Joural of the Acoustical Society of America*, vol. 88, no. 1, pp. 35–51, Januaray 1990.

[35] J. Kergomard, S. Ollivier, and J. Gilbert, "Calculation of the spectrum of self-sustained oscillators using a variable truncation method," *Acustica*, vol. 86, pp. 685–703, 2000.

Bioacoustics

[36] Philip M. Morse, *Vibration and Sound*, Acoustical Society of America through the American Institute of Physics, 1981.

[37] Mechteld R. Ballintijn and Carel Ten Cate, "Sound production in the collard dove: A test of the 'whistle' hypothesis," *Journal of Experimental Biology*, vol. 201, pp. 1637–1649, 1998.

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

[39] Henry C. Bennet-Clark and David Young, "A model of the mechanism of sound production in cicadas," *Journal of Experimental Biology*, vol. 173, pp. 123–153, August 1992.

[40] Henry C. Bennet-Clark and David Young, "Short communication: The scaling of song frequency in cicadas," *Journal of Experimental Biology*, vol. 191, pp. 291–294, Frebruary 1994.

[41] Henry C. Bennet-Clark, "Tymbal mechanics and the control of song frequency in the cicada *Cyclochila australasiae*," *Journal of Experimental Biology*, vol. 200, pp. 1681–1694, 1997.

[42] Henry C. Bennet-Clark and A. G. Daws, "Transduction of mechanical energy into sound energy in the cicada *Cyclochila australasiae*," *The Journal of Experimental Biology*, vol. 202, pp. 1803–1817, 1999.

[43] Henry C. Bennet-Clark, "Resonators in insect sound production: How insects produce loud pure-tone songs," *Journal of Experimental Biology*, vol. 202, pp. 3347–3357, June 1999.

[44] Arthur C. Ewing, *Arthropod Bioacoustics*, Cornell University Press, Ithaca, New York, 1989.

[45] M. S. Fee, B. Shraiman, B. Pesaran, and P. P. Mitra, "The role of nonlinear dynamics of the syrinx in the vocalization of a songbird," *Nature*, vol. 395, pp. 67–71, 1998.

[46] Neville H. Fletcher, "Bird song – a quantitative acoustic model," *Journal of Theoretical Biology*, vol. 135, pp. 455–481, 1988.

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

[48] Neville H. Fletcher and A. Tarnopolsky, "Acoustics of the avian vocal tract," *Journal of the Acoustical Society of America*, vol. 105, no. 1, pp. 35–49, January 1999.

[49] Neville H. Fletcher, "A class of chaotic bird calls?," *Journal of the Acoustical Society of America*, vol. 108, no. 2, pp. 821–826, August 2000.

[50] P.J.Fonseca and Henry C. Bennet-Clark, "Asymmetry of tymbal action and structure in a cicada: A possible role in the production of complex songs," *Journal of Experimental Biology*, vol. 201, pp. 717–730, February 1998.

[51] P.J. Fonseca and R.M.Hennig, "Phasic action of the tensor muscle modulates the calling song in cicadas," *Journal of Experimental Biology*, vol. 199, pp. 1535–1544, March 1996.

[52] Tim Gardner, G. Cecchi, M. Magnasco, R. Laje, and Gabriel B. Mindlin, "Simple motor gestures for birdsongs," *Physical Review Letters*, vol. 87, no. 20, pp. 1–4, November 2001.

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

[54] Rosemary Jellis, *Bird Sounds and Their Meaning*, Cornell University Press, Ithaca, New York, 1984.

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

[56] Roderick A. Suthers, "Contributions to birdsong from the left and right sides of the intact syrinx," *Nature*, vol. 347, pp. 473–477, October 1990.

[57] Roderick A. Suthers, "Variable asymmetry and resonance inthe avian vocal tract: A strutural basis for individually distinct vocalizations," *Journal of Comparative Physiology*, vol. 175, pp. 457–466, 1994.

[58] Ingo R. Titze, "The physics of small-amplitude oscillation of the vocal folds," *Journal of the Acoustical Society of America*, vol. 83, no. 4, pp. 1536–1552, April 1988.

[59] Ingo R. Titze, "The voice as a musical instrument: fundamental differences between man-made and biological designs," in *Proceedings of SMAC 03*. Stockholm Music Acoustics Conference, 2003, vol. 1.

[60] David Young, "Do cicadas radiate sound through their ear-drums?," *Journal of Experimental Biology*, vol. 151, pp. 41–56, 1990.

[61] D. Young and Henry C. Bennet-Clark, "The role of the tymbal in cicada sound production," *The Journal of Experimental Biology*, vol. 198, pp. 1001–1019, 1994.

Human-Computer Interaction

[62] D.J.Levitin, S.McAdams, and R.L.Adams, "Control parameters for musical instruments: A foundation for new mapping from gesture to sound," *Organized Sound*, vol. 7, no. 2, pp. 171–189, 2002.

[63] Charles Nichols, "The vbow: Development of a virtual violin bow haptic human-computer interface," in *Proceedings of NIME 2002*, Dublin, Ireland, May 2002, Conference on New Instruments for Musical Expression.

[64] M. Sile O'Modhrain, *Playing by Feel: Incorporating Haptic Feedback into Computer-Based Musical Instruments*, Ph.D. thesis, CCRMA, Stanford University, 2000.

[65] Tamara Smyth and Julius O. Smith, "Creating systained tones with the cicada's rapid buckling mechanism," in *Proceedings of NIME 2002*, Dublin, Ireland, May 2002, Conference on New Instruments for Musical Expression.

[66] Tamara Smyth and Julius O. Smith, "A musical controller based on the cicada's efficient buckling mechanism," *Journal of New Music Research*, December 2003.

[67] J. Rovan, M. Wanderley, S. Dubnov, and P. Depalle, "Instrumental gestural mapping strategies as expressivity determinants in computer music performance," in *Proceedings of the Kansei - The Technology of Emotion Workshop*, Genova, Italy, October 1997.

Signal Processing

[68] J. B. Allen and L. R. Rabiner, "A unified approach to short-time fourier analysis and synthesis," *Proc. IEEE*, vol. 65, no. 11, pp. 1558–1564, November 1977.

[69] Carl M. Bender and Steven A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, Inc, 1978.

[70] James H. McClellan, Ronald W. Schafer, and Mark A. Yoder, *DSP First: A Multimedia Approach*, Prentice Hall, Upper Saddle River, New Jersey, 198.

[71] Andrew Reilly, Gordon Frazer, and Boualem Boashash, "Analytic signal generation – tips and traps," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3241–3245, November 1994.

[72] Louis L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley Publishing Company, Inc., 1991.

[73] Julius Smith, *Techniques for Digital Filter Design and System Identification with Applications to the Violin*, Ph.D. thesis, Stanford University, Stanford, California, June 1983.

[74] Julius O. Smith, "A constant-gain digital resonator tuned by a single coefficient," *Computer Music Journal*, vol. 6, no. 4, pp. 36–40, 1982.

[75] Julius O. Smith and Jonathan S. Abel, "Bark and ERB bilinear transforms," *IEEE Transactions on Speech and Audio Processing*, November 1999.

[76] Julius O. Smith, "Introduction to digital filters," www-ccrma.stanford.edu/~jos/filters/, September 2003, Evolving online book.

[77] Ken Steiglitz, *A Digital Signal Processing Primer with Applications to Digital Audio and Computer Music*, Addison-Wesley Pulishing Company, Inc, Menlo Park, California, 1996.

[78] Udo Zölzer, *Digital Audio Signal Processing*, John Wiley and Sons, Inc., New York, 1999.

Miscellaneous Sources

[79] Jonathan S. Abel, Private correspondence.

[80] Y. Uny Cao, Alex S. Fukunaga, and Andrew B. Kahng, "Cooperative mobile robotics: Antecedents and directions," *Autonomous Robots*, vol. 4, no. 1, pp. 7–23, 1997.

[81] Henry C. Bennet-Clark, Private correspondence.

[82] Neville H. Fletcher, Private correspondence.

[83] *Field Guide to the Birds of North America*, National Geographic, U.S., third edition edition, 1987.

[84] Albert Glinsky and Robert Moog, *Theremin: Ether Music and Espionage*, University of Illinois Press, October 2000.

[85] Lafcadio Hearn, "Insect-musicians," in *Exotics and Retrospectives*, pp. 39–80. University Press, John Wilson and Son, Cambride, USA, 1898.

[86] Michael Hopkin, "Superglue from the sea," *Nature*, January 2004.

[87] Ronald C.Kube and Hong Zhang, "Collective robotics: from social insects to robots," *Adaptive Behaviour, 1994*, vol. 2, no. 2, pp. 189–218, 1992.

[88] "Pd," http://www.pure-date.org.

[89] Noralv Pedersen, "Follow that ant," http://www.ntnu.no/gemini/2001-06E/22-23.htm.

[90] Kendall Powell, "Gecko glue round the corner," *Nature*, August 2002.

[91] Dava Sobel, *Galileo's Daughter: A Historical Memoir of Science, Faith, and Love*, Turtleback Books Distributed by Demco Media, January 2000.

[92] Julius O. Smith, Private correspondence.

[93] "Stk," http://www-ccrma.stanford.edu/software/stk/.

[94] John Whitfield, "Ants shape europe's biggest carnival," *Nature*, August 2002.