

# Hybrid Percussion: Extending Physical Instruments Using Sampled Acoustics

by

Roberto Mario Aimi

S.B., Biology, Massachusetts Institute of Technology (1997)

S.M., Media Arts and Sciences, Massachusetts Institute of Technology (2002)

Submitted to the Program in Media Arts and Sciences,  
School of Architecture and Planning  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2007

© Massachusetts Institute of Technology 2007. All rights reserved.

Author .....  
Program in Media Arts and Sciences  
October 31, 2006

Certified by .....  
Tod Machover  
Professor of Music and Media  
Thesis Supervisor

Accepted by .....  
Andrew Lippman  
Chairman, Department Committee on Graduate Students



# Hybrid Percussion: Extending Physical Instruments Using Sampled Acoustics

by

Roberto Mario Aimi

Submitted to the Program in Media Arts and Sciences,  
School of Architecture and Planning  
on October 31, 2006, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Media Arts and Sciences

## Abstract

This thesis presents a system architecture for creating hybrid digital-acoustic percussion instruments by combining extensions of existing signal processing techniques with specially-designed semi-acoustic physical controllers. This work aims to provide greater realism to digital percussion, gaining much of the richness and understandability of acoustic instruments while preserving the flexibility of digital systems. For this thesis, I have collaborated with percussionists to develop a range of instruments, to refine and extend the algorithmic and physical designs, and to determine successful models of interaction.

Conventional percussion controllers measure and discretize the intensity of strikes into discrete trigger messages, but they also ignore the timbre of the hits and fail to track more ambiguous input. In this work, the continuous acoustic output of a struck physical object is processed to add the resonance of a sampled instrument. This is achieved by employing existing low-latency convolution algorithms which have been extended to give the player control over features such as damping, spectral flattening, nonlinear effects, and pitch.

One of the advantages of this approach is that light taps, scrapes, rubs, or stirring with brushes all take on a hybrid timbre of the real and sampled sound that is surprisingly realistic and controllable. Since part of its behavior is inherently acoustic, a player's intuition about interacting with physical objects can be applied to controlling it. The ability to transform the apparent acoustic properties of objects also suggests applications to HCI and product design contexts.

Thesis Supervisor: Tod Machover  
Title: Professor of Music and Media





**Hybrid Percussion: Extending Physical Instruments Using Sampled  
Acoustics**

by  
Roberto Mario Aimi

**Thesis Readers:**

Thesis Reader .....  
Joseph A. Paradiso  
Associate Professor  
MIT Media Laboratory

Thesis Reader .....  
Hiroshi Ishii  
Associate Professor of Media Arts and Sciences  
MIT Media Laboratory



## Acknowledgments

Thanks to:

my advisor, Tod Machover; my readers, Joe Paradiso and Hiroshi Ishii

percussionists: Rakalam Bob Moses, Jamey Haddad, Curt Newton, Dave Flaherty

Bill Gardner, Gili Weinberg, my friends, family, the Opera of the Future group everyone who had input on the document, the instruments, and the talk.

the Media Lab staff, students, faculty for their help, feedback, and daily inspiration.



# Contents

<b>1</b>	<b>Introduction</b>	<b>17</b>
1.1	Contributions . . . . .	19
1.2	How this approach differs from existing trigger-based electronic percussion systems . . . . .	19
1.3	Musical vision: realistic, physically grounded timbral behavior for digital percussion . . . . .	21
1.4	Structure of this document . . . . .	22
<b>2</b>	<b>Musical mappings, spectral continuity, and percussion.</b>	<b>23</b>
2.1	The mapping problem . . . . .	23
2.1.1	Simpler isn't always better. . . . .	25
2.1.2	An instrument as an extension of the body . . . . .	25
2.2	How is percussion special? . . . . .	26
2.3	Spectral continuity . . . . .	28
<b>3</b>	<b>Background</b>	<b>31</b>
3.1	Historical precedents for electronic/acoustic instruments . . . . .	31
3.1.1	The Stroh violin . . . . .	32
3.1.2	The National Resonator Guitar . . . . .	38
3.1.3	The Rickenbacker electric guitar . . . . .	41
3.1.4	Ondes Martenot . . . . .	42
3.2	Electronic percussion . . . . .	46
3.2.1	A brief history of electronic percussion . . . . .	46
3.2.2	Contemporary commercial approaches . . . . .	47
3.3	Convolution in computer music . . . . .	53

3.3.1	A graphical example of convolution . . . . .	56
3.4	Other approaches: physical modeling and modal synthesis . . . . .	58
<b>4</b>	<b>System design and software implementation</b>	<b>61</b>
4.1	How the system works . . . . .	61
4.2	Initial proof of concept . . . . .	62
4.3	Further implementations . . . . .	62
4.3.1	Pd patch architecture . . . . .	63
4.4	Nonlinear responses . . . . .	65
<b>5</b>	<b>Other expressive controls, extensions to realtime convolution</b>	<b>67</b>
5.1	Damping . . . . .	67
5.1.1	Simple damping model . . . . .	68
5.1.2	Frequency-dependent damping . . . . .	78
5.1.3	Pitch shifting . . . . .	80
5.1.4	Cross fading . . . . .	82
5.1.5	Inverse filtering . . . . .	83
5.1.6	Crashing a ride cymbal: pseudo- nonlinear processing . . . . .	85
5.2	Summary . . . . .	86
<b>6</b>	<b>Physical controllers</b>	<b>89</b>
6.1	Cymbal . . . . .	89
6.1.1	Assembly . . . . .	90
6.1.2	Function . . . . .	91
6.1.3	Brushes . . . . .	92
6.1.4	Wireless brush . . . . .	93
6.1.5	Wired brush . . . . .	93
6.2	Pad . . . . .	94
6.3	Frame drum . . . . .	95
6.3.1	Assembly . . . . .	96
6.3.2	Function . . . . .	97
6.4	Bass drum with speaker . . . . .	98
6.4.1	Assembly . . . . .	98

6.4.2	Function . . . . .	100
6.5	Scratch pad . . . . .	100
6.6	Touching sound . . . . .	100
6.7	Enhanced car console switches . . . . .	101
6.8	A Human-Computer Interface application: augmenting <i>Pico</i> with sound . .	101
6.9	Summary . . . . .	104
<b>7</b>	<b>Evaluation</b>	<b>105</b>
7.1	<i>Timing controllability</i> . . . . .	107
7.1.1	Convolver latency . . . . .	107
7.1.2	Audio subsystem latency . . . . .	108
7.1.3	Total system latency . . . . .	109
7.2	Observations and input from percussionists . . . . .	111
7.3	My own observations . . . . .	115
<b>8</b>	<b>Discussion / future directions</b>	<b>119</b>
8.1	Applications . . . . .	119
8.2	Future work . . . . .	121
8.3	Conclusion . . . . .	122
<b>A</b>	<b>Percussionist biographies</b>	<b>125</b>
<b>B</b>	<b>Offset deconvolution in Octave</b>	<b>129</b>
<b>C</b>	<b>Pin outs</b>	<b>131</b>





# List of Figures

1-1	System architecture . . . . .	18
1-2	Example of typical digital percussion instrument schematic . . . . .	20
1-3	Simplified convolution instrument . . . . .	20
1-4	Re-complexified convolution instrument . . . . .	20
2-1	Generalized controller - sound source schematic . . . . .	24
2-2	Production and perception feedback loop . . . . .	29
2-3	Pantograph . . . . .	29
3-1	Images from Stroh US patent . . . . .	34
3-2	Mechanism detail . . . . .	35
3-3	Charles D’Almaine . . . . .	35
3-4	Edison phonograph recording session . . . . .	36
3-5	Contemporary uses of the Stroh violin . . . . .	37
3-6	Vioara cu goarnă stamp [60] . . . . .	37
3-7	National tricone resonator guitar . . . . .	38
3-8	Tricone patent . . . . .	39
3-9	Image from Beauchamp pickup patent . . . . .	41
3-10	“Frying pan” design . . . . .	42
3-11	Ondes Martenot with diffusers . . . . .	42
3-12	Diffusers of the Ondes Martenot . . . . .	44
3-13	Moog Percussion Controller 1130 . . . . .	46
3-14	PAiA Programmable Drum Set . . . . .	46
3-15	Simmons SDS5 drumset . . . . .	47
3-16	Roland V-Pro TD 20S [73] . . . . .	48

3-17 Roland position sensing system . . . . .	49
3-18 Roland V-cymbal . . . . .	50
3-19 Roland Handsonic HPD-15 . . . . .	50
3-20 Pressure sensor and piezo . . . . .	51
3-21 Korg Wavedrum . . . . .	51
3-23 Wavedrum sensor . . . . .	52
3-22 Korg Wavedrum . . . . .	52
3-24 Rim assembly . . . . .	52
3-25 Rim . . . . .	52
3-26 Time smearing . . . . .	55
3-27 An impulse . . . . .	56
3-28 The room impulse response . . . . .	56
3-29 Two impulses . . . . .	57
3-30 Two impulses convolved with the room response . . . . .	57
3-31 A plucked string sample . . . . .	57
3-32 A plucked string sample convolved with the room response . . . . .	57
3-33 Multiplying in the frequency domain . . . . .	58
4-1 Block diagram of the basic system . . . . .	62
4-2 Minimal system schematic . . . . .	64
4-3 Latency of partitions corresponds to partition offset . . . . .	64
4-4 Three regimes of cymbal response . . . . .	65
4-5 Power spectrum of a practice pad being hit with a drumstick. . . . .	66
5-1 Dashpot . . . . .	68
5-2 Exponential decay ( $\lambda=0.006$ ) . . . . .	68
5-3 Superimposing the desired decay onto the stored response . . . . .	69
5-4 Superimposing the desired decay onto the stored response . . . . .	69
5-5 Exponential decay superimposed over varying block sizes . . . . .	70
5-6 Exponential decay plotted vs partition number . . . . .	71
5-7 Original and target decay envelopes . . . . .	72
5-8 Ideal damping, and less-than-ideal damping. . . . .	72
5-9 Ideal vs linear and quadratic cross fades . . . . .	73

5-10	What happens when you stop damping? . . . . .	74
5-11	Schematic with input and output gains . . . . .	74
5-12	“Double ramp” effect due to the FFT delay . . . . .	75
5-13	10 strikes per second: artifact for long stored impulses when $G_i = G_o$ . . .	75
5-14	Artifact when $G_i = G_t, G_o = \min(G_t, G_i(\text{delayed}))$ . . . . .	76
5-15	Fixing one artifact introduces another. . . . .	77
5-16	Slowing down the muting rate limits or removes the artifact for fast hits . .	78
5-17	Frequency-dependent damping . . . . .	79
5-18	Aligning the -3dB points of a one-pole filter and an exponential . . . . .	79
5-19	Schematic with input and output filters . . . . .	80
5-20	Tape-style pitch shift vs. fixed-frequency shift . . . . .	81
5-21	Schematic for crossfading two convolvers in series . . . . .	82
5-22	Deconvolving a typical hit from a frame drum sample . . . . .	84
5-23	Spikes at $0 \text{ FS}/2, 0.21 \text{ FS}/2$ . . . . .	84
5-24	Deconvolving a typical hit from a frame drum sample . . . . .	85
5-25	Crashable cymbal system diagram . . . . .	86
5-26	Spectra of soft and loud hit . . . . .	86
6-1	Cymbal controller . . . . .	90
6-2	Cymbal controller assembly . . . . .	91
6-3	Using the audio interface to measure the resistance of the FSR . . . . .	91
6-4	Cymbal pitch controlled with knob . . . . .	92
6-5	Wireless brush . . . . .	93
6-6	Wired brush . . . . .	93
6-7	Percussion pad controller . . . . .	94
6-8	Percussion pad cross section . . . . .	94
6-9	Frame drum controller . . . . .	95
6-10	Frame drum controller assembly . . . . .	96
6-11	Frame drum controller . . . . .	97
6-12	Bass drum controller . . . . .	98
6-13	Bass drum controller . . . . .	99
6-14	“Touching sound” . . . . .	100

6-15	James Patten’s <i>Pico</i> system with sound augmentation . . . . .	102
6-16	Contact microphone placement under actuation table . . . . .	103
7-1	Schematic for measuring latency within the convolver . . . . .	107
7-2	64 samples (1.45 ms) of latency between the left and right channel . . . . .	108
7-3	Schematic for measuring latency of the audio system, minus the convolver . . . . .	109
7-4	Schematic for measuring latency of the entire system . . . . .	110
7-5	Table of percussionists . . . . .	111
7-6	“Crazy” mallets . . . . .	112
C-1	Connector and cable pinouts . . . . .	131

# Chapter 1

## Introduction

Since well before recorded history, mankind has been making music and developing musical instruments. While music is not essential for survival, most of us are compelled to listen to and make it, and even spend considerable time and money on it. Music, and the tools for making it are important.

Much of the historical development of musical instruments could be characterized as a quest for new sounds to differentiate the player or composer, and to break the audience from the lull of familiarity. However, with the advent of synthesizers and digital samplers, literally any sound that could be recorded or rendered, captured or processed, could be played back by these instruments. Suddenly the technical ability to produce all possible sounds had surpassed our ability to make use of them.

Of course, the technical capability of producing any sound is not the only limitation. Much as a blank notebook and pencil can technically produce any novel, there is a lot of thought and creative judgments that have to be made to make the right sounds in the right way. With the technical roadblock largely removed, it becomes clear that the quest for new instruments was not just a desire for new sounds, but new tools for musical thinking, and new ways of thinking about sound.

This thesis proposes one new way to think about designing digital percussion instruments that can extend the capabilities of the player while maintaining much of the coherence

and understandability we find in acoustic instruments. The same techniques can also be extended to non-musical applications of Human-Computer Interaction (HCI) and design.

The goal of this work has been to make more realistic digital instruments, not so much in terms of emulating timbres, which the digital sampler has perfected, but in terms of realistic behavior that is similar to that of real objects. This thesis seeks a middle ground between the controllability of pure acoustic instruments and physical models, and the extensibility of the digital sampler.

This is achieved by joining realtime convolution algorithms with semi-acoustic physical objects, sensors, and mappings to change the apparent acoustics of the objects. These algorithms are well known in computer music, but have not yet been applied to creating realtime percussion instruments. This technique can either be viewed as pulling part of a synthesis algorithm out of the computer and into real world objects, or using computation as a way to extend the acoustics of those objects.

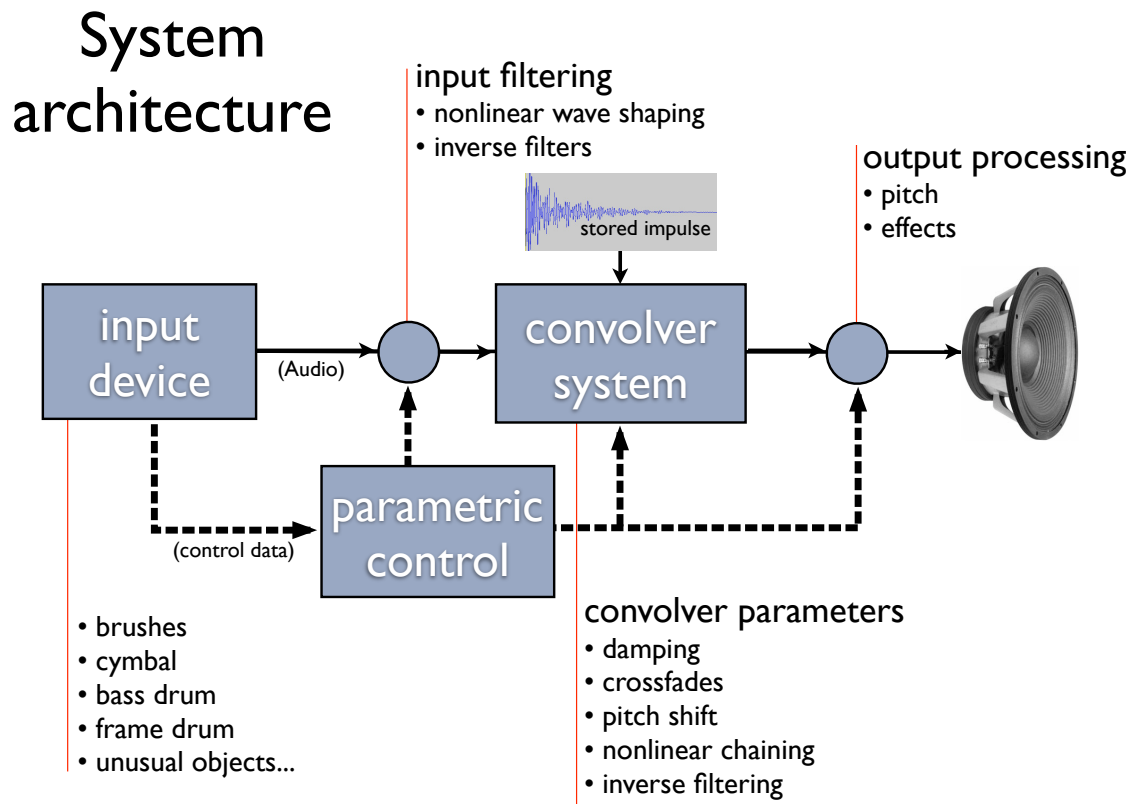


Figure 1-1: System architecture

## 1.1 Contributions

Specific contributions of this thesis are:

- A novel system architecture that allows players to apply their intuitions and expectations about real acoustic objects to new percussion instruments that are grounded in real acoustics, but can extend beyond what is possible in the purely physical domain.
- Extensions to the functionality of convolution algorithms to accommodate muting, pitch shifts, approximation of nonlinear effects, and inverse filtering.
- A range of semi-acoustic physical controllers designed to integrate with the system architecture and that illustrate design principles for future instruments.
- An implementation of these algorithms that can serve as a platform for future development and allow customization to meet future creative goals.
- Applications to the areas of human-computer interface and product design, exploring apparent acoustic properties as a design parameter and for information display.

## 1.2 How this approach differs from existing trigger-based electronic percussion systems

Conventional percussion controllers measure and discretize the intensity of strikes into trigger messages that specify how loud the output should be (Figure 1-2). However, they also lose the timbre of the hits and fail to track inputs that do not result in clear peaks. This thesis work can be described as a simplification of the typical percussion triggering model. Comparing figures 1-2 and 1-3 shows that the convolution method employed in this work removes several intermediate steps present in typical percussion instruments and makes no effort to represent the hit as a unique abstract event. Processing is continuous: since it processes the raw acoustic output of the physical object, all timbral variation achievable with that physical object is represented in the output signal.

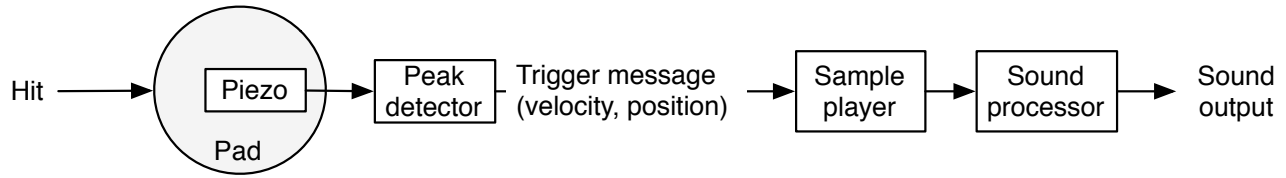


Figure 1-2: Example of typical digital percussion instrument schematic

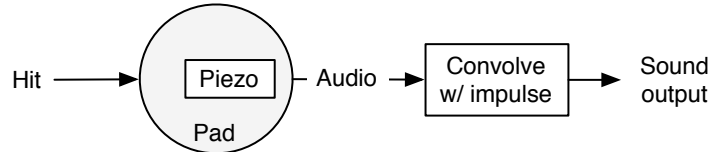


Figure 1-3: Simplified convolution instrument

The simplified design of figure 1-3 can also be extended to include sensors and a mapping layer that controls the processing of the audio (figure 1-4). Part of the challenge then is to identify meaningful ways to control the convolution algorithm that are consistent with the physical design while still expanding the musical options available to the player.

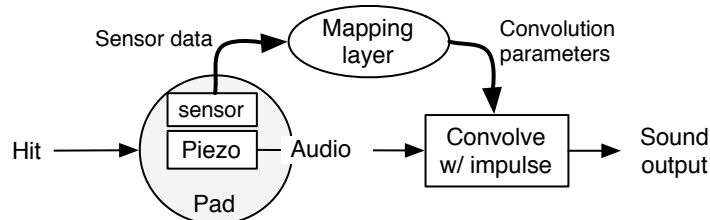


Figure 1-4: Re-complexified convolution instrument

Design of the physical interface and the corresponding digital processing differs from typical percussion controller design because the acoustic properties of the controller are a critical component of the sound. There is an inherent tradeoff between the generality of the controller, and its suitability for a particular task. For example, a weighted keyboard provides better tactile feedback for piano sounds, but not for organ tones. The systems proposed in this thesis are substantially less general-purpose than any percussion trigger unit, striking a middle ground between the nearly infinite reconfigurability of triggers and MIDI, versus the specialization but richer control of acoustic instruments.



### 1.3 Musical vision: realistic, physically grounded timbral behavior for digital percussion

The core musical vision for this work is to make a system that gets out of the way of great players and lets them do what they do best: explore the range of sounds possible, expand the timbres available, find new and surprising ways to play, and get good at playing it. All through acoustic interaction with physical objects.

**Explorability** Players naturally gravitate towards exploiting the more expressive aspects of an instrument. However, when playing digital drum sets, since the timbres are essentially fixed, players must focus instead on timing and accents. The vision for this work is to make systems that reward timbral exploration as well, even as the actual timbres depart from exact emulations of physical sounds. I want the intentionally non-physical percussion sounds of an FM synthesizer to be as playable and richly timbrally explorable as a real ride cymbal. Additional controls, knobs, pedals, are all fine for dialing in changes to the sound, but the fundamental percussion interaction of sound being produced in response to physical contact with an object is (I think) essential for maintaining this feature of timbral control and explorability.

**Expandability** Since the stored impulses are just sound files, they are very easily interchanged, processed, cross faded. Much like a digital sampler, those processes are almost unbounded, and easily experimented with, while the method of interaction and the physical controller remain grounded in the constraints of the physical world. What this means for a drummer is, for example, the ability to dial in different cymbals and “cymbal-esque” sounds that each behave realistically even as they sonically diverge from what is physically possible.

Found surfaces can be incorporated that enable other playing gestures. Wooden chairs, bristle brushes, cheese graters are all fair game. Specialized controllers can be designed to heighten their feel and physical response because they no longer have the sole responsibility for producing and shaping the timbre.

**Audio-driven** This work also represents a return to treating audio as a control signal, something that was inherent in the early modular synthesizers, but that was lost when MIDI essentially separated control data from audio signal. Only now has processing power reached a point where digital signals can have a high enough sampling rate that they are essentially continuous for the frequencies of interest.

## 1.4 Structure of this document

The following chapter will discuss some of the challenges of musical mappings, how it relates to the controllability and learnability of an instrument, and how these apply to percussion in particular. Chapter 3 discusses further related work, covering electronic percussion instruments, the current uses of convolution in computer music, and other related synthesis and processing approaches. Chapter 4 discusses the implementation of the software system and chapter 5 covers the physical controllers. Chapter 6 describes the evaluations of the work, and chapter 7 describes conclusions and future work.

## Chapter 2

# Musical mappings, spectral continuity, and percussion.

In this chapter, I will describe some of the ongoing challenges related to musical mappings as they apply to digital musical instruments, and to the specific problem domain of percussion.

### 2.1 The mapping problem

In acoustic instruments, the resonator is physically connected to the part of the instrument that the player touches and controls. In some keyboard instruments though, there is a mechanical linkage between the keyboard and the strings. This enabled the ergonomics and layout of the keys to be designed for ease of playing and representation (for example all octaves are the same distance apart, so players can shift registers while using the same technique in a different location on the keyboard). Probably the most extreme example of this mechanical coupling is in pipe organs, where the keys control valves that supply air to organ pipes that can be located many feet from the player.

But these linkage also limited the degrees of freedom that were controllable by the player. For the piano, one could specify an entire performance knowing only which note was hit, when it was hit, and how hard. Player pianos were able to play back such recorded performances punched on paper rolls, and the human performances could be abstracted

from any particular instrument.

With the development of the MIDI protocol [47] in the early 1980s, it became possible to literally and conceptually disconnect the physical part of the instrument that the player touched from the part that made the sound. The sensible place to make this break was at the mechanical linkage. Keyboards and tone generators could then be interchangeable.

Once the protocol was established, it was quickly discovered that controllers did not need to resemble keyboards, and a flurry of development of unusual controllers followed. It also became clear that the data stream from the controller could be easily re-mapped by putting a computer between the controller and the tone generator (represented by the mapping system box in figure 2-1).

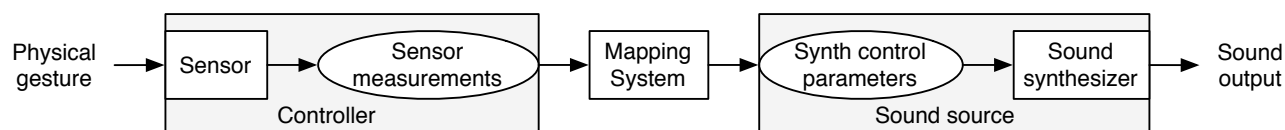


Figure 2-1: Example of typical generalized controller - sound source system schematic. Adapted from [99]

This approach had the promise of connecting any input behavior to any output behavior: any gesture could create any sound, limited only by the imagination of the person doing the mapping.

But making mappings that approached the complexity and continuity of acoustic instruments proved difficult. In acoustic instruments without mechanical abstraction, the relationship between the player input and the sound is quite complex, but since that relationship is constrained by physics, players can apply and further develop their intuition about the relationship between playing gesture and sound. The physical constraints also force the mapping between gesture and sound to be inherently continuous.

When the mapping occurs in the digital domain it is not just potentially complex, but the potential complexity is unbounded by the rules governing a physical object. The player doesn't know what rules that mapper has put into the instrument, or even the nature of those rules, and cannot readily "reverse engineer" those rules without searching the entire parameter space.

A great burden then rests on the shoulders of the person designing the mapping to anticipate all of the ways a player might use the instrument, and to make sure that the gesture parameters interact in learnable and consistent ways with the sound parameters. As the number of sensors increases, this task only gets harder.

### **2.1.1 Simpler isn't always better.**

One approach to the challenges of mapping is to create simple relationships between gesture and response such as a 1:1 mapping of a bend sensor output to the pitch of a sound. While this helps provide constraints to the instrument designer, there are some arguments suggesting that simple 1:1 mappings are not necessarily easier to understand than rich interdependent mappings provided that there is sufficient continuity.

Hunt et al [33],[34] make this case through a series of experiments. Even if the physical interface and sound sources are kept the same, subtle changes in mapping can radically alter the player's impression of the instrument. They show that it was easier for subjects to copy a complex musical task (after practice) using interfaces with "multi-parametric cross mappings" more like those found in acoustic instruments than for simple 1:1 mappings, even if the cross mappings were not exact emulations of those found in real instruments. This suggests that players are quite capable of making sense of complex mappings, possibly because of their extensive experience with such systems in the real world.

Rovan [76] modified the simple 1:1 mapping of a commercial Yamaha Wind controller to create more complex interactions between the sensor parameters, similar to what is found in real reed instruments. Experienced players preferred these more complex mappings. In these examples, while the mappings were complex, they were repeatable, and provided consistent feedback to the player. It is this consistency of feedback that is an important part of how tools are learned and even integrated into our sense of body.

### **2.1.2 An instrument as an extension of the body**

An instrument, like any tool, can become an extension to the body schema if the sensorimotor feedback loop (auditory, visual, tactile) is sufficiently tight. Neuroscientist Jacques

Paillard describes this phenomenon:

Consider here the interesting phenomenon, studied in man, of the assimilation of a stick, a tool, a prosthesis, and even a car in the 'body schema' (Paillard 1971) to the extent that these objects are literally incorporated in body space (e.g. the tactile sensitivity of the tip of the blind person's stick). It has been shown that this incorporation is effective only if the subject actively experiences the 'prosthesis' by associating his own movements with the sensory impressions to which they give rise: in other words, closing the sensorimotor loop which, in intentional movement, adjusts the motor command in anticipation of the sensory consequences of its execution. [54]

Any discontinuities in the feedback loop work against this extension of body schema. Percussion instruments in particular provide especially tight sensimotor feedback due to direct contact with the resonator.

## 2.2 How is percussion special?

This thesis is primarily focused on the specific domain of digital percussion instruments.

Percussion instruments are unique among the families of instruments in that there is no clear distinction between any object and an instrument. Every object has an acoustic response to being handled, scraped, or hit. This isn't true of other instrument families; a coffee table can never be played like a brass instrument, but it can always become a percussion instrument.

Perhaps because of this continuum of object to instrument, when unusual items are used in the orchestra, they usually end up in the percussion section. Percussionists are not just trained as specialists of one instrument, but are expected to be proficient at the range of common and uncommon instruments [27]. Their core skill may be rhythm, but understanding how to coax interesting timbres from a wide variety of objects is also central. Exploration of a range of timbres is a fundamental component of percussion and is repeated each time the percussionist encounters a new object. Such exploration is important to develop mastery. Anyone who has seen Max Roach's "Mr hi hat" routine (in which he plays a hi hat as a solo instrument, hitting nearly every surface of it) can recognize his skill at obtaining an incredible range of sounds from a single object.

Although the physical materials and construction of all acoustic instruments are closely related to how they sound, percussion instruments are unique in that the playing impulse is transferred directly to the resonator (often comprising the entire instrument) rather than through a complex driving system such as a vibrating string or reed. Because of this, the shape and material of the resonator has significant influence over the timbre of the instrument [74]. In this thesis work, the physical resonator can be augmented or even replaced by a digital resonator derived from another sampled source.

**Timbre** Since much of the role of percussion is rhythmic, it can be easy to lose sight of the importance of timbre. The drum set, with its range of discrete instruments: snare, toms, bass, cymbals, and hi hat suggests discrete timbres, clearly a cymbal sounds very different from a snare. And in many contexts, the focus of the drummer is on using these default timbres as a basic palette to drive the rhythm forward. A ride pattern can be accented with the snare and crash, and there is tremendous potential for expression in control of timing, accents, and phrasing.

But within each part of the drum set, there is a fantastic range of timbral nuance possible as well. In hand drumming, this is perhaps more apparent. The Djembe, for example, has three primary timbres: bass, tone, and slap. Since these three are achieved from the same physical instrument, it is more apparent that there can be in-between sounds as well. Even though a beginner must practice to keep the three tones clear and distinct, variation into that middle timbral space can be a very expressive domain for a master drummer. Each element of the drum set has that same degree of timbral potential as well.

Everett “Vic” Firth, long-time principal timpanist of the Boston Symphony Orchestra and drumstick maker describes the depth of a snare drum in his introductory method text:

As well as control and fast articulate hands, the snare drummer must develop a sense of rhythmic phrasing, interpretation, and concept of sound. He must understand tone production and sound projection as related to sticking and technical execution. He must not treat the instrument as a noisy rhythm maker, but as a musical instrument capable of countless musical subtleties [24].

**Commercial drum kits** But despite the merits of timbral variation, the predominant feature of most percussion music is rhythm and accent. These two elements are well represented in the state of the art commercial electronic drum kits such as those offered by Yamaha and Roland [73]. For many styles of music, those features are sufficient, and by having multiple pads each with a different sound, the range of the timbral palette of a real drum set is covered. Since many drummers playing on real drum sets use each drum to produce only one or two timbres, the Yamaha/Roland approach goes far to meet their needs, but what is missing is the ability to control the nuance of timbre by varying the impulse.

Digital drumsets have seen wide adoption in the commercial market, showing that there is a real demand for such devices among the general public. These commercial digital percussion instruments also represent a very high degree of iterative incremental refinement of their triggering and sound generation systems while maintaining the same basic system design.

Percussion instruments (real ones) represent extreme examples of the immediacy and understandability inherent in acoustic instruments. It is frustrating then, that the highly-refined triggering systems common in electronic percussion introduce discontinuities between the playing gesture and the sound.

## 2.3 Spectral continuity

As discussed earlier, when a typical digital drum pad is struck, the vibration is captured by a sensor and the peak is detected. A discrete sampled or synthesized drum sound is played back at a corresponding amplitude to the height of the peak.

While this system works very well for clear stick strikes, it lacks the nuance of a real drum: it doesn't allow the player to rub the drum with brushes, and it misses the smaller scrapes and hits that comprise real drum playing. None of the spectral information from the hit is represented in the output; hitting a drum pad with a foam mallet would sound the same as hitting with a hard stick.

In a real drum, hitting with something hard or something soft would make two different



sounds, and there is a continuum of possible output between those two sounds running from bright to dark. Knowing what two different hits sound like immediately informs the player about what to expect when he plays something in between. More hits help the player to further chart the sonic territory, but because of the continuity of response, the player does not need to test the entire parameter space. Though the acoustic drum builder can control many aspects of the sound, the continuity of possible sounds and gestures is nearly unavoidable, and any discontinuity (engaging a snare, for example) is quite deliberate.

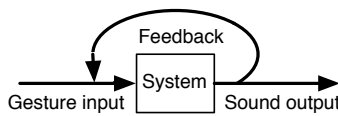


Figure 2-2: Production and perception feedback loop

One way to get this continuity in a digital instrument is to make the sound of the physical controller object important. Every physical interaction with an object makes sound that can be captured and processed. Hitting the object with a stick or the palm of your hand creates not only different intensities of sound, but radically different sound spectra. A small tim-

bral space of the physical object can be expanded and transformed, but what is important is that the system is continuous and predictable. This continuity helps the player make corrections to the input gesture based on what is heard in the output sound to converge on what the player intends (figure 2-2).



Figure 2-3: Pantograph

Playing such an instrument is a lot like freehand drawing with a pantograph (figure 2-3); you can be looking at the larger output while controlling the small pointer. Looking at one gives reasonable feedback over the control of the other because the transform preserves the relationship between points.

One example of a instrument that has its timbral space transformed would be the electric guitar. The player is playing (and listening to, correcting) the system of guitar, pedals, amplifier, speaker – even the room acoustics and feedback, not just the wood and strings. The instrument isn't just the physical guitar, It's likely that the player would play quite differently if unplugged because an electric guitar without the rest of the system is a different instrument.

In this thesis I have been working to develop ways of designing digital percussion instruments that have much of the predictability, spectral continuity, and understandability

present in acoustic percussion instruments, but with the advantages of digital techniques. In the next chapter, I will discuss some background work relevant to the approach I have taken.

## Chapter 3

# Background

This chapter presents some of the related work on instruments and signal processing. In the first section, four historical instruments that span the gap between purely acoustic and electronic instruments are considered as possible examples for future work. The next section is a discussion of electronic percussion with a focus on three contemporary instruments that most strongly relate to this thesis work, the Roland V-drums and Handsonic, and the Korg Wavedrum. The third section looks at the role of convolution in computer music, and the fourth is a brief discussion of physical modeling and modal synthesis.

### 3.1 Historical precedents for electronic/acoustic instruments

The beginning of the 20th century, prior to the Second World War, was a period of incredible innovation in musical instruments that was not equaled until the popularization of the electronic synthesizer in the 1960s. Composers and musicians were looking for new sounds, and performers needed new ways to project their sound to larger audiences. This motivated several unusual instrument designs. Traditional instruments were extended to create new and louder timbres, while advances in electronics allowed the entirely new categories of electro-acoustic and electronic instruments to develop [83].

This section looks at the designs of four prewar instruments; two purely acoustic, two electric/electronic: the Stroh violin, the National resonator guitar, the Rickenbacker electric

guitar, and the Ondes Martenot. These four instruments are unusual in the degree that they treat the resonator as a distinct element that can be modified in the instrument design process to affect the volume and timbre. But since they are from a time before high fidelity amplification, the resonators can not be treated as pure abstractions, there is still an intimate connection between the instrument and its amplification system. This tension between abstraction and integration is part of what makes these instruments so interesting.

The first electronic amplifiers were not powerful, so speaker systems needed to be very efficient, usually at the expense of fidelity. However, with careful design, it was possible to create instrument-amplification systems that had compelling timbres despite limited fidelity. This was especially true for early electronic instruments, which often produced thin or harsh-sounding direct signals. The Ondes Martenot, for example, had special wooden speaker cabinets fitted with sympathetic strings to give extra resonance to its thin electronic sound. The amplification systems were not general-purpose high-fidelity P.A. systems; instead, they were designed for the particular instrument, significantly reshaping the timbre enough to be considered an essential part of that instruments timbral identity.

Even though these designs were shaped by technical limitations that are no longer in place, they represent examples of a middle ground between purely acoustic and purely electric or electronic instruments, maintaining a grounding in the physical world and providing a template for how one might think about making digital acoustic hybrid instruments now and in the future.

### **3.1.1 The Stroh violin**

The story of the Stroh violin is intimately connected to the early development of recorded music. Inspired by the phonographic methods of sound playback, it also became an essential instrument in acoustic phonograph recording sessions.

#### **Phonograph recording**

The enormous commercial success of Edisons phonograph, invented in 1877, led to the establishment of a thriving industry to meet the demands of the growing legions of phonograph

owners for recorded music [96].

Recordings in this period (from 1877 until 1925) were purely acoustic. The recording system was essentially the reciprocal of the playback system: Sound was directed by a horn to a diaphragm which was connected to a cutting stylus that would carve on a blank wax cylinder, which replaced the tin foil cylinders in Edisons original invention [29].

Players had to crowd around a recording horn and play quite loudly to be heard. This was one reason for the popularity of louder instruments like the banjo and horns in early recordings.

The violin was at a particular disadvantage due to its relatively bright sound. The frequency response of the phonograph was limited to about 3000 Hz, so instruments like the violin with significant content above 3000 Hz were particularly difficult to record. Given the popularity of violin music and the limitations of phonograph technology, there was a need for a violin that could be more easily recorded [14].

## **J. M. Augustus Stroh**

John Matthias Augustus Stroh was a prolific inventor. Born in Germany in 1828, he emigrated to England where he worked as a watchmaker from 1857-1861. Although he made many contributions to the field of watch making, he also applied his mechanical skill to a broad range of other challenges [14].

With Sir Charles Wheatstone (best known for his Wheatstone bridge circuit for measuring electrical resistance), Stroh developed and manufactured an improved high-speed telegraph. The two also created an accordion that could slide between pitches. Stroh became widely known for his mechanical skill, and received a Gold Medal from the International Jury of the Paris exhibition of 1878 for his telegraph.

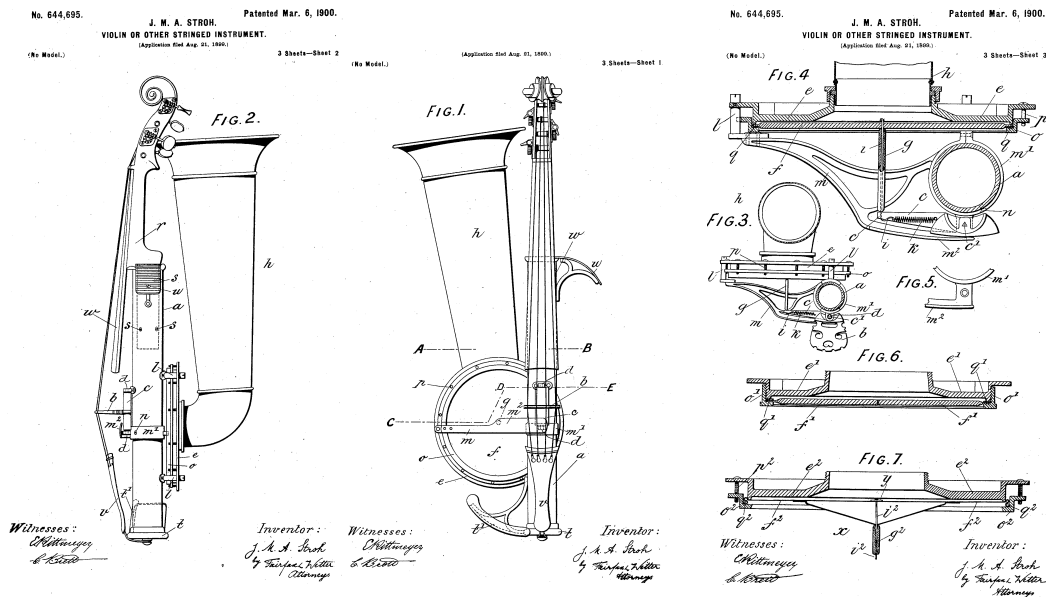
First word of Edisons phonograph reached England in a London Times newspaper article on January 18th, 1878. Stroh began work immediately, and on February 1st (less than two weeks); he presented what was essentially a working copy of Edisons design. By February 27th, Stroh had extended and improved Edisons design through the addition of

a fan governor, flywheel, and clockwork motor mechanism to decrease the speed variations intrinsic to the hand-cranked design [61]. Strohs phonograph design was manufactured by the London Stereoscopic Company. Stroh retired from the phonograph business in 1880, but continued to pursue many interests, including making his own cameras with a special color filter system and continuing to enhance the phonograph.

Through his experience with recording technology, Stroh was aware of the need for a violin that could be loud enough for recordings, and he began working to apply phonograph technology to the violin. Stroh received his first patent for his violin design in 1899 (UK patent 9418) and a patent for the conical diaphragm (UK patent 3393/1901) in 1901. His US patent for the violin was issued in 1900 (number 644,695). His son, Augustus Charles Stroh began manufacturing his fathers design in London in 1901 [14].

## Design and adoption

The Stroh violin (figure 3-1(a)) features some clever mechanical designs. The bridge of the Stroh violin pivots at its base, and the pivoting motion is carried by a lever to the conical aluminum diaphragm (figure 3-1(b), label y) by way of a thin connecting rod (figure 3-2, label g). The neck and fingerboard are essentially the same as in a traditional violin.



(a) Stroh violin patent

(b) Cone and rod mechanism

Figure 3-1: Images from Stroh US patent [37]

The diaphragm opens into a large aluminum horn (figure 3-1(a) label a). Some models featured an additional smaller horn pointed at the players ear to help him hear himself in loud recording sessions.

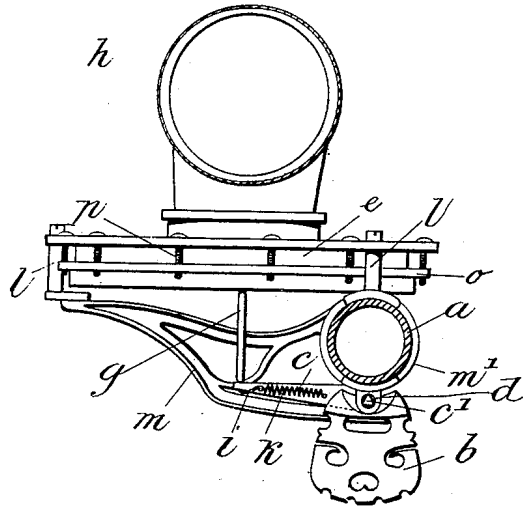


Figure 3-2: Mechanism detail

The Stroh violin became widely used in phonograph recordings, holding its own against horns and banjos. The first known recording of the Stroh violin was by Charles D'Almaine, shown in a recording session in figure 3-3 playing a conventional violin [57].

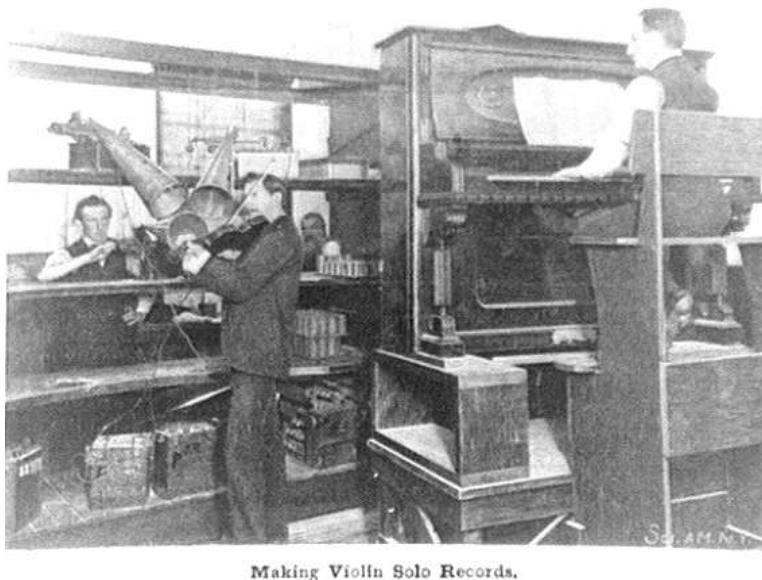


Figure 3-3: Charles D'Almaine (photo courtesy of the collection of Glenn Sage, Portland, Oregon <http://www.tinfoil.com/>)

Figure 3-4 shows a recording session of the singer Harry Anthony. Stroh violins are

visible in the band along with the usual brass instruments.



Figure 3-4: Harry Anthony recording session at the Edison recording studio. Note the Stroh violins in the background [22]

Under the leadership of George Evans and Co, the line of Stroh instruments was expanded to include cello, bass, mandolin, ukulele, and guitar, though the violin remained the most popular model.

### Electrical phonograph recording

On February 25, 1925, Columbia records recorded pianist Art Gillham using a new system of condenser microphone, tube amplifier, and electrically actuated recording stylus developed at AT&T's Bell labs. The new system had approximately double the frequency response, and the microphones were able to pick up much quieter sounds. Musicians were able to gather around a microphone in the center of a room, rather than crowding around the recording horn. Almost immediately, recording studios returned to traditional violins. Even the acoustic holdout Edison switched to electrical recording in 1928, before going out of business in 1929. The Stroh violin that had been a mainstay of the recording studio was now out of its main line of work.

All along, a smaller fraction of Stroh instruments had also been used in live performances, where there was also a need for greater output, and stage PA systems of the time were



inadequate and rarely present. These became the core of Strohs business. By the early forties, interest in the violin as an instrument in popular music waned, diminishing the market for Stroh violins even further. Production of the Stroh instruments ended in 1942 .

### Continued use and manufacture of Stroh-style instruments in Romania

Although production of Stroh instruments stopped in 1942, a Stroh-style violin called the Teibel became popular across Romania for traditional Romanian music [46].

Known as the vioara cu goarnă, or literally violin with horn, it often displaced the traditional vioara dulce, or soft violin. Romanian instrument makers began manufacturing their own variations on the design, usually featuring a trumpet or bugle horn. Once popular throughout Romania, the vioara cu goarnă is now primarily found in the Bihor region of northwest Transylvania.

Like the Stroh violin, the vioara cu goarnă is significantly louder than the conventional violin, but has a more pronounced, nasal quality than the Stroh. It is played both by amateurs and professional popular musicians, especially at fairs and outdoor events, weddings, or baptisms (figure 3-5(a)).



(a) player in the village of Berzunti in Moldavia.  
[63]



(b) recent example,  
made in the Bihor  
region in 1998 [63]

Figure 3-5: Contemporary uses of the Stroh violin

The Romanian post office issued a series of commemorative stamps in 2003 celebrating the Romanian musical heritage. Among them was a stamp featuring the vioara cu goarnă.



Figure 3-6: Vioara  
cu goarnă stamp [60]

The Stroh instrument design was both enabled by phonograph technology and necessitated by the limitations of phonographic recording, so it is remarkable that its use has outlasted the original reason and means for its existence. It is also an important step in the path to what would become the modern electric guitar.

### 3.1.2 The National Resonator Guitar

Originally designed for Jazz and Hawaiian music styles, National guitars had a significant impact on the developing sound of acoustic blues music in the 1920s and 30s. The story of National guitars also literally connects the Stroh Violin to the Rickenbacker electric guitar that followed.

#### George Beauchamp and John Dopyera

George Beauchamp was a young guitar player working on the Vaudeville circuit in the mid 1920s playing Hawaiian-style slide guitar. Like many players, he wanted a louder guitar that could be heard in crowded venues. At this time, the guitar was emerging as a popular instrument and the use of the louder banjo was in decline. The guitar was also beginning to be used in jazz ensembles as a solo instrument, rather than simply as part of the rhythm section.

Around 1925, he began searching for a way to make a louder guitar. Having seen a Stroh violin, Beauchamp wanted someone to build him a Hawaiian guitar based on the same principle, unaware that Stroh had already manufactured a similar guitar. Beauchamp sought out several instrument makers to meet his requirements. He eventually chose the Dopyera brothers, whose workshop was close to Beauchamp's Los Angeles home [11].



Figure 3-7:  
National  
tricone res-  
onator guitar  
[50]

Though their first effort was unsuccessful, Beauchamp continued to seek a louder guitar. He was intrigued by the mica disc diaphragm of the phonograph, and thought that it might be applied to a guitar design and again enlisted John Dopyera to pursue the idea. Dopyera experimented with several materials and shapes, eventually settling on a resonator design

that used thin conical aluminum discs instead of mica [11].

The strings of the guitar rested on a wooden saddle attached to a cast aluminum bridge that connected to the center of the cones. After trying several configurations, Dopyera chose a three-cone system with a T-shaped bridge that connected to all three cones (figure 3-8).

Beauchamp solicited investment and founded the National String Instrument Company company which hired Dopyera as Factory superintendent, and in 1927, they began manufacturing Spanish and Hawaiian style Tri-cone guitars. As business expanded, production of metal bodies and cones was moved to the nearby Rickenbacker Tool and Die company. The owner, Adolph Rickenbacher invested in National, and National even gave him the title of engineer in a 1930 catalog.

Personal and creative conflicts emerged between Beauchamp and Dopyera, coming to a head when Beauchamp sought a patent for a single-cone resonator guitar, a design that Dopyera had considered but discarded. Dopyera resigned and gave up his shares of National.

The National single-cone guitar could be produced much more cheaply, and it found its way into the hands of many blues musicians, becoming an essential part of the emerging acoustic blues music.

Dopyera also went on to create a new single-cone guitar with his brother, and formed the Dobro Corporation (short for Dopyera Brothers, dobro also means good in Czech). The Dobro became popular in Bluegrass music for its bright and twangy sound.

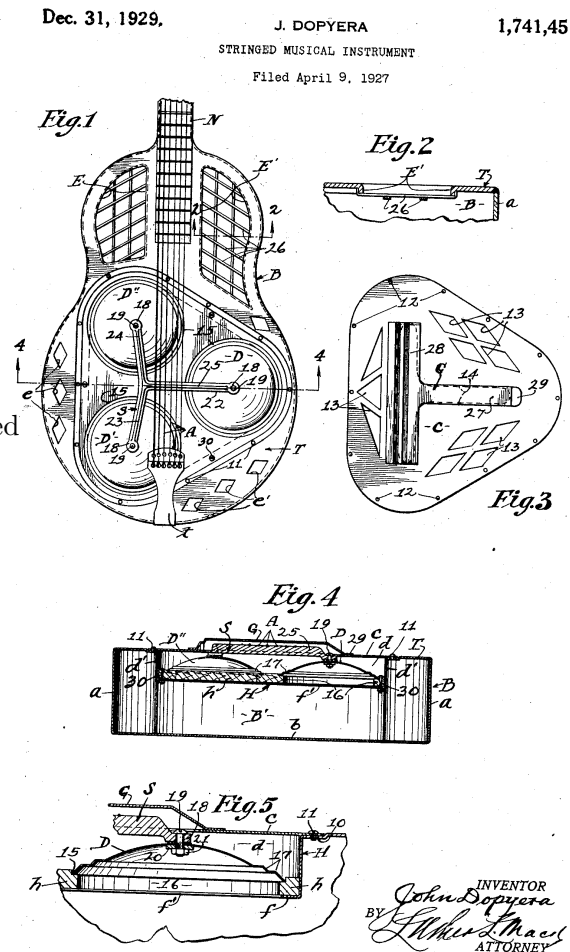


Figure 3-8: Tricone patent [19].

## Design of the National resonator guitars

The bodies of the National guitars were most commonly made of German silver while brass and sheet steel were also used (sheet steel was primarily used in the cheaper single-resonator models). Steel bodies had a faster attack and brighter sound, at the expense of sustain and smoothness. Brass and German silver bodies were heavily nickel plated, while Steel bodies were painted. (German silver is an alloy commonly used in fret wire, consisting of 65% copper, 10-23% zinc, and 10-20% nickel. German silver is also called nickel silver or white brass.)

The body of the Tricone model was made in three parts: back, top, and one-piece sides, with all edges soldered together. The single cone bodies were in two pieces, with either a flat front or back, and the other side deep-drawn. All metal parts were made in Adolf Richenbachers shop, which had one of the largest deep-drawing presses of the time. Production of metal-bodied nationals stopped in 1941 [11].

Necks were made from standard guitar woods such as mahogany and maple. Hawaiian models featured a square neck and were intended to be played face-up on the players lap, while the Spanish models used a triangular neck profile that allowed the guitar to be played facing out like a Martin or classical guitar.

The National guitars were much louder than wood guitars, with a great dynamic range, making them well suited for playing in tents and bars, as well as in recordings. The Tricone was considered to have a smoother sound, preferable for slide and Jazz playing. It was noted for having good sustain and resonance in open tunings. While Bluegrass players prefer the Dobro single-cone models, Blues players such as Son House, and Bukka White used single-cone Nationals almost exclusively.

All resonator guitars are considered to sound better when played with a pick, which results in less string extension and a brighter sound. Because the guitar is so resonant, more right-hand damping is often required.

### 3.1.3 The Rickenbacker electric guitar

The story of the first true electric guitar (with an electromagnetic pickup sensing string motion) starts where the National Guitar story left off, again with George Beauchamp. Before starting National, Beauchamp had experimented unsuccessfully with placing carbon microphone buttons on his guitar to amplify it. In 1930 or 1931, he returned to the topic, taking electronics classes at night. With Paul Barth, another National employee, Beauchamp began working on a single-string prototype guitar using coils wound around a washing machine motors magnet and the amplifier circuit from a Brunswick phonograph [5].

Refinements to the pickup system resulted in the design featured in his 1934 patent that is, apart from the use of a horseshoe magnet (Figure 3-9 label 22), essentially the same as modern-day electric guitar pickups. A coil of wire (label c) surrounds six pole pieces (label 11) that guide the magnetic field through the strings. When the strings move, they create variations in the magnetic field, which causes current to flow in the coil, creating a signal that can be amplified [6].

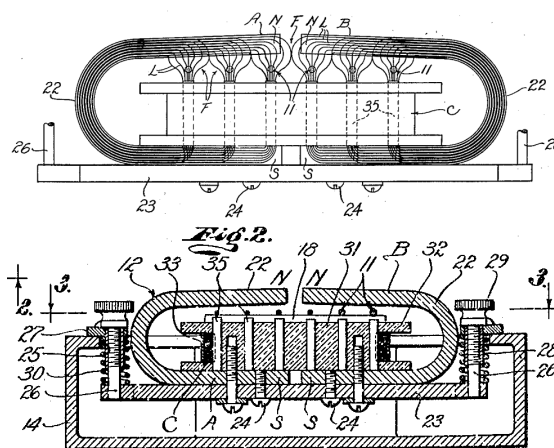


Figure 3-9: Image from Beauchamp pickup patent [6]

Beauchamp recruited Harry Watson (another National employee) to build a wooden lap steel guitar on which to mount the prototype pickup. This guitar was known as the “frying pan for its unusually small round body (shown in the 1934 patent, figure 3-10)

In 1931, Beauchamp, Barth, C.L. Farr (a National board member), and Adolph Rickenbacher (owner of the nearby metalworking company that made all of the metal parts for National guitars) formed Ro-Pat-In corporation (later renamed the more sensible Electro String Instrument Corporation).

Soon after, Beauchamp was fired from his post at National for unspecified reasons, though he continued to serve on the board through 1934.

In 1932, Ro-Pat-In began making versions of the electric frying pan guitar in Rickenbackers factory. These guitars were made from cast aluminum, and featured the electro brand on the headstock. By 1934, the Rickenbacker name was added to the headstock (the spelling changed). They also began manufacture of a Spanish-style wood-bodied guitar (starting with conventional wood guitars from Harmony or Kay, and refitting them with the electric pickup).

Aluminum turned out to be a problematic material; due to its high coefficient of thermal expansion, changing temperatures caused it to go out of tune. This led to experiments with other materials such as Bakelite, before eventually adopting more conventional wood necks.

**Impact** The Rickenbacker frying pan established the basic approach that is still used in electric guitars today. String-driven coil pickups, a solid body, even a  $\frac{1}{4}$ " output jack are all part of what we expect from a modern electric guitar. Though unpopular at first (they sold only ~12 Frying Pans in 1932) later models including possibly the first solid body, the Bakelite Spanish guitar, sold much better. The Beauchamp pickup also set the stage for Leo Fender's development of the Broadcaster guitar, and all modern electrics [10].

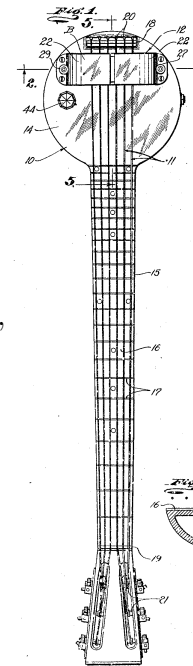


Figure 3-10:  
"Frying  
pan"  
design [6]

### 3.1.4 Ondes Martenot

If the development from the Stroh Violin to the National guitar to the Rickenbacker electric guitar represents the transition from acoustic instruments to pure electric sound, the development of the Ondes Martenot can be seen as almost the opposite; an electronic instrument that took on more and more acoustic aspects over time through connection to specially designed resonators.

The Ondes Martenot is one of the earliest electronic instruments. First conceived in 1917, it went through several significant design changes



Figure 3-11: Ondes Martenot with  
diffusers

before finally arriving on a keyboard design with additional controls. Like many early electronic instruments, the Ondes Martenot used circuitry to create electrical oscillations that could be amplified to produce musical tones. However, the Ondes differs from its contemporaries in several ways.

Later versions of the Ondes presented a mix of continuous and discrete pitch control that gave a unique method of playing that was well suited to music that had an almost vocal quality. This ethereal quality was enhanced by a variety of special speakers that added acoustic resonance to the sound through sympathetic strings and metal plates.

Also, quite unusual for an early electronic instrument, there is an established repertoire of over 1200 compositions for the Ondes Martenot, including many by famous composers such as Varese, Messiaen, and Jolivet. It was also frequently played with symphony orchestras [12][7]. Even today, many composers write pieces for it, and 20-40 new pieces for the Ondes are performed each year.

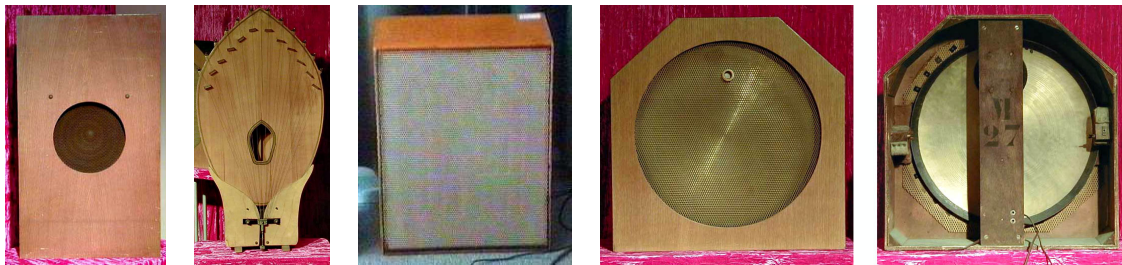
### **Maurice Martenot (1898-1980)**

In 1917, working as a young radio engineer, Maurice Martenot found that he could create oscillations using vacuum tubes and variable capacitors attached to them, a property that Leon Theremin was discovering at almost the same time in Russia. Starting in 1917, Martenot developed several instruments, and in 1928, he presented his Ondes Martenot or Martenot waves at the Paris opera. Soon he was presenting the instrument all over the world.

### **Generations of the Ondes:**

1. 1917 – The first Ondes Martenot was very similar to the Theremin; it used changing capacitance in the air to control the amplitude and frequency of a single oscillator.
2. 1928 In the second-generation instrument, which was presented at the Paris Opera, pitch was controlled by pulling on a cord that was connected to a pulley system with a constant distance per octave. Articulation was provided by a key which was controlled by the left hand
3. 1929 The third generation added a painted keyboard to help the player more quickly find the desired notes.

4. Before 1932 A real keyboard was added so the player could choose to play glissandos using the cord or to play discrete pitches using the keyboard. A new metallic loudspeaker was developed that consisted of essentially a metal gong driven by a voice coil, giving the sound a non-harmonic, metallic sustain. A special 66-note-per-octave model was made for Indian poet Rabindranath Tagore to play ragas.
5. 1937 - Glissandos were now possible over the entire range of the instrument, and the articulation key moved to a retractable drawer at the front of the instrument.
6. 1950 - Another loudspeaker, the Palme was added. The Palme featured two sets of twelve sympathetic strings, one on each side of the enclosure to provide harmonic resonance.
7. 1974 The seventh-generation system was transistorized, and the keyboard could slide left and right to obtain microtones.



(a) D1 “principal”, (b) D4 “palme”, (c) D2 “resonance”, (d) D3, “Metallique” front and back view.  
loudspeaker loudspeaker loudspeaker

Figure 3-12: Diffusers of the Ondes Martenot. a,b,d images used with permission, <http://www.audities.org>

Through the development of the instrument, different loudspeaker designs were created to increase the timbral range of the instrument. The final versions of the instrument featured four distinct diffusers, D1 through D4 that applied different sound qualities to the output. Combinations of diffusers could be engaged through controls for the left hand.

- D1** “Principal”, a conventional loudspeaker, used since the original invention, provides clear, relatively uncolored sound.
- D4** “Palme”, a loudspeaker with twelve sympathetic strings was introduced in 1950. A voice coil drives the strings that are then acoustically amplified by the cabinet.
- D2** “Resonance”, a loudspeaker that provided reverberation, was introduced in 1980. The reverberation was achieved by coupling the speaker to a dense grid of springs, an unusual design since the reverberation is purely acoustic. In some models it shares the same cabinet with D1 (image from <http://www.cslevine.com/>).
- D3** “Metallique”, invented around 1930, used a voice coil to drive a flat brass cymbal, yielding elongated metallic timbres.



## Playing technique

Typically, the right hand plays the keyboard or controls the cord, while the left hand controls the intensity, timbre, and articulation. There is a ring on the cord that the player can put his finger in, and position of the ring relative to the keyboard indicates its approximate pitch. Adding vibrato is possible by moving the cord, or by pushing the keys from side to side, while glissando playing can be achieved with the cord alone.

The Ondes was not an overwhelming commercial success; approximately 400 Ondes Martenot exist in the world; however, the instrument was readily adopted by the musical establishment and was used with symphony orchestras. The existing body of work for the instrument, and the steady stream of new work suggests that the Ondes Martenot will continue to be used for some time.

Recently, Johnny Greenwood of Radiohead has used an Ondes Martenot in some of their studio recordings, and commissioned an Ondes-style keyboard from UK-based Analogue Systems for use in their live shows [65].

## 3.2 Electronic percussion

### 3.2.1 A brief history of electronic percussion

Although some electronic percussion instruments existed before 1960, notably Leon Theremin's Keyboard Electronic Timpani (1932) [30], modern efforts to incorporate electronics into percussion instruments began in the late 1960s as modular synthesizers became more common. Musicians and engineers began experimenting with attaching transducers to pads that they could hit. By plugging the resulting waveform into a modular synthesizer, they could use the trigger output to gate a synthesizer sound [18], or be used as an input to any part of the synthesis. Since all signals in these systems were analog voltages, there was no distinction between control data and audio.

In 1973, Moog introduced what was possibly the first commercial percussion controller. The Moog Percussion Controller Model 1130 was a drum with a sensor in the drum head that could drive the Moog modular synthesizer (figure 3-13) [55]. The PAiA Programmable Drum Set, released in 1975, was one of the first self-contained electronic drum devices (Figure 3-14). Its sounds were made by sending impulses into almost-oscillating filters to make them ring. Also credited with being the first programmable drum machine, it featured touch pads that responded to skin capacitance. A modification was available to approximate velocity-sensitive pads by measuring changes in skin capacitance and resistance, taking advantage of the fact that the skin's complex impedance is approximately proportional to finger pressure on the pad [53].

In 1982, Simmons released a new electronic drum kit, the SDS5 (Figure 3-15), which quickly became popular with emerging new wave bands, in part because of their recognizable hexagonal pad design. Bass, snare, and tom sounds were produced using analog oscillators and noise generators [3]. Modules that used sampled audio for Hi-hats and cymbals were added to the lineup, but their acoustic counterparts were usually preferred.



Figure 3-13: Moog Percussion Controller 1130



Figure 3-14: PAiA Programmable Drum Set [53]

Some percussion controllers have ventured further from traditional drum designs. Don Buchla has made several percussion interfaces that depart substantially from simple emulation of drum kits. Buchla’s Thunder is a drum intended to be played using fingers, and it can track the position of the depression made by the finger using an optical system on the back of its reflective mylar drum head [66, 55]. The position and velocity information can be mapped to any MIDI control, opening up a range of sounds and mappings not previously possible. Buchla’s Marimba Lumina is a marimba-like controller that can sense which mallets struck it, as well as where on the bars it was hit, enabling different mallets to be mapped to different timbres [90].

The Mathews/Boie radio drum used capacitive sensing to track the 3D position of transmitter batons in the air close to the playing surface [8], giving the player free-gesture in addition to the discrete hits typical of a drum pad.



Figure 3-15: Simmons SDS5 drumset

The Yamaha Miburi was a suit that (among other things) allowed the player to perform on a virtual drumset in the air [55], and Laurie Anderson’s drum suit, made drum triggers wearable [26].

The Rhythm tree [55], by Joe Paradiso et al. was a 320-pad percussion installation that toured as part of Tod Machover’s Brain Opera project. Each pad could sense velocity and discriminate between top and side hits. To a lesser degree, it could also sense the difference between damped hits and undamped hits.

Recently there has been interest in using percussion controllers not just as performance instruments, but as a way to understand the dynamics of drumming. Ajay Kapur et al have been working to capture and analyze traditional north Indian percussion technique through designing specialized interfaces using impact and pressure sensors and also through classification of the audio spectra of real drums [39].

### 3.2.2 Contemporary commercial approaches

With the development of MIDI [47] in 1983, a new set of percussion controllers became available that could be plugged into any synthesizer module, making them truly generic

controllers. Simmons, DrumKat, and Roland offered a variety of drum controllers modeled after drum kits, marimbas, and in some odd multi-pad arrangements like the Roland Octopad and DrumKat’s eponymous DrumKat [95]. Generic drum trigger boxes that could accept a range of triggers, usually including a built-in set of drum samples, include the Yamaha TMX, and the Alesis DM series.

Three percussion instruments are particularly relevant to this thesis. Two, the Roland V-drums and Handsonic, are market successes and are the dominant electronic percussion instruments of their kind today. The other, the Korg Wavedrum is a much more radical, but short-lived design that provided the inspiration for much of the work in this thesis.

## Roland V-drums

The electronic percussion market had seen a short burst of interest with the success of the Simmons SDS drum system, which prompted companies such as Roland to build their own electronic drum sets. But as the fad passed, the market went into what Roland founder Ikutaro Kakehashi described as an “irreversible decline” which he attributed to the poor tactile response of existing electronic percussion instruments [38].



**Figure 3-16:** Roland V-Pro TD 20S [73]

After their success with their V-guitar system, Roland began work on a “V-drum” system. Roland combined wavetable synthesis with the physical modeling techniques they call “COSM” or Composite Object Sound Modeling, that can allow the user to edit the default drum kits to change parameters such as kick drum beater material, cymbal size, and snare tension [73].

At the same time they began work on improving the tactile response of their drum pad designs. The first V-drum, the PD-7 released in 1997, used rubber-covered plastic pads with separate rim and center sensors. However, they were unhappy with the feel of these pads, and wanted something much more like a real drum head. Roland engineer, Hiroyuki Nishi describes the origin of the mesh drum head design:

One day, our Mechanical Designer, Mr. Yoshino, visited a do-it-yourself shop a supermarket for carpentry and gardening goods and coincidentally found a small trampoline, which used a mesh-type material for the bouncing mat. He had an inspiration: Roland could use a mesh surface for the drum pad.[79]

The mesh could bounce the stick back like a drumhead, but since air passes through the mesh freely, it made very little sound. After several design iterations, Roland had their design produced by the drum head manufacturer Remo.

Roland also developed a method for sensing the annular location of drum hits. (figure 3-17). They found that the period of the first half-wave of output from the sensor decreased as hits occur further from the center of the drum [101].

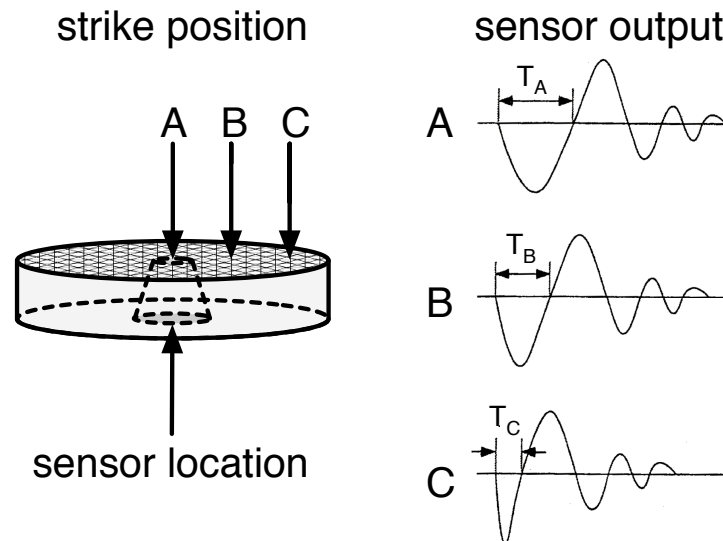


Figure 3-17: Roland position sensing system: The period of the first half-wave decreases as hits get further from the center (figure based on [101]).

The V-drums can therefore measure velocity and stick position, and dual-trigger pads can detect rimshots and cymbal chokes [78]. The position and damping information can be used to control the synthesis and sample playback to give much more realistic output.

Cymbals were another challenge. The first V-drum set, PD-7 used the same rubber pads for cymbals as well as drums. Roland engineers experimented with using damped real cymbals, stretched mesh, and finally converged on a system using a plastic frame and a hard rubber cover. By varying the thickness of the cover, they were able to make the impact sensor output uniform over the entire surface of the cymbal. A second sensor was added at

the rim to detect chokes and crashes (labeled "3" in figure 3-18) [100]. The V-cymbals were released in 2001, followed by the V-hi-hat in 2004. These developments are quite recent, and highlight the continued importance of electronic percussion development.

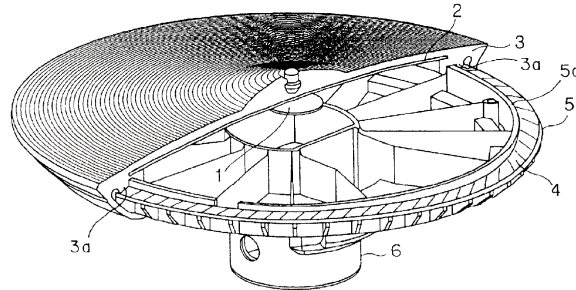


Figure 3-18: Roland V-cymbal [100].

The V-PRO TD 20S 3-16 is the current top of Roland's product line, incorporating mesh pads, V-cymbals, and a V-hi-hat, representing the state of the art for electronic drum kits. Their triggering and control is widely considered to be sufficient for many popular music applications which would have required an acoustic drumset. [71].

## Roland Handsonic

The Roland Handsonic (figure 3-19) is a self-contained percussion system for hand drummers. What looks like one large circular rubber pad is actually divided into 15 zones, each of which can trigger a different sound. Sensing is achieved by a set of piezoelectric elements around the perimeter and force sensing resistors (FSRs) under the center pads (figure 3-20). By comparing the piezo and FSR output, position of the strike can be computed [81]. Several of the built-in patches also take advantage of pressure sensing to mute or detune the drum sounds.



Figure 3-19: Roland Handsonic HPD-15

An infrared range-finder controller called the "D-Beam" allows for free-gesture control by measuring the distance between the player's hand and the unit, and two ribbon controllers

provide additional expressive control. The synthesizer section is a conventional wavetable followed by an effects processor with variable parameters that can be controlled via the ribbon or D-beam depending on the preset patch. The Handsonic also works as a MIDI controller; trigger, pressure, ribbon and D-beam, along with knob positions and an optional footswitch can be recorded using a sequencer, or can be used to control another MIDI device [72].

One of the weakness of both the v-drums and the Handsonic is that they still depend on a trigger. This means that despite their sensitivity, because all of the hits are discretized, they can fail to recognize small hits, and actions like stirring with brushes or sliding a hand on the pad won't produce any reliable output [42].

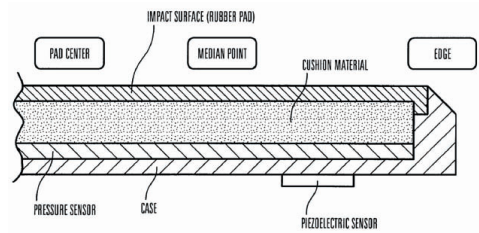


Figure 3-20: Pressure sensor and piezo (from [81])

The only commercial system that has been robust to these kind of playing gestures is a short-lived instrument called the Korg Wavedrum

### Korg Wavedrum

The Wavedrum (figure 3-21) was one of the first commercial applications of physical modeling, but it is unique in that it used the acoustic sound of the drum head to directly drive the synthesis algorithm. A team of former Sequential Circuits engineers (makers of the Prophet synthesizer) were responsible for designing the synthesis system. Some modes used simple triggering, while others filtered the audio from the drumhead through various waveguide drum models and effects. By using some of the actual sound present in the drumhead, the wavedrum was far more responsive and predictable than existing percussion controllers [64].



Figure 3-21: Korg Wavedrum [40]

A piezo sensor located under the drum head, below the sensor cover picks up vibration from the drum head. The cover keeps players from accidentally hitting the sensor directly and damaging it. A rubber cushion located under the head provides damping, and a sensor below the cushion measures pressure applied by the player. The spacing between the head

and the cushion is very delicate, and care must be taken not to damage the sensors when the head is off [41].

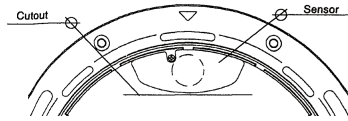


Figure 3-23: Wavedrum sensor [41] image courtesy Korg corporation

A rim can be bolted on (figure 3-24) to the Wavedrum, or used independently on its own stand. The rim contains another piezo sensor, and has ridges with two different spacings to allow for guiro sounds and scrapes in addition to rim shots and other sounds (figure 3-25).

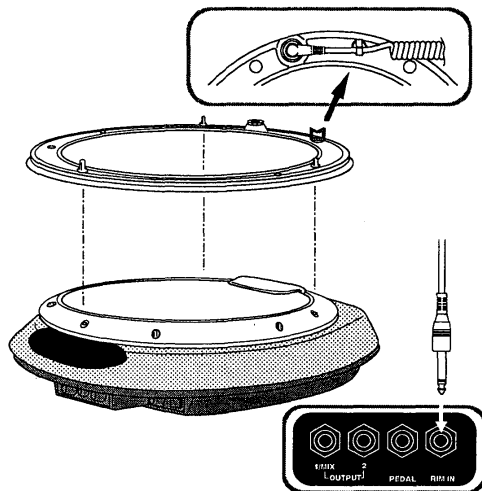
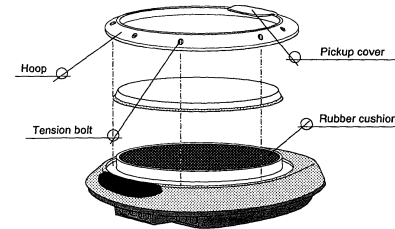


Figure 3-24: Rim assembly [41]

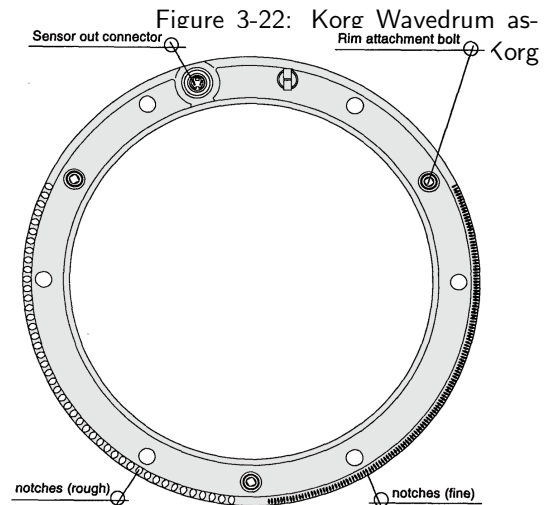


Figure 3-25: Rim [41]

Images courtesy of Korg corporation

Extensive editing was possible on the unit, but due to the constrained interface, more detailed modifications were much easier with the optional RE-1 remote editor [15]. Though it was only sold for a short time, the few that were made command high prices in the used market, and the uniquely expressive sounds are featured in sample libraries for drum loops [77]. Its high price (\$2499 not including the editor or stand) was likely a contributing factor to its quick demise.

Musician and writer, Gordon Reid describes his view of the market failure of the wavedrum:



Many musicians seemed to think that the Wavedrum was merely a more expensive way to obtain sounds similar to those produced by PCM-based drum pads, but it was far more than this. It could produce all manner of sounds, ranging from traditional percussion through to overdriven lead guitars. The Koto patch, for example, was a true 'string' model, with control over pluck position, string damping, plucking noise, and more. Unfortunately, Korg never managed to convince enough players of the potential of the Wavedrum, and few, if any, plumbed its depths. What a waste! [64]

Although the Wavedrum can send and receive MIDI data, it can not send the actual sound of the drumhead over MIDI, so much of its expressiveness is lost in that mode. Other limitations are that each patch has to be handcrafted (parameters of the system are modifiable, but creating new patches is limited to the building blocks already in the system [15]). Creating a particular sound that is not very similar to the preset sounds is difficult if not impossible, requiring knowledge of the instruments' physics as well as understanding of how to reduce it to a tractable waveguide representation that can be calculated efficiently.

One of the main ideas of this thesis is to use the concept of direct audio processing for percussion, as in the Wavedrum, but to use realtime convolution instead of filters and waveguides to allow the flexibility to use recordings of real instruments as the resonator.

### 3.3 Convolution in computer music

Convolution is a well known algorithm in physics and engineering that can be used to evaluate the similarity of two functions, among a variety of other signal processing tasks. In computer music it has been primarily used for filtering, adding resonances, reverberation, and cross synthesis. Although possibly the earliest discussion of its musical applications was by Richard Boulanger in 1986 [9] in which he described musically relevant methods for processing speech and other signals, convolution remained largely unrecognized as a musical tool.

Curtis Roads' 1993 ICMC review paper describes the state of convolution in computer music at that time:

Convolution occupies an odd position today. It remains unknown to most

musicians, yet to signal processing engineers, it is a basic topic. The mathematical theory of convolution was nailed down long ago, so that signal processing textbooks inevitably present it abstractly in the first few pages, reducing it to a handful of mathematical clichés . . . Unfortunately, The musical significance of these equations is not well known or appreciated, either by engineers or musicians [69].

Since the mid 1990s thanks to the increase of power and availability of computation, and descriptions of musical applications [69],[70], convolution has become a much more common tool in computer music.

Convolution (represented by the symbol  $*$  of two functions  $x$  and  $y$  is defined as

$$(x * y)_n \triangleq \sum_{m=0}^{N-1} x(m) y(n - m) \quad (3.1)$$

where  $N$  is the length of the signal  $y$  [84]. If you know the response of a linear system to an impulse, you can obtain the system's response to an arbitrary function by convolving that function with the impulse response of the system.

This technique is widely used to implement filters of known impulse response, and specialized DSPs have been designed to perform the necessary multiplication and summing quickly enough to achieve filtering in real time. Since this algorithm is of order  $NM$  ( $N$  is the length of signal  $x$ ,  $M$  is the length of signal  $y$ ). working with long impulse responses in the time domain can still be prohibitive.

Shortly after the discovery of the Fast Fourier Transform (FFT) by Cooley and Tuckey in 1965<sup>1</sup>, Stockham [89] used the FFT on short sections of signal to implement fast convolution. This algorithm is of order  $N + M \log(N + M)$  making it substantially more efficient for long convolutions.

Both the time domain and frequency domain approaches have been used in computer music to achieve a variety of filtering and resonance effects, and spatialization [20].

---

<sup>1</sup>Later research by Cooley discovered a previous description of the algorithm by Gauss in 1866 written in neoclassic Latin, which likely limited its influence!

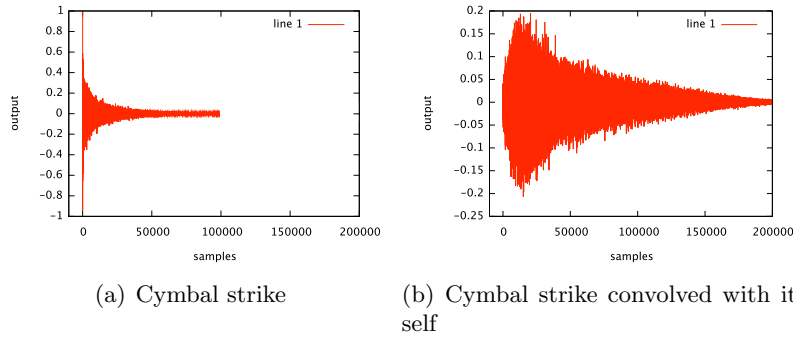


Figure 3-26: Time smearing

Some common uses:

**Time smearing** Convolving a signal with itself or another signal prolongs the sound, slowing the attack and decay, and resulting in output the length of the sum of the signals (- 1 sample). This effect is show in in figure 3-26.

**Cross synthesis** When two signals are convolved, the common frequencies in both are boosted, while frequencies present in only one signal are cut. This can give a sense of a hybrid signal that is neither totally one signal or the other, but a combination of their common spectra, such as making the sound of the wind speak. Similarly, the rhythmic aspect of one signal can be melded with the timbre of another.

**Filtering** Convolving a signal with the impulse response of a filter is the same as applying that filter to the signal. In some cases it is easier to obtain the impulse response than to know the actual mechanisms at work in the filter.

**Reverberation** The imulse responses of concert halls can be recorded and applied to any signal, giving the sense that the sound was played in that particular space. Multiple microphones can be used to capture impulses for surround sound or stereo effects.

One of the first instances of convolution in a commercial instrument was the E-mu Emax SE (an upgrade option for their popular Emax sampler) which was released in 1987. “Transform Multiplication”, their term for direct convolution, could be performed on any two stored samples in a very non realtime fashion, with the process often taking many hours [21].

But it wasn’t until personal computers became powerful enough to perform convolutions in reasonably short times that it saw more widespread adoption. In 1991, Soundhack [23], a useful audio conversion and editing tool for the Macintosh computer, added convolution to its set of (non-realtime) audio processing algorithms, making it readily accessible to a broader range of users. Curtis Roads’ 1993 ICMC paper [69] was among the first to describe the musical applications of convolution to the computer music community.

By the early 90s, some realtime musical applications were also beginning to take shape. Lippe and Settel described their realtime (moderate latency) convolution using the IRCAM spatial workstation and Max [82].

One drawback of FFT convolution is that in its realtime application, it requires at least two frames of latency. Gardner [28] and also McGrath [43] developed zero latency methods that use direct convolution for the first part of the impulse response, and fast convolution for the remainder, with progressively larger windows. This approach allows true realtime low latency processing (limited by the audio hardware) with modest hardware requirements.

The primary application of low latency convolution has been in reverberation plugins for digital audio workstations such as Protools and Digital Performer, though there has been some interest in using the same plugins for more general sound processing [88]. Convolution (for reverberation) has recently been added as a built-in effect in samplers such as Tascam's Gigasampler, and reverberation impulse recordings are available on CDs for people wishing to expand the number of virtual environments available to them for reverberation.

### 3.3.1 A graphical example of convolution

It can be easier to think of convolution as a series of impulses. In this example, we will show one way of looking at convolution in the time domain. Consider the case of reverberation, a common use of convolution: if you bring a starter pistol into a concert hall and record its firing at a distance, you will have the impulse response of the hall. The pistol is approximately an impulse (figure 3-27), that is, it excites all frequencies of the hall equally.

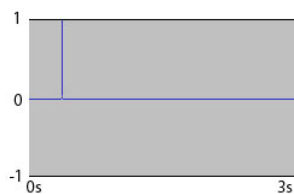


Figure 3-27: An impulse (waveform plot, time goes from left to right on the x-axis)

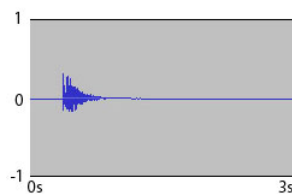


Figure 3-28: The room impulse response

If you wanted to apply that room response to a second impulse, you could multiply the room impulse response by the new impulse, yielding exactly the room response again, as in figure 3-28.

Suppose you had two impulses close together (figure 3-29). you could multiply each by the room response, and add the samples where the responses overlap (figure 3-30).

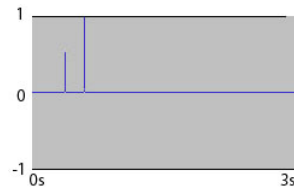


Figure 3-29: Two impulses

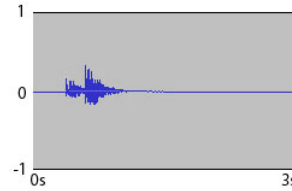


Figure 3-30: Two impulses convolved with the room response

Now consider a continuous stream of impulses at different amplitudes and at a rate of one per sampling interval, which can represent any audio signal. In figure 3-31 and 3-32, the input audio is the sound of a plucked string.

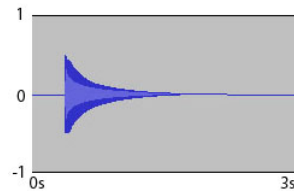


Figure 3-31: A plucked string sample

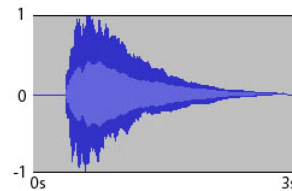


Figure 3-32: A plucked string sample convolved with the room response

Unfortunately, performing this process in the time domain is computationally very intensive; Each incoming sample of audio is multiplied by the entire impulse response and the result is added to the previous output. A less computationally intensive method is to perform the same task in the frequency domain.

**Fast convolution** Convolution of two signals in the time domain is equivalent to multiplication in the frequency domain. In the case of convolving a stored signal with an incoming signal, first the stored signal is prepared by windowing and taking the FFT of each window. Then the incoming signal is windowed, Fourier transformed, and multiplied by the transform of the stored signal.

Figure 3-33(a) shows the magnitude of the FFT of 512 samples of a plucked string sample, and figure 3-33(b) shows the magnitude of the FFT of 512 samples of bandpassed noise. Figure 3-33(c) shows the result of multiplication of the spectra: common frequencies are boosted, while frequencies that are low or absent in either signal are cut. In this case,

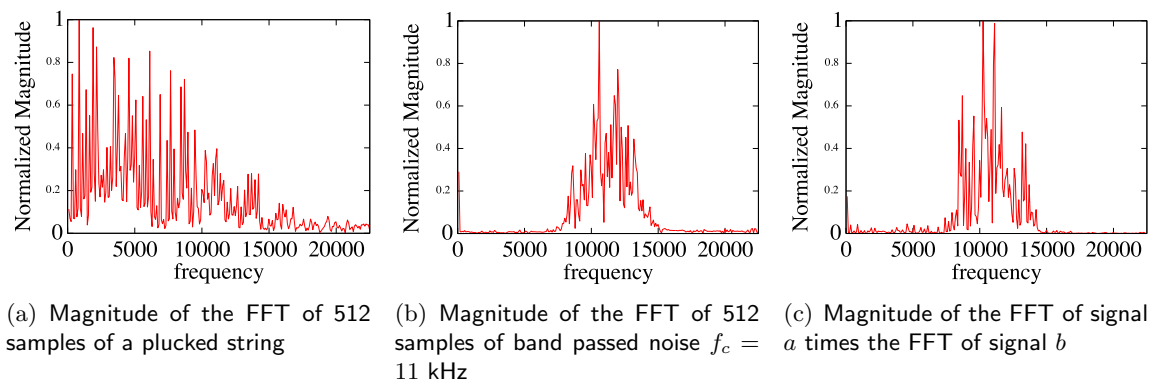


Figure 3-33: Multiplying in the frequency domain

the overall spectral shape of the filtered noise predominates in the output, but the modal structure of the plucked string persists.

By varying the window size, latency can be reduced. This is the method used in the techniques outlined in this thesis.

### 3.4 Other approaches: physical modeling and modal synthesis

One ongoing commercial and research approach to the challenges of realistic synthesis has been to create instruments using models of the physics of real instruments for sound production. Since these models rely on the same parameters as real instruments, they have the potential to be just as expressive and controllable, given appropriate physical parameters as inputs. Commercially, in addition to the Korg Wavedrum, the Yamaha VL1 and the Korg prophecy were two other significant physical modeling synthesizers. In addition to providing rich control over the sound, physical modeling also provides a way to recreate some of the physical constraints of real systems. The challenge of this approach is that the physics of the instrument have to be well understood, and to be useful for realtime instruments, the models have to be computationally tractable.

Commercially, physical modeling synthesizers have not been a tremendous success. The ability to control so many degrees of freedom of a sound is very different from hitting a single key and hearing a rich pad sound. The Yamaha VL1 for example, is a keyboard instrument

with an additional breath controller. Part of the difficulty is that the presence of a keyboard suggests discrete interaction. A Yamaha product manager describes the dilemma:

The VL1 is an instrument that actually like its acoustic counterparts requires some practice. You really have to develop some technique with the breath controller, three wheels, pedals, aftertouch, and realtime sliders to extract the maximum expressiveness from the technology ...it's very much like picking up a clarinet for the first time [92].

The VL70m, a similar physical modeling synth has done better, in fact it is still in production. It found a niche with players who use the WX-5 wind controller. The use of a controller specifically designed for the physical models likely helped it gain traction, along with the fact that the controller is intended to be used by reed players, who are already more comfortable controlling many degrees of freedom simultaneously.

Digital waveguide models are systems of delay lines that can be combined with linear and nonlinear filters to approximate traveling waves in various media. Real waves encounter losses and filtering throughout their travel, but those effects can be summed and moved to the junctions of lossless delay lines, greatly reducing the computational requirements. [85]

Physical modeling is extensively described by Julius O. Smith [86]. Some systems are easier to model than others; cymbals and gongs have been particularly difficult to model due to their complex behavior [13] and computational requirements.

**Modal synthesis** A second method, called “modal synthesis” uses the fact that complex resonating structures can be modeled as a sum of simpler structures that are each a “mode” of the more complex system, each with its own fundamental frequency and damping coefficient. For realtime applications, banks of tuned resonators such as second order IIR filters are used to represent each mode. By controlling the frequency and decay time of the resonators, the response of the system can be tuned to match that of any (linear) target system. Spectra of recorded sounds can also be analyzed to obtain the frequency and damping coefficient of each mode, without necessarily needing to understand the physics of the particular instrument.

Additional expressive control is achieved by allowing frequency and decay to be varied for each resonator during playing, allowing for glissandos and complex spectral modification [93]. Modal synthesis works best when there are a small number of modes, but becomes computationally impractical for systems with very complex spectra, such as cymbals.

Work has been done on connecting audio output of physical objects to physical models and modal synthesizers to make percussion instruments [93], [77]. Van den Doel [93] suggests using contact microphones for this task.

An interesting interface to this type of synthesis is a sensor that measures real interaction forces. This can be demonstrated with a contact microphone. When touching and scraping real objects the audio signal can be sent to a synthesis process, where this audio signal is then interpreted as a force to whatever vibration model is currently loaded. We can then scrape some interface object and transfer the measured signal to the audio synthesis to create the impression of touching a virtual object [93].

The work in this thesis differs from this modal approach in that *any* recorded impulse can be used as a resonator, not just ones that can be modeled by a filter bank. This comes at a greater computational expense and does lose some of the spectral control possible in modal synthesis but in exchange for greater generality. The next chapter describes the specific design and implementation of that system.



## Chapter 4

# System design and software implementation

For this thesis, I have constructed several examples of hybrid digital / acoustic instruments. This chapter describes the design and implementation of the main signal processing section.

### 4.1 How the system works

In this method, sound from the physical object is transduced by contact microphones or other pickup design connected to the audio input of a computer. The computer continuously digitizes, filters, and convolves the incoming audio with the pre-recorded impulse response of a desired instrument. The resulting transformed audio is routed to the output of the computer and can be played over loudspeakers. This basic system is shown in figure 4-1.

The “impulse response” in this case is a recording of the particular instrument of interest being struck clearly. A clear elastic strike is a reasonable approximation of the impulse function in that it excites nearly all modes of the instrument. One should note that the result is not the impulse response of the entire instrument, but the response to an impulse applied to the particular location where it was struck. Striking it in other locations could produce different modal structures.

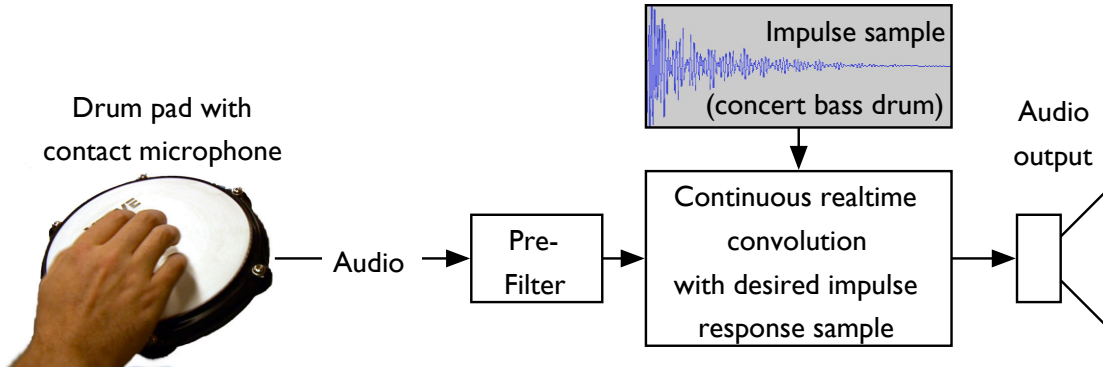


Figure 4-1: Block diagram of the basic system

## 4.2 Initial proof of concept

In the first trial implementation, convolution was performed using a commercial realtime VST plugin hosted by the graphical audio programming environment Max/MSP [16]. Rough pre-filtering was performed in Max before the audio was sent to the convolver to flatten the typical spectra of the physical object and contact microphone. A MOTU 828 audio interface was responsible for all audio input and output.

For a physical controller, a PZT piezoelectric element was taped to a table top to pick up vibrations when the table was struck. The impulse responses used were samples gathered from the Web, as well as from the Sound Ideas library [87] and a percussion ensemble CD which had some sounds in sufficient isolation.

While this system had significant advantages in its efficiency due to the use of optimized commercial software, it was lacking in realtime control over the processing, and any modification of the parameters of the plugin would cause long gaps in the audio.

## 4.3 Further implementations

The next step was to re-create the convolver function inside Max/MSP. The primary goal of the implementation was to minimize latency while still allowing realtime control over the processing. I used the MSP internal `pfft~` (MSP's system for FFT subpatches) to implement partitioned convolution with variable window size as described by Bill Gardner

[28] , outlined in section 3.3. Max/MSP (and Pd, to be discussed later) both allow for reblocking of sub patches with different block sizes, enabling the output of different sized blocks to be processed separately and for their outputs to be added together.

Unfortunately, although it was possible to perform realtime modification of the sounds, this system proved to be too slow on the hardware I had available at the time. Luckily, there was a partitioned (fixed block size) convolver external [80] for Pd [62], an open-source graphical programming environment similar to Max/MSP. This had improved performance over implementation using Max/MSP's built-in functions.

Later I was also able to gain use of a much faster computer, a 2.5 GHz quad G5 Macintosh, which enabled the system to be re-implemented to work with a combination of the Pd externals, and an implementation using internal Pd objects `rfft~` and `rifft~`, which allowed greater flexibility for additional expressive controls (to be described in the following chapter) While the faster machine removed the original reason for switching to the Pd environment, I had already committed enough time to it that I decided to continue working in Pd.

### 4.3.1 Pd patch architecture

To minimize processing, stored samples are Fourier transformed at the time they are loaded. When a new stored impulse response file is loaded, it is placed in a buffer and cut up into partition-sized lengths. These slices are windowed (using a square window), fast Fourier transformed, and loaded into tables to be processed by the convolver (figure 4-2). These slices are of increasing size to minimize latency, which is equal to double the block size. The first 128 samples are transformed as two blocks of 64, the following 256 are transformed with a block size of 128, and so on up to the maximum block size (typically 4096 samples) at which point the block size repeats until the end of the recording (figure 4-3). Since each partition requires a real FFT and IFFT, its total latency is twice its block size. By convolving two blocks of each size, for a single impulse, the shorter blocks finish playing exactly as the next-longer block begins playing, giving a seamless output. This does require adding a delay to the audio input going to the second block of each partition, and adding progressively longer delays before the same-sized partitions at the end of the recording.

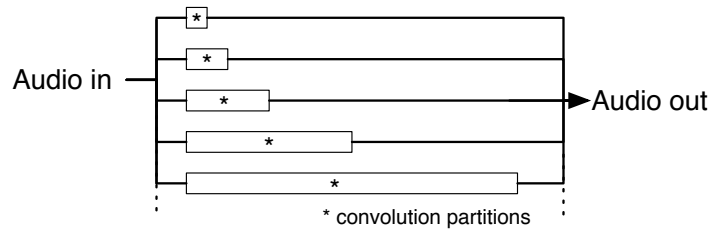


Figure 4-2: Minimal system schematic

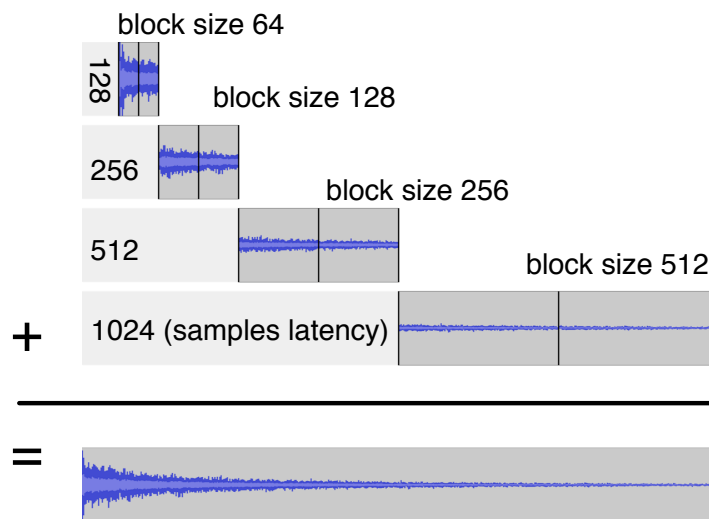


Figure 4-3: Latency of partitions corresponds to partition offset: When a stored sample is convolved with an impulse, the outputs of each of the partitions exactly line up so that the next partition starts as the previous one finishes.

In the convolver, each pair of convolution partitions resides in its own subpatch, and each subpatch can have its audio block rate set independently using the `switch~` object (though it needs to be a power of 2). This requires only one FFT per convolution partition. New audio coming in from the physical interface is fed into all of the partitions, with additional delays for the repeated partitions.

Control data (for damping, cross fades, pitch shifts) is sent at data rate to each sub-patch where the appropriate levels are calculated.

## 4.4 Nonlinear responses

One weakness of the technique of using impulse responses to represent physical systems is that it does not account for nonlinearities. Some percussion instruments such as cymbals and gongs have significant nonlinear responses that are amplitude-dependent, resulting in their rich spectrum. Because of their complex behavior, cymbals and gongs are also particularly hard to model.

For gongs, the modal frequencies can shift with amplitude, with as much as 20 percent frequency variation as the sound decays [25]. When driven with a fixed tone, gongs will develop subharmonics and overtones as the displacement increases. [13]

When driven sinusoidally, cymbals exhibit three distinct modes of operation (figure 4-4): at low amplitudes, harmonics of the driving frequency develop, with greater amplitude as the driving signal increases. At medium amplitudes, subharmonics develop, filling in the spectrum, yielding a non-harmonic sound. At high levels, the cymbal exhibits chaotic behavior, with a very complex spectrum [91],[25], [75]. This accounts for why crashing a cymbal sounds different from a louder ride sound.

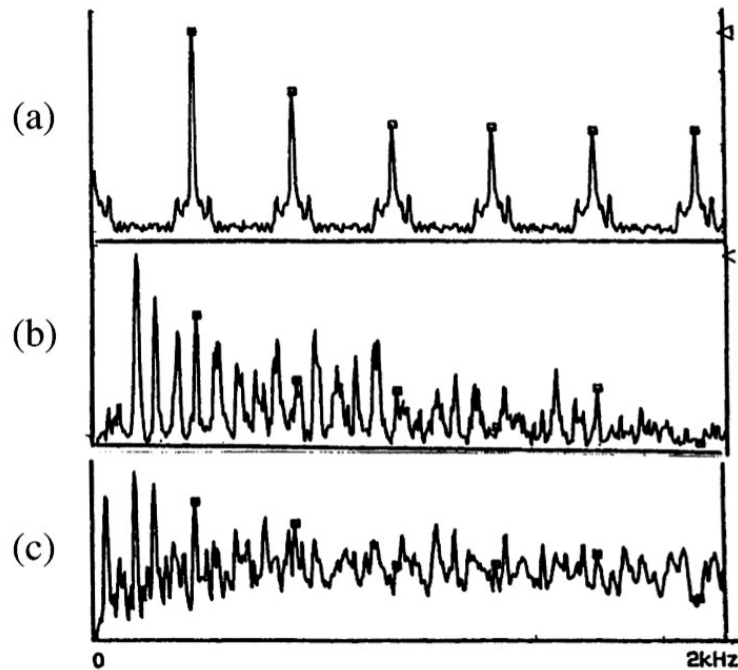


Figure 4-4: Three regimes of cymbal response to a sinusoid of fixed frequency at low, medium, and high amplitude, respectively: (a) harmonic spectrum, (b) subharmonics, (c) chaotic response. From [75]

If one were to send a louder impulse through the convolver, it would have no effect on the spectrum, but would just result in a louder output. If one convolves with a cymbal sample in which the first part is in the chaotic regime, decaying to the subharmonic, and finally harmonic regimes, all output will be in those same regimes, following the same time profile, regardless of hit intensity.

One approach I have used to make a convincing crash cymbal has been to run two convolutions, one of a standard ride hit, and the other of a crash, and to only send the driving signal to the crash if it is above a set threshold. This method is discussed in section 5.1.6.

Since the driving signals in the system typically are impulses that cover many frequencies, the fact that harmonics and subharmonics do not occur is largely obscured by the broad input spectrum. If subharmonics are present in the stored sample, it is likely that there is some energy in the input signal at the same frequencies as the subharmonics. Figure 4-5 shows a typical spectrum of a hit. Although the power at various frequencies varies, there is some energy in every frequency band.

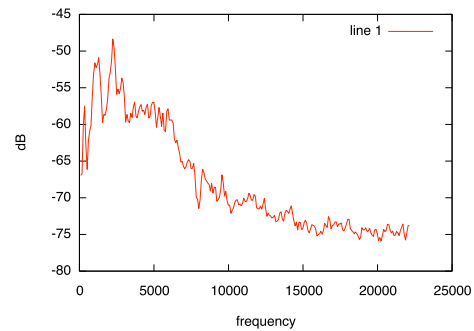


Figure 4-5: Power spectrum of a practice pad being hit with a drumstick.

Similarly, in the case of a gong, even if the modes shift by 20 percent, it is unlikely to cause gaps because of the relatively broad spectrum of the driving signal.

## Chapter 5

# Other expressive controls, extensions to realtime convolution

Beyond varying the spectrum of the hits, players of real percussion instruments often have control over other features of the instrument including damping and pitch, which can play significant roles in the player's control of the sound and musical expression. To perform such modifications to the sound would ideally occur by changing the stored impulses. This would work if we could run fixed block size of 64 samples, for example. Unfortunately, relatively long block sizes are required (and their associated higher latencies) to make the system computationally tractable. Because of the latency, simply switching out the stored impulse is not an option. The bulk of this chapter is then devoted to figuring out how to approximate the desired effects *without* being able to change the stored impulses.

### 5.1 Damping

One very important property of real percussion instruments that they can be damped. The player can press on the drumhead or grab a cymbal and the sound will decay more quickly. In physical systems, energy losses can occur internally or in transfer to a part external to the system.

Viscous losses (such as air resistance) are proportional to velocity, such as seen in a dashpot (figure 5-1) yielding an exponential decay. However, other damping mechanisms do not behave as exponentials. For example, internal friction in a non-viscous material provides a constant force opposing the direction of movement, but independent of velocity, resulting in a linear decay [68]. This is referred to as hysteretic, or coulombic damping. The observed decay for any system is the sum of all of the damping mechanisms. In percussion instruments, viscous damping tends to predominate at the attack and early decay due to higher velocities, while hysteretic damping dominates the tail. If a player further damps the system by resting a hand on it, the hand acts as an additional damper, increasing the rate of decay of the system.

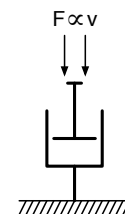


Figure 5-1: Dashpot

### 5.1.1 Simple damping model

In the convolution percussion system, we would like to let the player damp the sound in the same manner as with an acoustic instrument. Ideally, we would multiply the stored impulse by a known function that yields a decay curve that is similar to that of the damped instrument, for example the exponential decay in figure 5-2. By superimposing a new decay curve on the original signal, we can obtain a new apparent degree of damping.

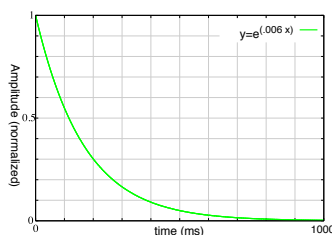


Figure 5-2: Exponential decay ( $\lambda = 0.006$ )

The sampled impulse responses already exhibit approximately exponential decay (except for the very end of the sample where there is usually a linear fade out to zero.) This is both because of hysteretic damping in the object, which is more prominent at lower amplitudes, and because a linear fade out is often necessary when editing the audio samples to keep their duration reasonably short. To make it sound as if the damping coefficient ( $\lambda$ ) of the real instrument were higher, one can multiply the recording by another exponential, as seen



in figure 5-3.

$$e^{-\lambda_1 t} e^{-\lambda_2 t} = e^{-(\lambda_1 + \lambda_2) t}$$

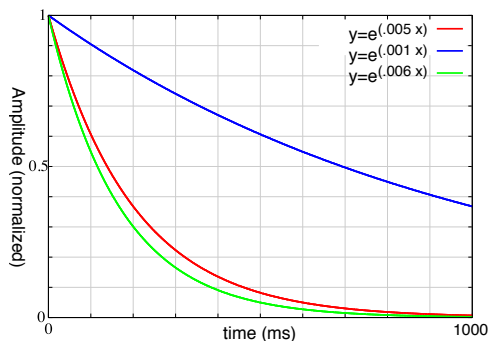


Figure 5-3: Superimposing the desired decay onto the stored response (multiplying two exponential decays yields another exponential)

Unfortunately, the system works by storing the FFTs of the various impulse partitions to avoid having to recalculate them. Any operations performed on the stored impulse in the time domain would require an additional FFT. In addition, any changes to the impulse would require at least one block of latency for the FFT and IFFT before they were heard by the player. This presents a problem: multiplying two time domain signals is equivalent to convolution in the frequency domain. For large signals this is not computationally tractable.

One solution is to control the gain of each block at its output, so the early sounds are louder than the later ones. Recall that the system uses variable-size convolution partitions to limit the overall system latency (figure 5-4).

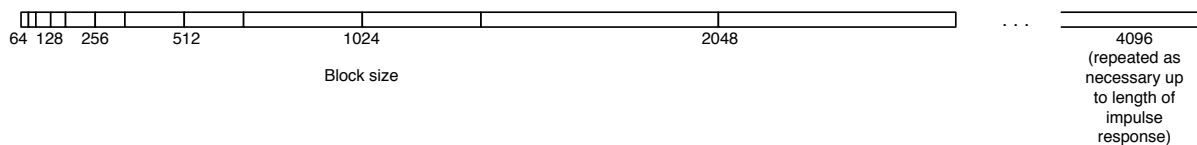


Figure 5-4: Variable sized windows to reduce latency (from [28])

The block gains can be set to approximate any function, but since the gains are constant within each block, the output takes on a staircase shape, shown in figure 5-5.

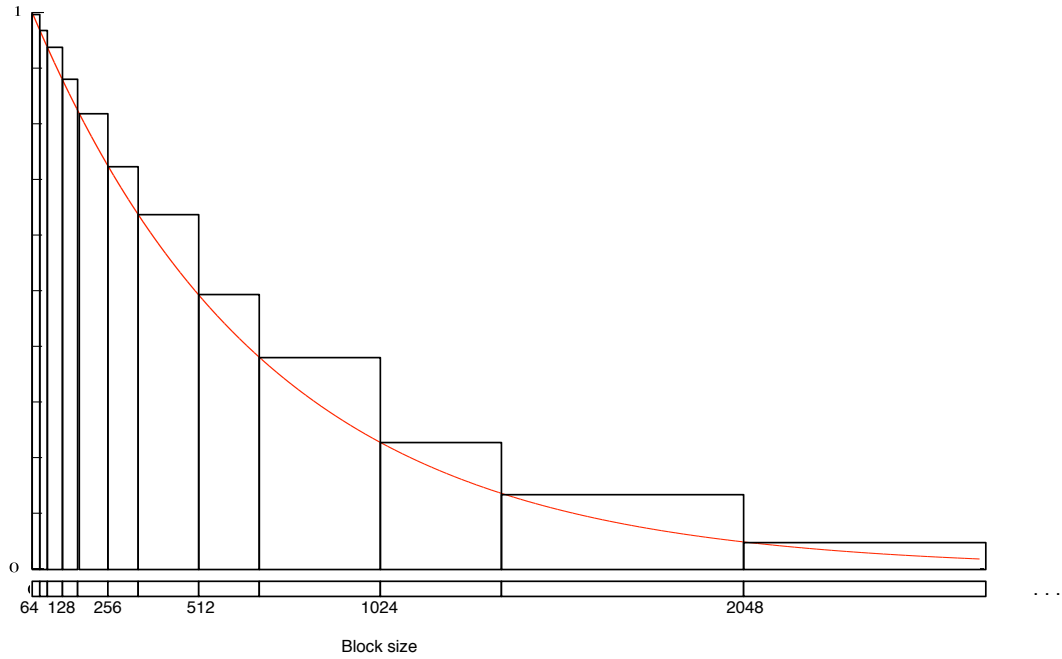


Figure 5-5: Exponential decay superimposed over varying block sizes

### Calculating block gains

The convolution blocks start out with two 64-sample blocks, two 128-sample blocks, etc., as shown in figure 5-4. The sample location  $t$  relative to the start of the impulse response recording is given by the sum of the previous blocks:  $2(64) + 2(128) + 2(256) + 2(512) + 2(1024) \dots$ , or  $128 + 256 + 512 + 1024 + 2048 \dots$ , the sum of a geometric series, also given as

$$a + ar + ar^2 + ar^3 + \dots + ar^{n-1} = \sum_{k=1}^n ar^{k-1} = \frac{a(1 - r^n)}{1 - r}$$

in this case  $r = 2, a = 128$  so

$$t = \frac{128(1 - 2^n)}{1 - 2} = 128(2^n - 1)$$

The exponential decay we would like:

$$y(t) = e^{-\lambda t}$$

expressed in terms of  $n$  is therefore

$$y(n) = e^{-128(2^n-1)\lambda}$$

giving us figure 5-6 which shows the gain vs partition number for an exponential decay (plotted both as quantized to partition number and smooth).

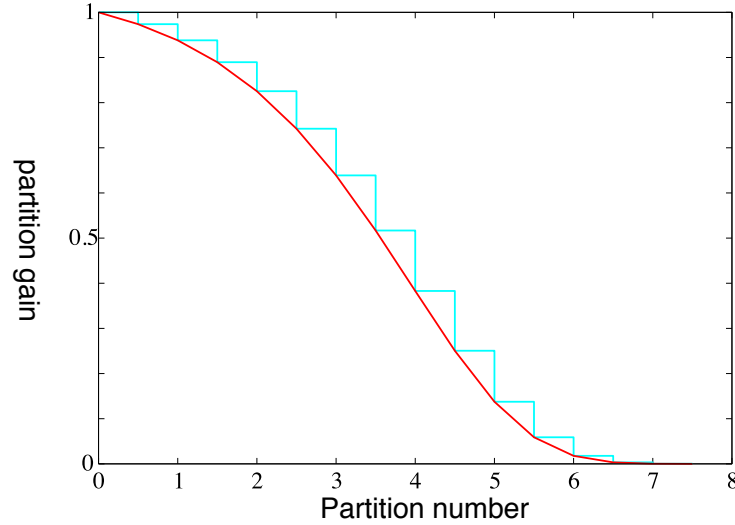


Figure 5-6: Exponential decay plotted vs partition number. Cyan (lighter) plot is the value quantized by partition.

Transitions between the block gains can introduce artifact, but is usually not audible, and using a Hanning window instead of a square window can remove that artifact, but also increases the computational requirements. The steady state response can then be made to approximate any desired decay curve. However, things get more difficult when we look at the dynamics of changing damping.

### Dynamic continuity problems

Controlling the gains of each block gives a realistic-sounding damping at steady state. Unfortunately, changing the damping causes considerable artifact.

Figure 5-7 shows two decay curves, the red plot indicates the original decay of the instrument, and the green shows what we would like the decay to be at steady state. Ideally, at the time the new damping is applied ( $t_1$  in figure 5-8(a)), the decay curve should continue

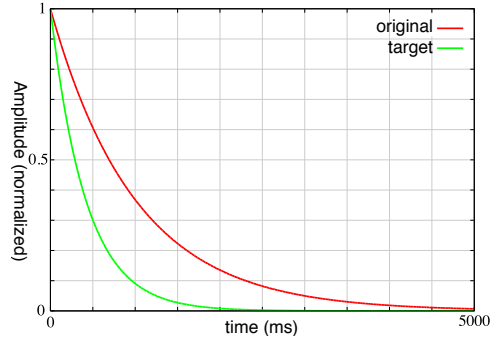


Figure 5-7: Original and target decay envelopes

from its current level, but with a new damping factor.

Ignoring for a moment the stair-step response caused by the block convolution, if the new damping curve is applied at time  $t_1$ , the output jumps to the level that the system would have been at if it had been using the new damping coefficient at the time when the object was originally struck (figure 5-8(b)).

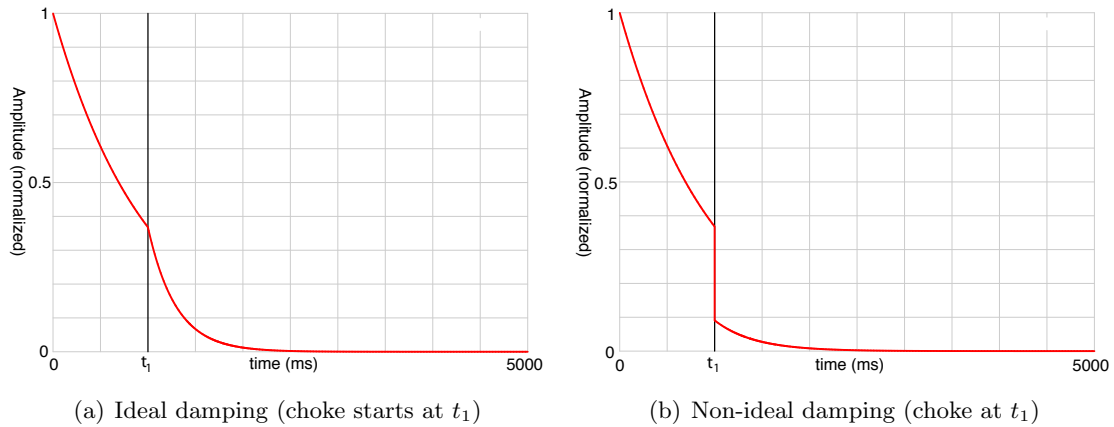


Figure 5-8: Ideal damping, and less-than-ideal damping.

However, by cross fading between the two curves, the discontinuities due to switching damping coefficients can be minimized (figure 5-9). Neither the linear nor the quadratic cross fade are very good fits, but the main goal is to minimize transients during the transition. For all subsequent hits, the actual decay curve will match the target curve.

## A second dynamic problem: undamping

While using the above method to control the gain of the output of each convolution partition results in an immediate change in the decay curve, it exhibits quite unrealistic behavior when

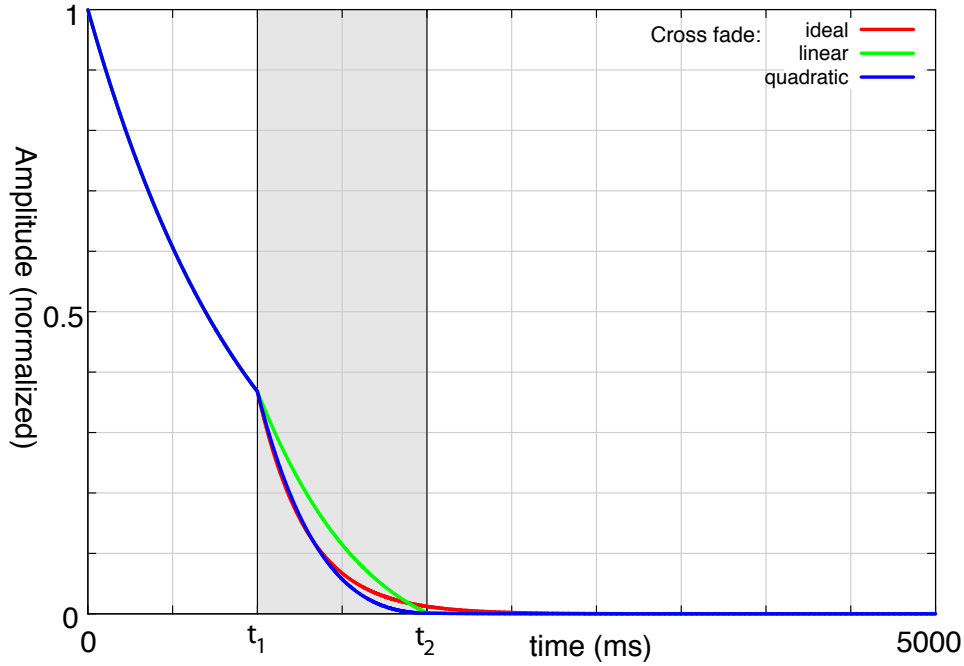


Figure 5-9: ideal vs linear and quadratic cross fades (in the region between  $t_1$ , and  $t_2$ )

undamping.

Striking a real cymbal while holding on to it will result in a short decay. Let go of the cymbal, and it will continue to decay with its previous un-choked time constant. In our virtual cymbal when we only control the output gains, if the player releases the cymbal before it is completely decayed, the level jumps back to the previous decay curve, creating an unnatural echo. Figure 5-10 shows such a situation where additional damping is applied at time  $t_1$  and released at time  $t_2$ .

If there are additional hits that happen while the system is damped, when the player releases, the output jumps to the accumulated volume of those hits, just as if the system had never been damped to begin with.

One partial solution is to decrease the gain of each convolution partition at its input as well as at its output (figure 5-11). This would completely eliminate the echo as long as the damping is held for the duration of the 2x the longest partition, typically 4096 samples (93 ms). Any changes made to the gain at the beginning of the partition (say at time  $t_1$ ) won't be heard until the convolved result emerges from the partition at time  $t_1 + \delta$

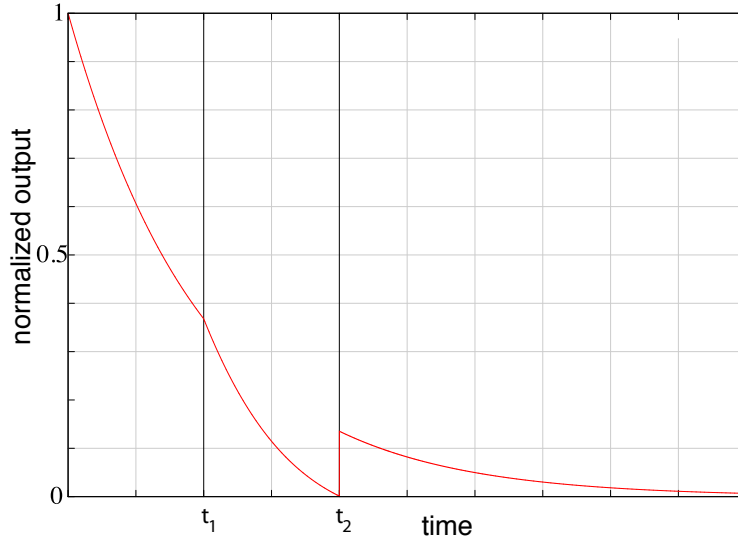


Figure 5-10: What happens when you stop damping?

$G_i$  = partition input gain

$G_o$  = partition output gain

$G_t$  = desired total gain

To achieve a total gain  $G_t$  at steady state,

$$G_i, G_o = \sqrt{G_t}$$

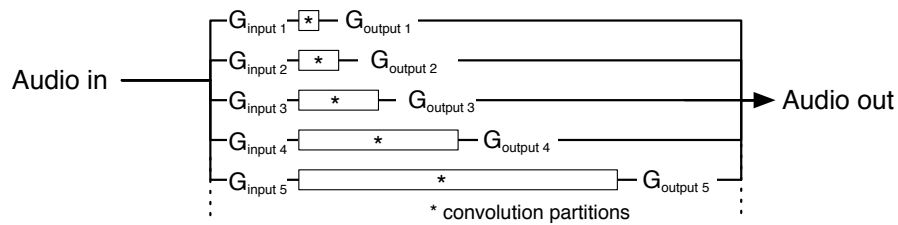


Figure 5-11: Schematic with input and output gains

One advantage is that the longest partitions processing the end of the impulse also are already at the lowest volumes, minimizing the significance of any artifact. However, although both the input and output gains are reduced immediately, because of the latency due to the FFT of each partition only the effect of the output gain is perceived immediately, while the change in input gain becomes audible one partition size later. This actually causes the overall gain of the partition to go through two different reductions. as long as the gains

are nonzero, resulting in the “double ramp” shown in figure 5-12.

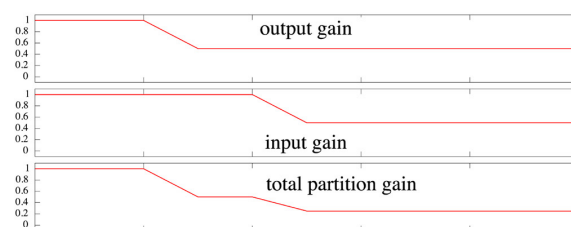


Figure 5-12: “Double ramp” effect due to the FFT delay

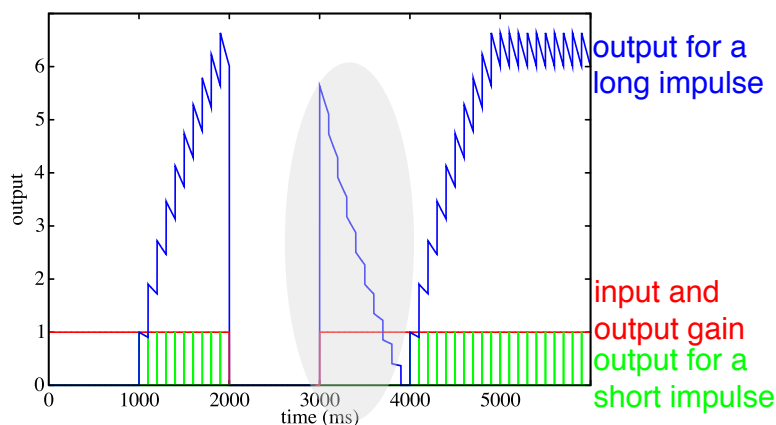


Figure 5-13: 10 strikes per second: artifact for long stored impulses when  $G_i = G_o$

A bigger problem with using the same input and output gains comes when the system is muted for less than the sum of partition size plus the length of the stored sample in that partition (usually occupying the whole partition).

Consider only one partition with a latency of 1000 ms that is receiving 10 strikes per second starting at  $t = 0$  (figure 5-13). The output gain is shown plotted in red, and is either 1 or 0 for the sake of simplicity. We first hear output at  $t = 1000$  ms. If the sample is very short, it tracks the output gain times the delayed input gain (green line). But if it is longer, it slowly builds up (blue line) which matches the behavior of real instruments. When the system is muted, output goes to zero as expected, but if it is un-muted before two partitions have elapsed, the output jumps to the level that is still decaying inside the convolver. After that short burst, it behaves properly again, slowly building up as seen at  $t=1000$ .

We do better if we set  $G_o$  to be the minimum of the target gain  $G_t$  and the input gain  $G_i$ ; however, there is still artifact if the duration of muting is less than 1/2 of the partition

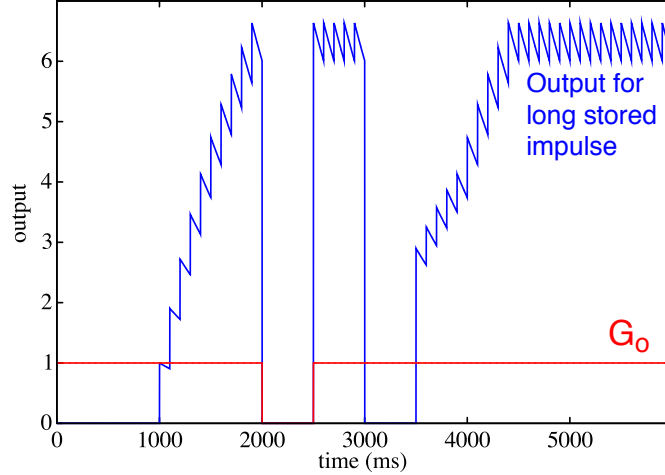


Figure 5-14: Artifact when  $G_i = G_t$ ,  $G_o = \min(G_t, G_i(\text{delayed}))$

duration, as seen in figure 5-14.

Here  $G_o$  is plotted in red, output drawn in blue. The decay seen in figure 5-13 is masked by the output gain, but there is still a burst of noise between the two mutings. This problem is solved by making  $G_o$  equal to the minimum value of  $G_i$  over the duration of the partition:

$$G_o(t) = \min_{t-\delta \leq \tau \leq t} G_i(\tau)$$

However, this solution just reveals yet another problem. Figure 5-15(a) shows the output of one partition of the convolver (4096 samples, or 93 ms of latency) where the gain is reduced from 1 to 1/2 for a duration of 500 ms. The result for input hits at less than the partition frequency, every 100 ms, is shown in red. When the inputs are below the partition frequency, the output does not build up, since the result of each hit stops playing before the next hit occurs. If the hits are above the partition frequency, (every 2 ms, shown in green, normalized to fit on the same graph) the outputs do accumulate. At both input frequencies, there is a step artifact due to the lag in changes to the input gain propagating through to the output. This lag is equal to the partition duration.

For infrequent (less than the partition frequency) inputs, this artifact can be removed by setting the output gain to be equal to the minimum of the input gain (over the duration of the partition) divided by the delayed input gain:



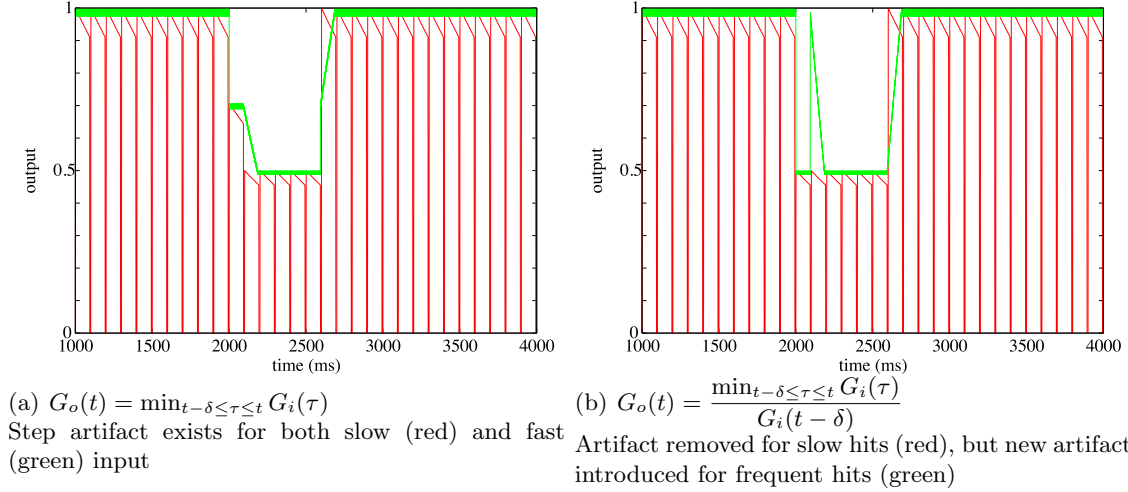


Figure 5-15: Fixing one artifact introduces another.

$$G_o(t) = \frac{\min_{t-\delta \leq \tau \leq t} G_i(\tau)}{G_i(t-\delta)}$$

Figure 5-15(b) shows the result in red. The artifact is removed successfully for infrequent hits. Unfortunately I had not considered the effect of more frequent hits. When hitting clearly with a stick, I heard no artifact when damping and undamping, but stirring with brushes while changing damping created a series of pulses. The cause of this intermittent artifact was not clear until I attempted to plot the output for more frequent hitting. The output for frequent (every 2 ms) hits is shown in figure 5-15(b) in green. There is a significant spike 93 ms after the muting begins which is only apparent for frequent hits, due to the accumulation that occurs because the period between hits is faster than the duration of the stored impulse.

So although the artifact is completely removed for infrequent hits, the new artifact generated for frequent hits is much more objectionable due to its spike shape and sharp transitions which make it sound like a click. In practice, however, the artifact in figure 5-15(a) is not readily apparent, and is mitigated by slowing down the rate of change of muting. If it is slowed to the partition duration or slower, it disappears completely (figure 5-16).

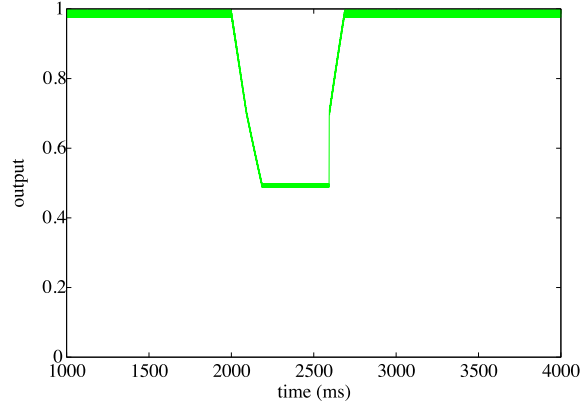


Figure 5-16: Slowing down the muting rate limits or removes the artifact for fast hits

### 5.1.2 Frequency-dependent damping

The muting mechanisms by controlling partition gain described so far act equally on all frequencies. However, viscous damping acts more strongly at higher frequencies, so we would like to implement a faster decrease in high frequencies than in low ones.

Although losses in real materials occur through a variety of complex mechanisms, they can be approximated as the sum of viscous and frequency-independent losses [67]. As in the case of frequency-independent muting, latency and block size are still going to introduce some artifact, and although the ideal steady state solution would be to filter the recorded impulse, the latencies involved in changing the filter are again too long to give a convincing result.

In viscous damping, any particular sinusoid will decay as an exponential, and at any particular time, the rates of decay will increase exponentially as a function of frequency such that sinusoid gain  $\propto e^{-\lambda f t}$ , shown in figure 5-17.

For ease of calculation, the exponential frequency curve will be approximated using a one-pole filter by matching their -3dB points. For the exponential,  $y = e^{-\lambda f}$ , the -3dB point is half the power, or  $\frac{1}{\sqrt{2}}$ . The equivalent cutoff frequency  $f_0 = \frac{\ln(\frac{1}{\sqrt{2}})}{-\lambda}$ . The filter response is shown in figure 5-18.

**Minimizing artifact when changing the damping values** As in the case of frequency-independent muting, changes to the filter cutoff at the input to each partition take one

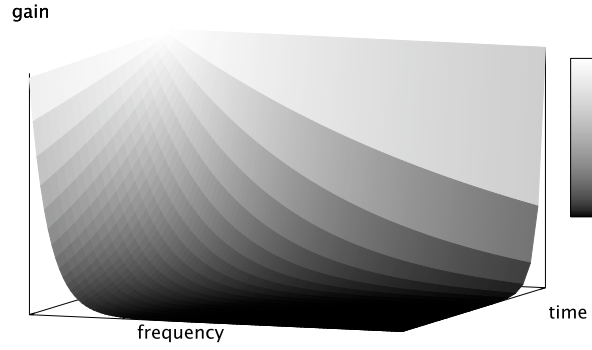


Figure 5-17: Frequency-dependent damping

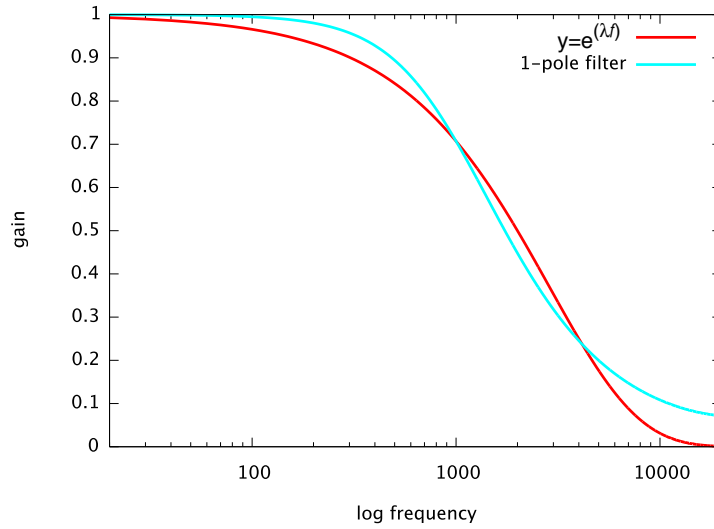


Figure 5-18: Aligning the -3dB points of a one-pole filter and an exponential

partition length to be heard. Similarly, we can temporarily apply another filter at the output, and set its cutoff to be the minimum value of the input filter cutoff over the duration of the partition ( $t - \delta \leq \tau \leq t$ ). Figure 5-19 shows the system schematic.

$$F_o(t) = \min_{t-\delta \leq \tau \leq t} F_i(\tau)$$

The amounts of frequency-dependent and frequency-independent damping can be controlled independently, enabling the player to dial in a particular default decay profile, and

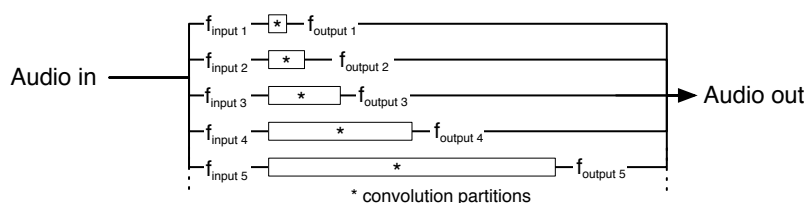


Figure 5-19: Schematic with input and output filters

also control the effect of choke and pressure sensors (described in chapter 6) to allow for intermittent, expressive damping. For a stored impulse response like a cymbal, increasing the frequency-independent damping results in a dryer sound, more like a change in the properties of the cymbal itself, while increasing the frequency-dependent damping sounds as if the player was applying a manual choke.

Both systems can also be abused to provide progressively larger boosts as the stored impulse decays, giving much brighter or simply extended decays relative to the original recording. Similarly, crude multi-tap and tremolo effects are also possible simply by controlling the partition gains.

### 5.1.3 Pitch shifting

Some drums, such as timpani and many hand drums, allow for changes in the tuning of the head. Since we only have a sample of the instrument to start with, and not a physical model, we can't simply vary model parameters to gain the new pitch. Further complicating matters is that unlike in a digital sampler with which a sample can be played out slower or faster to achieve tape-style pitch shifting (figure 5-20), we are stuck with partitions that have a fixed duration. Slowing down or speeding up the playback of a partition, or stretching its spectrum will result in gaps or discontinuities at the partition boundaries. Shifting the partitions in time to accommodate and conceal these gaps would also require an additional partition's length of latency. Using Hanning or raised cosine windows instead of square windows hides the gaps, but at the expense of doubling the computation.

One advantage of working with percussion sounds is that they are largely non-harmonic, which allows the use of spectrum shifting to achieve changes in pitch. The chief advantage

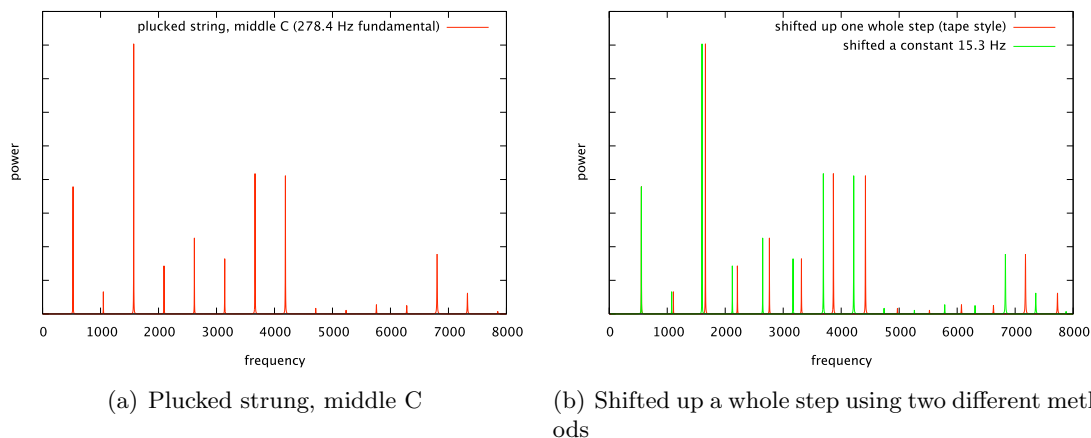


Figure 5-20: Tape-style pitch shift vs. fixed-frequency shift

of this method is that the timing remains constant while the pitch changes. The primary disadvantage is that the spectrum is shifted by a fixed number of Hz, so the ratios of frequencies do not sound constant. For example, a plucked string has overtones that are multiples of its fundamental. Shifting the string spectrum will cause those overtones to no longer be multiples of the fundamental, giving a more metallic, non-harmonic sound. Luckily, many percussion sounds lend themselves to this kind of manipulation due to their lack of aligned harmonics.

In the convolution percussion system, due to efficiency constraints, I have primarily used spectrum shifting to achieve changes in pitch. Since this is operating on the stored FFTs of the impulse response, there is still some latency (half of a partition length) to hear the pitch change effect. For very fast pitch changes, this is an audible artifact. Limiting the rate of pitch change and limiting the maximum partition size helps control this artifact. Shifting the spectrum of the input has surprisingly little impact on the output sound for relatively broad band input, but might be useful in limiting audio feedback.

A second approach is to perform the pitch shifting on the output only. For this purpose, any of the established pitch shifting algorithms can be applied, with the usual tradeoffs of latency, jitter, and artifact.

### 5.1.4 Cross fading

To perform cross fades, the most straightforward method is to literally cross-fade the pre-transformed stored impulse with another. This works for very slow fades, but as with damping and pitch shifting, it falls apart for faster manipulations. There are several other options, all have their good and bad points.

**Parallel – gain set at outputs** In this method, two convolvers are going all the time, and there is a simple cross fade of their outputs. The effect is one of switching between listening to two different instruments that are ringing down differently. Unless the sounds are very similar, there is not fusion into one instrument.

**Parallel – gain set at inputs** This method gives each convolver time to ring down when the input is switched to the other. This primarily gives the impression that the player is switching between playing two instruments, or two distinct parts of one instrument.

**Series** Connecting two convolvers in series raises additional challenges for where the control should occur. One option is to leave the first convolver on all the time, and control how much signal goes through the second convolver, either by controlling its input, output, or both. When both are engaged, only frequencies in common to the input and both stored impulses will pass through. For full cross fading between convolvers, something like the system in figure 5-21 is required. When  $G$  is around 0.5, both convolvers are active, and frequencies in common are boosted, but also some signal is still allowed to bypass each convolver.

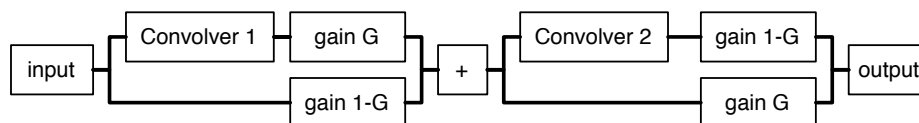


Figure 5-21: Schematic for crossfading two convolvers in series.  $G$  goes from 0 to 1

### 5.1.5 Inverse filtering

Since the spectra of the object being struck and the stored sample are being multiplied, when there are strong resonances in both signals overlap at the same frequencies, there can be extremely high output at those frequencies, often causing feedback, clipping, and an unpleasant sound. Also, to make the resulting sound more resemble the original recording, some of the physical controllers need their output to be equalized to boost or cut the bass or treble to flatten out their response. It would be useful to be able to automate this process through an inverse filter.

There are several ways to achieve this. The first attempt was to deconvolve a recording of a typical hit from the stored sample, so that when the new sample is reconvolved with another hit, the output is flattened, and any timbral variations center around the timbre of the original recording. The advantage of deconvolution over other filtering methods is the high resolution of the filter, which requires only a recording of the sound to be filtered out. If the hit is typically excessively bass-heavy for example, bass is removed from the stored sample. The Fourier transform ( $\mathcal{F}$ ) of the instrument recording was divided by the Fourier transform of a typical hit. For signals  $y$  and  $h$ :

$$h \text{ deconvolved from } y = \mathcal{F}^{-1}[\mathcal{F}[y]/\mathcal{F}[h]]$$

This worked fine for some samples and physical objects, but when the object had no output at a particular frequency we would be dividing by nearly zero, resulting in a large peak. If the subsequent hits were identical to the reference hit, then the peak would be flattened again, but any small variation in the spectra of the hits resulted in sharp resonances, and any noise in that band was boosted.

In one specific case, I attempted to deconvolve the sound of tapping on my laptop (picked up by the internal mic) from a recording of a frame drum. After deconvolution, the sample had a sharp metallic resonance almost like a cowbell. After reconvolving with more taps, the output still maintained a strong metallic coloration. Figure 5-22 shows normalized log plots of the original frame drum sample (red), the hit (green), and the result of the deconvolution. The frequency axis is from 0 to 1 FS/2 (in this case FS=44100 Hz).

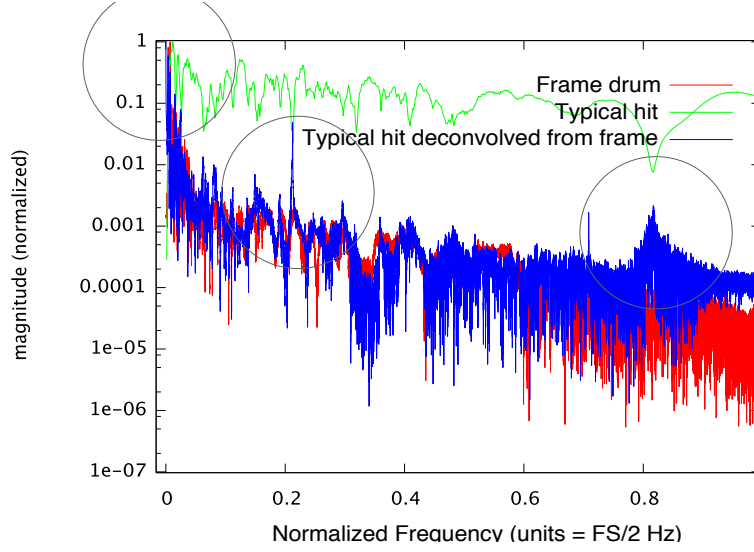


Figure 5-22: Deconvolving a typical hit from a frame drum sample. Artifacts are circled.

Some problems with this technique are apparent. In figure 5-22, Around zero FS/2, there is a large spike in the deconvolved output (blue). There are two others around 0.21 and at 0.7 FS/2 (figure 5-23).

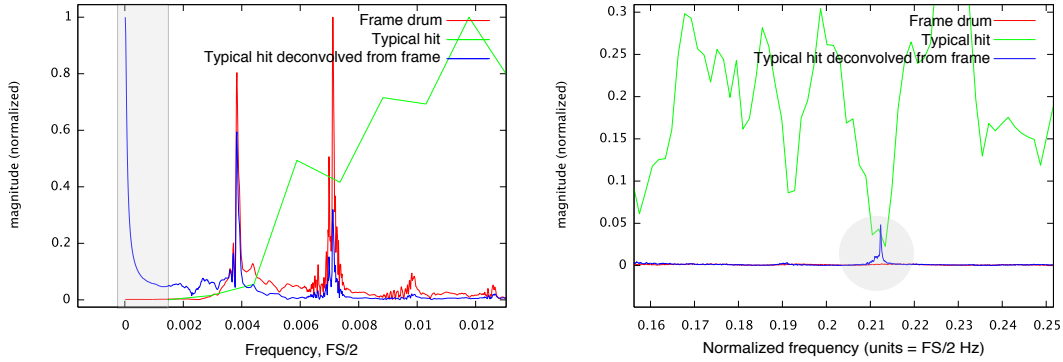


Figure 5-23: Spikes at 0 FS/2, 0.21 FS/2

After converting the complex output of the Fourier transform to phase ( $\Theta_{\mathcal{F}[h]}$ ,  $\Theta_{\mathcal{F}[y]}$ ) and magnitude ( $|\mathcal{F}[h]|$ ,  $|\mathcal{F}[y]|$ ), an offset ( $C$ ) can be added to the magnitude of the hit.

$$h \text{ deconvolved from } y = \mathcal{F}^{-1} \left[ \frac{|\mathcal{F}[y]|}{|\mathcal{F}[h]| + C \max(|\mathcal{F}[h]|)} e^{(\Theta_{\mathcal{F}[h]} + \Theta_{\mathcal{F}[y]})i} \right]$$

Good results were achieved with  $C = 0.1$ , providing sufficient filtering to flatten the output while avoiding the artifact of the previous method.



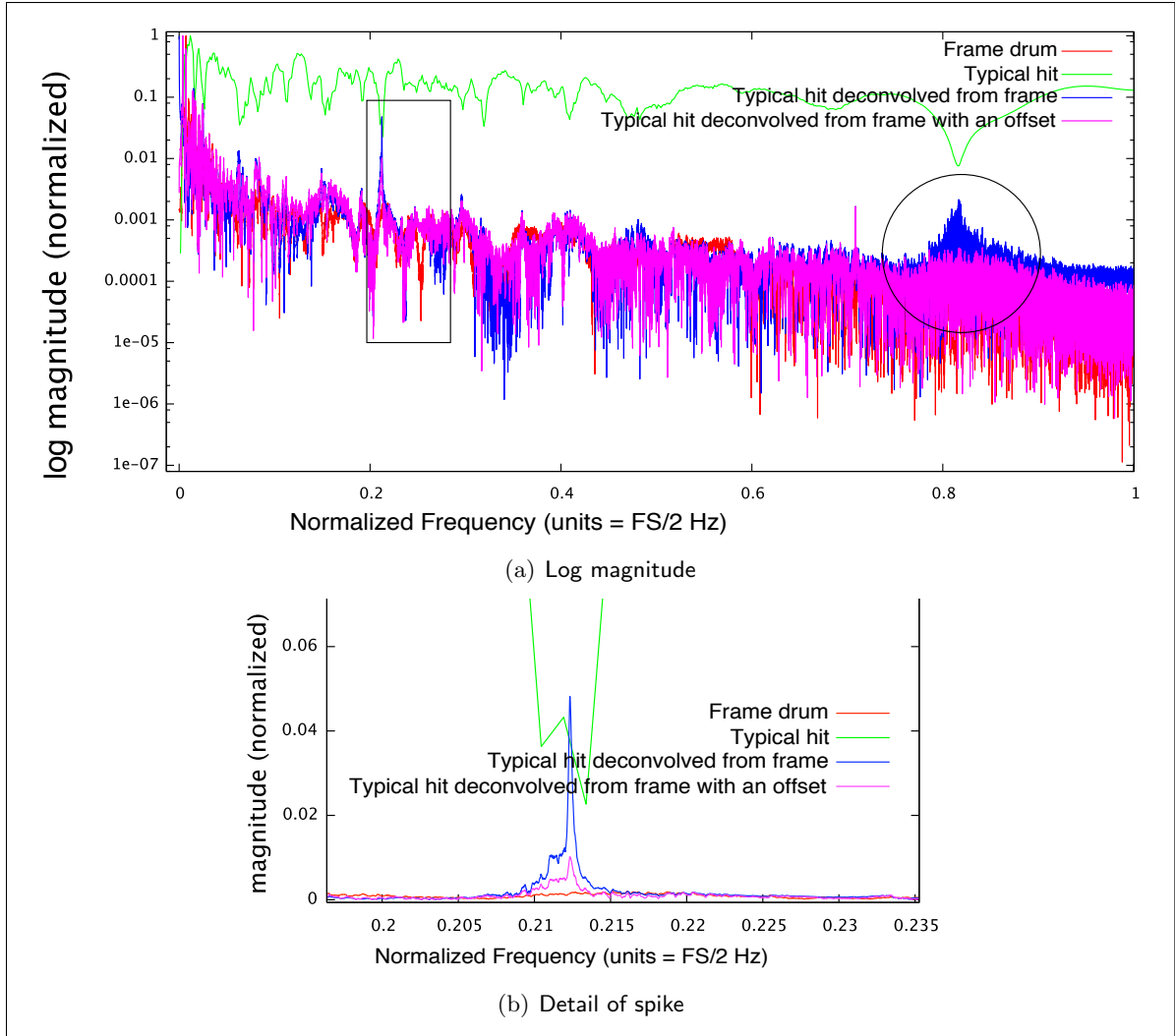


Figure 5-24: Deconvolution with offset in the denominator (shown in magenta)

### 5.1.6 Crashing a ride cymbal: pseudo- nonlinear processing

As was discussed in section 4.4, Cymbals exhibit nonlinear transitions between regimes. While convolution can emulate the response within a particular regime, the transitions are problematic. For example, playing a real ride cymbal with progressively louder hits will bring out more dense harmonics as the total output increases. With the convolution system and a single ride cymbal sample, there is no way to obtain modes other than what was already in that recorded sample. To get around this problem, some knowledge of the real system is required, and a solution will have to be customized for a particular application.

To approximate the cymbal crash, two convolutions were performed simultaneously.

Sound from the taps was waveshaped by an exponential to increase its harmonic content with increasing amplitude and then convolved with a ride cymbal sample. The output of the first convolution was then assymetrically clipped, and output above or below the clipping thresholds was then convolved with a sample of a cymbal being crashed (figure 5-25).

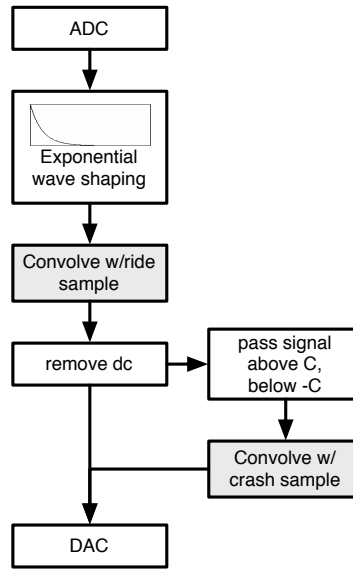


Figure 5-25: Crashable cymbal system diagram

By sending only signal above a clipping threshold, discontinuities were introduced in the signal that produces many more frequencies than were in the output of the first convolver (figure 5-26).

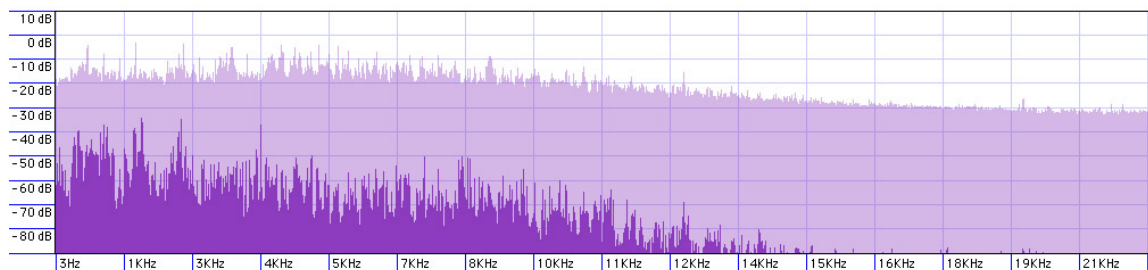


Figure 5-26: Spectra of soft and loud hit

## 5.2 Summary

This chapter has focused on several methods to provide realtime control over the sound. Because of the limitations of block size, artifact occurs when the parameters are varied

faster than the block size. This highlights a bandwidth constraint of the system; to limit such artifact, changes to the controls over the sound have to be slowed to a time constant closer to the block size. As computers get faster, this will be less and less of a problem, since all blocks will be able to be smaller.



## Chapter 6

# Physical controllers

Because of the nature of the processing, the physical part of the instrument is at least as important as the algorithmic part. For this chapter I will refer to the physical part of the instrument as a *controller* for convenience, though its acoustic properties and conception differ from typical controller schema. These controllers are designed to exploit the fact that the convolver is acting as a resonator. By varying the degree of damping, physical resonances can be removed and replaced with any desired resonance.

The controllers described in this chapter can be represented on a continuum based on the degree to which their own acoustics influences the output. At one extreme, the practice pad controller is highly damped, and although it does impart a “plastic” sound, it is a minor coloration. In the middle, the various brush controllers give a clear impression that the stored impulse is being performed with a brush, taking on the dense time texture of the metal tines. At the other extreme, the cymbal controller provides significant coloration to any sound, enough so that it can sound like a cymbal bolted to a bass drum, or a cymbal attached to a snare.

### 6.1 Cymbal

This cymbal controller started out as a budget brass student cymbal, and it is designed to accommodate normal cymbal playing gestures such as hitting the bell or shell and choking

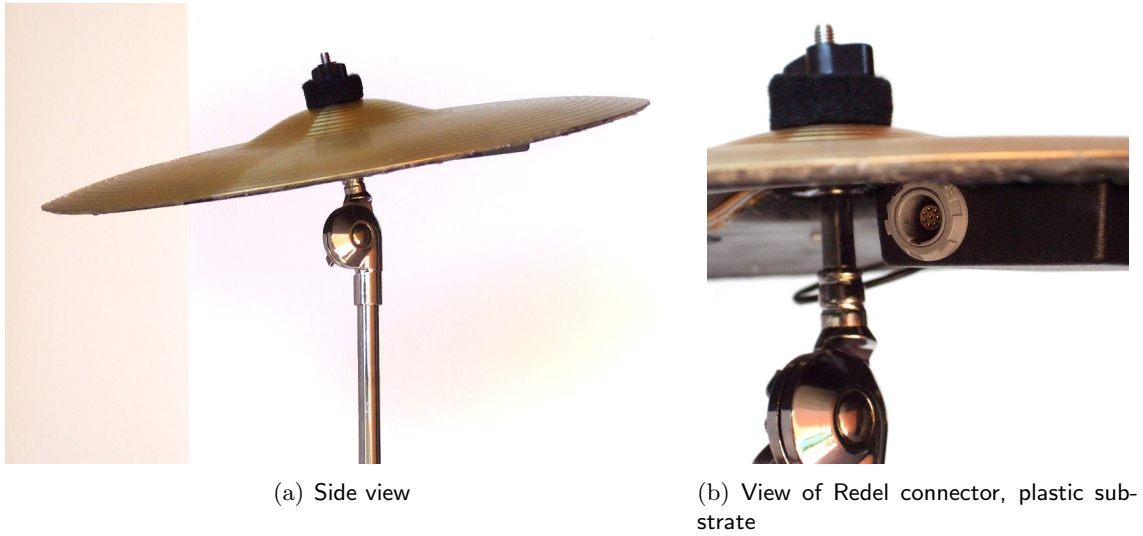


Figure 6-1: Cymbal controller

the cymbal by grabbing the front of it. Since it is built around a modified real cymbal, it can sit on a standard cymbal stand. A foam grommet limits contact with the cymbal stand, and allows the cymbal to swing normally.

### 6.1.1 Assembly

The cymbal controller is assembled in layers, from top to bottom, as shown in figure 6-2, the layers are:

- A real brass cymbal
- PVDF element (MSI FLDT1-052K [44]) bonded to the cymbal underside, away from the playing area
- A thin foam layer to damp the cymbal and transfer choke force
- Force sensing resistor (interlink #406 [35]) to detect choke force at edge of the playing surface
- Molded plastic cymbal substrate (pintech XT practice cymbal [58]) to support the assembly and further damp vibration

The edges are sealed with silicone caulk. The FSR is connected directly to the computer audio interface, and audio signals are sent through the FSR and change in the signal levels

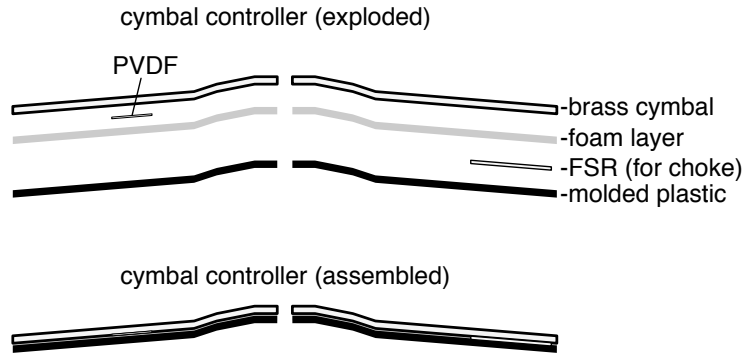


Figure 6-2: Cymbal controller assembly

is measured to determine the sensor's resistance. An advantage of this approach is that no additional hardware is needed, but it does take up another channel of input and output 6-3.

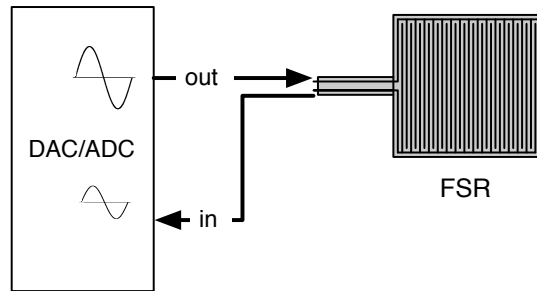


Figure 6-3: Using the audio interface to measure the resistance of the FSR

The signals used have been in the 150 – 500 Hz range to minimize capacitive coupling while maintaining sufficient time resolution for controlling the damping.

### 6.1.2 Function

Since there is significant spectral contribution from the cymbal, hits on the bell, rim, or edge sound substantially different from each other. I originally expected to need multiple contact microphones to get enough variation from hits in different locations, but it turns out that one microphone is sufficient because of the range of sounds achievable by hitting different parts of the cymbal. When convolving with a cymbal sound, the effect is that the lost resonance of the cymbal (due to damping) is restored, and it is quite surprising when the processing is turned off to hear that the real cymbal only sounds like a dull clank.

One drawback to allowing the controller to provide more of the spectrum is that while it heightens the realism of cymbal sounds, it will always impart a cymbal-like quality, even to non-cymbal sounds. For example, when convolved with a concert bass drum sound, the output sounds as if a cymbal was somehow joined to the drum head.

**Extensions** To allow for cymbal crashes, two convolvers can be chained (described in section 5.1.6 , approximating some of the nonlinearity of the real cymbal (section 4.4). In addition to the FSR circuit, the surface of the cymbal was also electrically connected to the audio interface to pick up the 60Hz hum from when the player touched the surface. The hum was filtered and the envelope was used to control damping. Even though it provided essentially only one bit of data, having the cymbal be sensitive to damping over its entire surface proved to be more important than having a range of damping in one location. A potentiometer knob was added to the top of the cymbal (figure 6-4) to control pitch. The knob's resistance was measured by Pd using the same method as for the FSR. This allows the player to dial in a particular cymbal sound from the cymbal itself.



Figure 6-4: Cymbal pitch controlled with knob

### 6.1.3 Brushes

Two kinds of brush controllers were developed for use with this system, one wireless and one tethered. Drum brushes in either configuration were fitted with a PVDF contact microphone



to pick up the sound in the metal tines. Any surface can be played with the brushes, and the resulting output sounds as if the sampled instrument is being played with brushes, but has the texture of the surface being played. By stirring the brush on a surface, a sustained broad band noise can be produced that results in quite different timbres than were observed with the pads or cymbal controller. Different combinations of surface textures, brush movements and stored impulse are possible.

#### 6.1.4 Wireless brush



Figure 6-5: Wireless brush

The wireless brush used the circuit board and part of the enclosure of a handheld VHF wireless microphone (Nady DKW-1H [49]) to transmit its audio signal. Up to four wireless brushes can be used simultaneously on four different VHF channels. As with the wired brush, a piece of PVDF (digikey MSP1006-ND) was threaded through the tines to pick up the brush sound. Kapton tape was used to protect the piezo element from abrasion from the brushes.

#### 6.1.5 Wired brush



Figure 6-6: Wired brush

The wired brush controller started with a rubber-handled drum brush, and added a 3” bend sensor [1] to detect when the brush was pressed against a surface. The bend sensor was placed in line with the tines, while a PVDF tab was threaded through the tines. The rubber covering was split to make room for the wiring.

A Redel-compatible connector was added to the end of the brush to allow quick connections to a multiconductor cable. This connector was common to several of the controllers built, allowing easy interchangeability.

## 6.2 Pad



Figure 6-7: Percussion pad controller

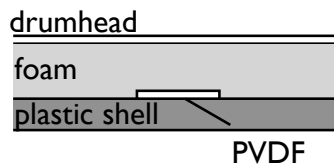


Figure 6-8: Percussion pad cross section

This is a simple controller derived from a drum practice pad. Since one of the goals of a practice pad is to be quiet, it was already well damped. A piece of PVDF foil [45] was applied under a layer of foam located beneath the drumhead and was connected directly to the audio interface (figure 6-8).

The pad proved a surprisingly versatile controller, working well with most impulses. Due

to the head material and the high degree of damping, treble had to be boosted to maintain reasonable sound. Unfortunately this also made it much more susceptible to noise.

The practice pad had a somewhat uniform sound due to its thick plastic head and highly damped design. The foam itself made some noise when compressed, creating unrealistic artifact for loud hits. Players had to work to produce a meaningful range of impulses. Sanding the head helped the sound somewhat, as did maximizing the tension of the head. Hitting the metal tension ring around the perimeter of the head gave more of a metallic clank, which was quite different from any of the sounds achievable by hitting the drum head.

### 6.3 Frame drum



Figure 6-9: Frame drum controller

Based on the preliminary results of early versions of the cymbal controller, I wanted to apply the same technique of using more of the acoustic response of the physical object to the construction of a drum controller. Starting with a wooden frame drum, I added contact microphones, damping material, and pressure sensors (figure 6-10). This drum was much less damped than the practice pad, ensuring that more of the spectrum of the drum was carried through the processing.

Drums struck in different locations can excite different modal structures. For example,

striking location helps create the differences between Djembe bass, tone, and slap sounds. Unfortunately, the convolution system is stuck with the one set of modes that are in the sampled sound. One way around this problem is to run multiple convolutions at once, and to have contact microphones at multiple locations on the drum head, or one could also track the location of the hit and control a cross fade. In this case, multiple contact microphones were used to be able to process hits on the center and edges of the drum differently.

### 6.3.1 Assembly

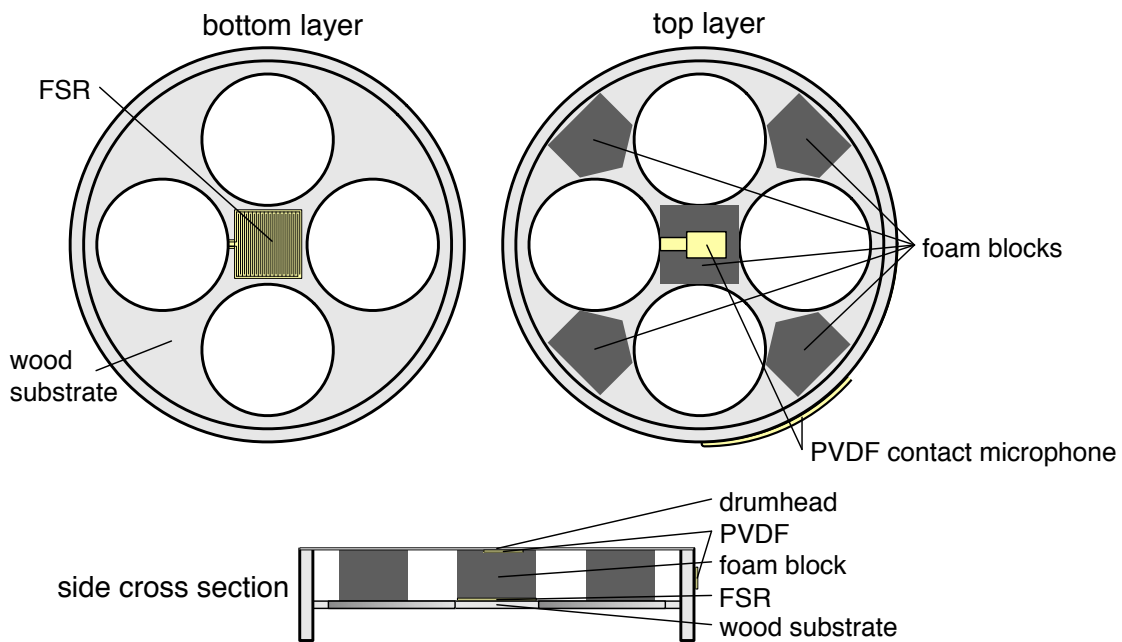


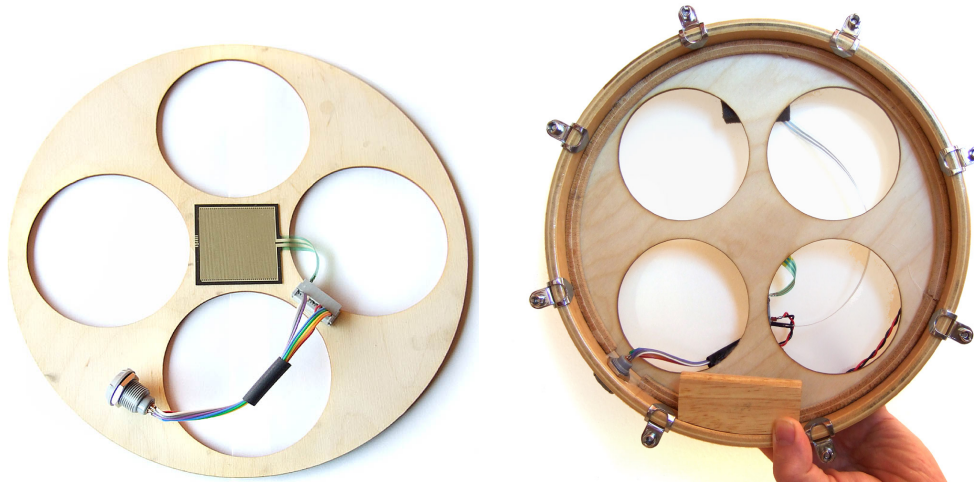
Figure 6-10: Frame drum controller assembly

One PVDF element (MSI FLDT1-052K [44]) was mounted to the underside of the center of the drumhead, and another was mounted to the frame.

A force sensing resistor (“FSR”, interlink #406 [35]) was mounted to a wooden substrate at the center of the drum and covered by a foam block to provide control of damping. The compressibility of the foam block allowed for a greater displacement of the drum head over the active range of the sensor, and also served to protect the sensor by spreading any forces over its whole area. The order at the center, from top to bottom is:

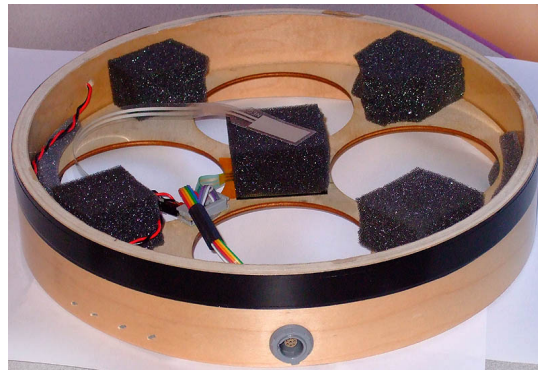
- Drumhead

- PVDF
- Foam block
- FSR
- Wooden substrate



(a) Substrate w/ FSR, connector

(b) Bottom view



(c) Drum without head

Figure 6-11: Frame drum controller

### 6.3.2 Function

Since the FSR is mounted at the center of the drum, it responds to pressing anywhere on the drumhead (although much more strongly at the center). This gives good subtle control of damping by pushing at the edges, while still allowing sudden and immediate damping by pushing at the center.

Pushing on the drum head also raises the pitch of the drum slightly. Originally I had intended to have a small pitch bend controlled by a second pressure sensor, but for many drum sounds, there is enough of a pitch effect due to the changes in tension in the real drum head, even though the stored impulse is not shifted.

Separate processing of the rim signals from the center works particularly well for djembe sounds. Since there is an increase in low frequency output of the center PVDF sensor when it is hit directly, I found that I could combine djembe bass and tone sounds into one sample, and obtain more of one or the other entirely based on where and how the drum was hit, while using the edge sensor just for djembe slap sounds.

## 6.4 Bass drum with speaker



Figure 6-12: Bass drum controller

For this controller, I was interested in having the sound emit from the object, to provide a stronger illusion that the player was interacting with a physical object rather than a computer. I converted a bass drum shell into what is essentially a speaker cabinet in which the speaker is located behind the drum head. This provided both a sonic and tactile feedback to the player.

### 6.4.1 Assembly

The assembly is shown in figure 6-13. A circle of medium density fiberboard (MDF) was used to seal one end of the drum, and an MDF ring supports and centers a 15-inch bass

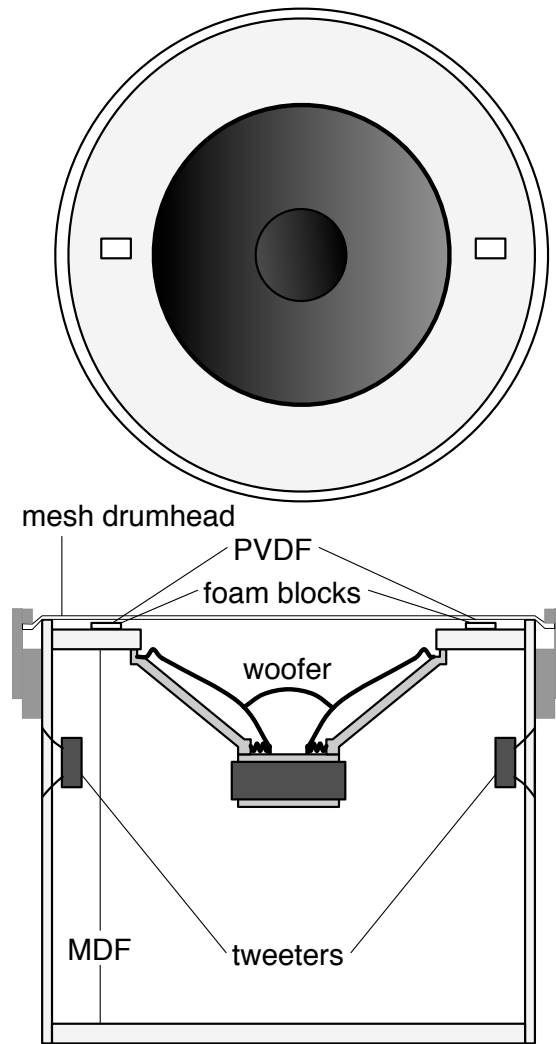


Figure 6-13: Bass drum controller

speaker at the other end. Internal MDF bracing was also added. A nylon mesh drumhead was stretched over the end with the speaker, allowing sound to pass through the mesh. Vibration in the mesh is picked up by two PVDF elements (LDT0-028K, [44]) supported by foam glued to the MDF ring.

Side-mounted piezo horn tweeters (Pyle PSN1167) were added to improve the system's high frequency response.

Audio output from the computer was routed to the speaker in the drum, semantically re-coupling the resonator to the playing surface, though thanks to the mesh they stayed essentially acoustically uncoupled.



### 6.4.2 Function

The bass drum controller, because of its appearance, loud output, and low bass extension, was well suited for the obvious role of large drum sounds, along with thunder, prepared piano soundboard, as well as for large gongs and cymbals. Due to the resonance of the mesh, some equalization was necessary to control feedback, making it an ideal candidate for using deconvolution to pre-filter a typical hit from the the stored impulses. One surprising outcome is that it is actually well-suited for snare drum sounds, provided that the head is given a high enough tension to provide proper stick bounce.

## 6.5 Scratch pad

## 6.6 Touching sound



Figure 6-14: "Touching sound". Photos courtesy of David Merrill

Media Lab graduate student colleagues Hayes Raffle and David Merrill have begun using my Pd convolution system in a series of explorations of playful objects. They constructed a broad range of physical objects with unusual sound interactions such as arrays of porcupine quills, a ridged plastic bracelet, and wood bristles all connected to contact microphones to allow users to explore the connection between the objects and the associated sound processing.

I have also been working with David Merrill on integrating a microphone into a brush



controller to allow the users to gather impulses by pushing a button, and then to use the brush to play the impulses like an instrument. This system currently requires a laptop to perform the sound processing, but it would be great if the processing could be integrated into the brush itself.

## 6.7 Enhanced car console switches

For this application, I was interested in how the same processing techniques could be applied to the sound of physical switches and controls. As a proof-of concept, the sound of switches in a car’s center console were processed with various resonances.

This can be seen as a physical instantiation of work to represent the acoustics of virtual objects in a computer interface. Darvishi et al [17] proposed using physical modeling to represent interaction elements of a graphical user interface (“GUI”). Their system modeled the sounds of spheres impacting various materials and geometries to aid visually impaired users to differentiate between the parts of the GUI by tapping on them.

The system as implemented in the car has so far only been used for aesthetic purposes, but could also be used as a secondary channel for user feedback about the state of the system. In some automotive interfaces, the function of a particular button or switch changes depending on the mode it is in. Each mode could have a unique sound to remind the driver of the mode. BMW’s I-drive controller, for example, provides different haptic feedback based on the mode it is in, changing detent positions and strength, resistance, and spring-return to suggest more familiar automotive controls.

## 6.8 A Human-Computer Interface application: augmenting *Pico* with sound

Working with James Patten, we augmented his Pico tabletop actuated display system [56] to provide audio feedback in addition to the tactile and visual feedback already present in his system (figure 6-15). Pico is a tabletop *tangible user interface* [36] in which users

can manipulate physical puck-like objects on a flat surface to represent data and processes. Graphical information is projected onto the surface and the objects. The objects themselves can be physically moved by the user, or under software control by an array of electromagnets located under the table.

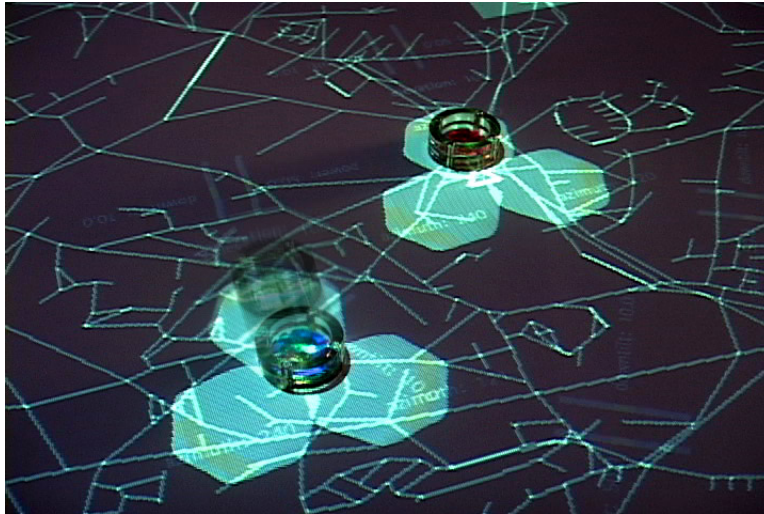


Figure 6-15: James Patten's *Pico* system with sound augmentation

The interface is based on a tabletop interaction surface that can sense and move small objects on top of it. Computation is merged with dynamic physical processes on the tabletop that are exposed to and modified by the user in order to accomplish his or her task. The system places mechanical constraints and mathematical constraints on the same level, allowing users to guide simulations and optimization processes by constraining the motion of physical objects on the interaction surface. The interface provides ample opportunities for improvisation by allowing the user to employ a rich variety of everyday physical objects as interface elements [56].

We were interested in providing each puck with unique *apparent acoustic properties*, for example to indicate that one was heavier than the other, or that they were made of different materials such as steel or stone.

Although Pico's control of the pucks could occur simultaneously, we chose to limit the number of simultaneous sliding pucks to make the relationship between pucks and sound more clear.

## Implementation

A single contact microphone was placed on the underside of the work surface (figure 6-16). The microphone picked up the sound of anything sliding on the surface. Since the system knows what is moving at any moment, the sound from the table can be sent through several different realtime convolutions depending on which puck is moving. Although we experimented with controlling the individual convolution gains based on the velocities of each puck, it turned out to work much better to use a simple rule of turning on the input to a puck's associated convolution when it was moving, turning it off when it was not. This still conveyed puck velocity because faster-sliding pucks created more sound. If multiple pucks were moving at once, then multiple inputs would be opened. The various convolutions were also statically panned in a stereo mix to make it easier to pick out the individual sounds. Ideally in the future we would like to dynamically control spatialization based upon each puck's position on the table to help fuse the sounds with the corresponding objects.

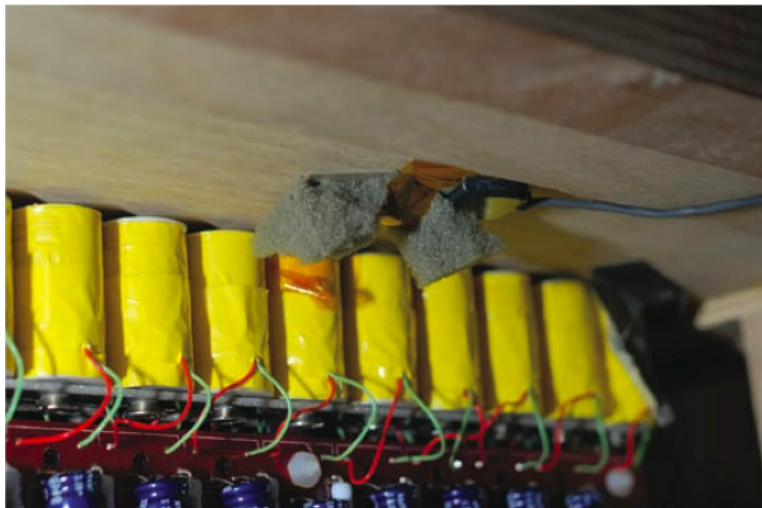


Figure 6-16: Contact microphone placement under actuation table (Photo credit: James Patten)

The sound processing was on a different machine than the rest of Pico, so motion data was sent via MIDI between the computers.

The most successful sounds were samples of a single scrape of different materials. Gravel, steel, stone, and plastic were sufficiently different from each other to provide useful feedback to the user. One idea that we had not pursued was to associate the sounds with regions of the table rather than individual pucks, so for example, as a puck traverses water, it could

make splashing sounds, or sound as if it is sliding through sand, grass, or mud.

## 6.9 Summary

In this chapter I have presented several different controller designs. Since the acoustic qualities of the controllers are so critical to their function, I think these represent just a tiny slice of what can be realized through extended development. In the same way that existing percussion instruments have constantly been extended and refined, the physical controllers in this system can also benefit from time and iteration.

The fundamental tension of the system is that for the output to sound exactly like the stored sample, you would like the input to be a perfect impulse with no timbral contribution from the physical controller. But for there to be sufficient variation in the timbre, the acoustic contribution of the controller has to be significant. In a system like this, the placement and design of the secondary controls such as pressure, bend, and touch sensors not only have to be consistent with the use of the instrument, but have to allow the controller to still function as an acoustic object.

The controllers I've built (excluding the non-musical examples) differ greatly in how their own acoustics influence the final sound. For the bass drum and pad, I saw that influence as a potential liability. The range of timbres was small, and the typical timbre had strong resonances requiring work through equalization and filtering to mitigate its impact. For the frame drum and cymbal, it was possible for the player to extract a much broader variation of timbre, giving an extra element of realism and variation to the final output.

The non-musical applications, particularly the sound extensions to Pico also begin to touch on what I see as an opportunity for almost any object we interact with to take on any apparent acoustic property. This could be used to represent hidden states of a system, convey low-priority information, and provide another degree of freedom for designers to explore the apparent quality of materials. Simple sample playback or synthesis is sufficient for discrete interactions like switches, but for less-constrained inputs, there is potential application of these techniques beyond the instrument domain.

## Chapter 7

# Evaluation

Many of the new instruments created in the field of computer music exist only for the designer. Some are only shown at conferences, or even more likely, a video is shown at the conference. Many are ephemeral.

The real evaluation of a musical instrument is how it gets used – whether it catches on, is a brief fad, or never leaves the inventor’s workshop. The lasting value of an instrument takes many years to identify, and frequently the most popular use is entirely unforeseen by the designer. The most important thing a designer can do to know if the approach is successful is to get the instrument into the hands of players.

To that end, during the development of these instruments, I have sought the input of multiple percussionists and gathered their responses and impressions. The feedback has been important both to evaluate the work and to guide its ongoing development, and I feel it has done the most to improve the instruments and to understand their strengths and weaknesses. Though the results of this kind of interaction are fundamentally anecdotal, they give more insight into the nature of the instrument-player interaction. These results are discussed in section 7.2.

Wanderley and Orio [94] suggest using the tools of Human Computer Interface design to evaluate musical input devices, though they do express some ambivalence about the significance of quantitative measures of a musical device.

The question here is whether this measurement must necessarily be quantitative, as in the case of HCI. In music, it must be noted that controllers cannot be evaluated without taking into account subjective impressions of performers, ruled by personal and aesthetic considerations. In fact, when skilled performers try a new instrument, rarely is a quantitative measurement of the instrument’s characteristics the goal.

That caveat aside, they outline seven possible contexts for evaluation, of which only the “instrument” context is relevant to this thesis. In that context, the player can control parameters of the sound through real-time gesture.

They break down the usability of a musical input device into four areas:

- *Learnability* – How long does it take to gain the skill necessary to perform a particular task.
- *Explorability* – The degree to which the capabilities of the controller can be recognized by the player through exploration. This relates to the consistency and precision of the sensing system as well as the particular mapping.
- *Feature controllability* – The degree to which the player feels they can control perceptual features of the output.
- *Timing controllability* – The precision to which a player can match timing.

Of these, the concept of *Explorability* is most relevant to the difference between this technique and simple triggering. Wanderley and Orio describe explorability as best measured by asking subjects to replicate expressive musical sound examples using the interface.

Measurement of *feature controllability* depends on the subjective response of the players, but that subjective response can be compared between different algorithms.

*Timing controllability* is strongly affected by system latency and the variability of that latency. Latency for both a simple triggering system and the convolution system can be measured using a variation of the protocol described by Wright [98]. The acoustic attack

and the resulting sound output from the audio interface will be recorded simultaneously on two channels, and the timing between their onsets will be measured by comparing the recorded audio data. Since there is no variation in the latency in this system, system performance can be represented by a single latency value.

## 7.1 *Timing controllability*

Latency and jitter was measured for the convolution algorithm, the computer input and output, and the combined system.

### 7.1.1 Convolver latency

Latency of the convolution algorithm was measured by generating clicks and convolving them with an impulse function, allowing the clicks to pass through the system unmodified, but delayed. Figure 7-1 shows the system schematic, the dotted line demarks the region responsible for the difference between the left and right channels. The clicks are recorded directly to the left channel of an audio file, and the right channel records the clicks after they emerge from the convolver. By looking at the offset between channels, one can measure the number of samples of latency introduced by the convolver (figure 7-2).

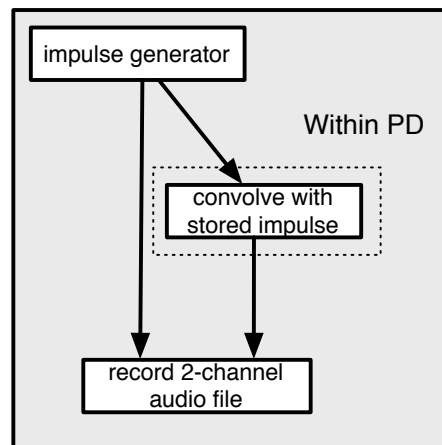


Figure 7-1: Schematic for measuring latency within the convolver

**Results** 20 trials were performed, and for each, the latency was exactly 64 samples (1.45ms). This is consistent with expectations since the smallest buffer size is 64 samples long. Figure 7-2 shows the resulting audio file. Measurement between peaks was performed manually in Audacity [4].

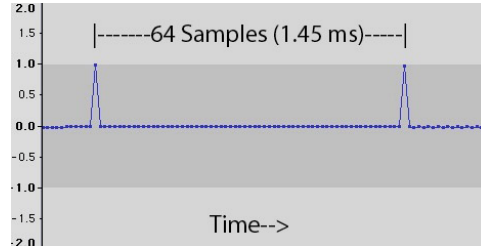


Figure 7-2: 64 samples (1.45 ms) of latency between the left and right channel

### 7.1.2 Audio subsystem latency

For these tests, I was interested in seeing what the round-trip input/output latency was for the audio interface and Pd without performing the convolution. Figure 7-3 shows the system schematic. Again, the dotted line demarks the region of interest. Clicks were generated in Pd and sent out DAC 1. DAC 1 was connected by cable to ADC 1. The input from ADC 1 was sent to DAC 2 and also recorded on the left channel of an audio file. DAC 2 was wired to ADC 2, and the input to ADC 2 was recorded to the right channel of the same file. The difference between the left and right channels is then the time it takes for audio to be sent out a DAC and received by the ADC. The order is the opposite of what we would usually think of, since it is measuring round trip from the computer and back, rather than from the outside and back, but the duration of the round trip is the same regardless of where it starts.

**Results** Delay in Pd was set to 25, 30, and 50 ms, and for each setting, 20 trials were performed. The results were identical for all trials at a particular setting, but they were not proportional to the set latency within Pd. Since Pd works with a block size of 64 samples, if we round up the Pd delay to the next multiple of 64 samples, we get a round-trip latency equal to the rounded Pd delay plus a constant, 496 samples. This extra delay is attributable to a combination of the operating system, drivers, and audio interface, and was constant for all settings of Pd delay:



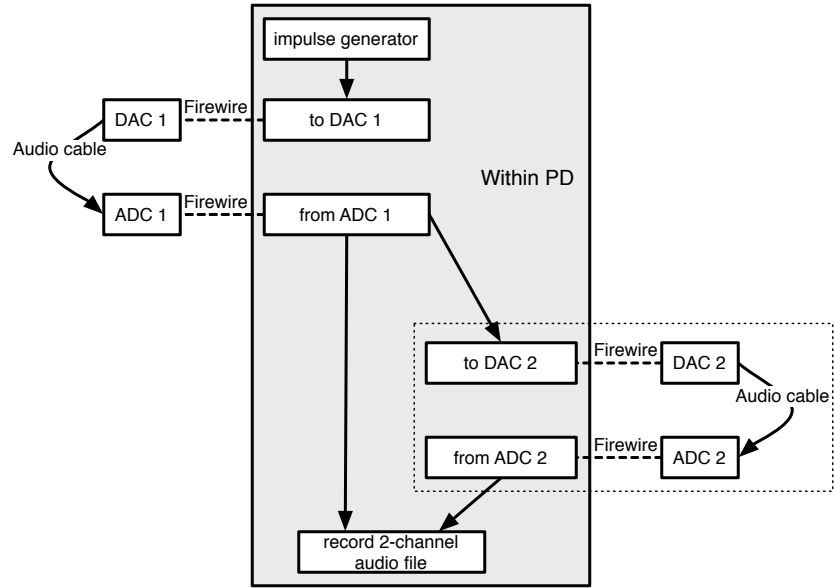


Figure 7-3: Schematic for measuring latency of the audio system, minus the convolver

Pd delay setting in ms	in samples	rounded up to a multiple of 64	measured latency (samples)	difference (samples)
25	1102.5	1152	1648	496
30	1323	1344	1840	496
50	2205	2240	2736	496

### 7.1.3 Total system latency

To measure the latency of the whole system, the experiment of section 7.1.2 was repeated but with the convolver inserted back into the system, as shown in figure 7-4. The region of interest is marked with a dotted line.

**Results** For Pd delays of 25, 30, and 50 ms, 20 trials were performed. Again, the results were identical within each trial group, and matched the results of 7.1.2 plus 64 samples, which was the measured latency of the convolver.

Pd delay setting in ms	in samples	rounded up to a multiple of 64	measured latency	difference
25	1102.5	1152	1712	560
30	1323	1344	1904	560
50	2205	2240	2800	560

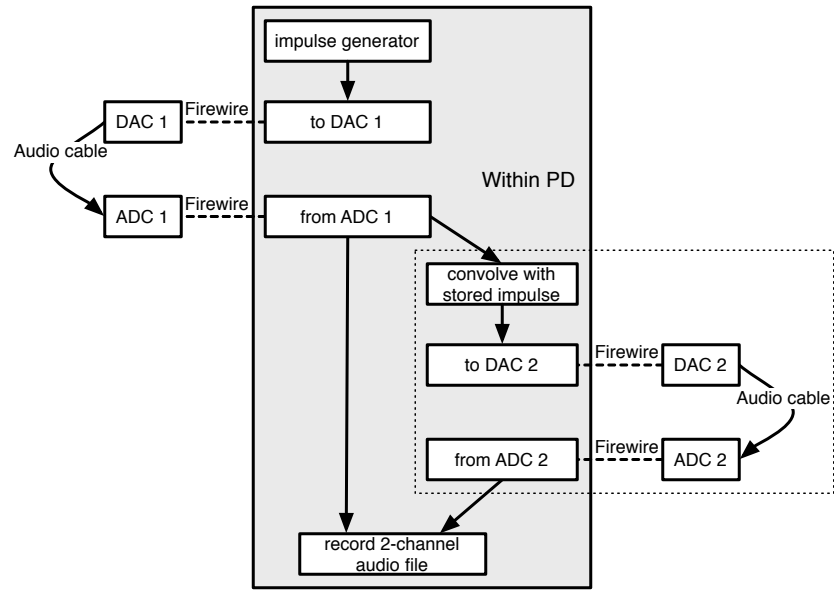


Figure 7-4: Schematic for measuring latency of the entire system

Using audio as a control system has the advantage that as long as there is no jitter in the audio system, then there is no jitter in the processing. Latency was limited by the audio subsystem, primarily by Pd, with an additional 496 samples due to the audio interface. Conceivably on dedicated hardware, the latency could be reduced to a single sample, but using general-purpose computers limits how low the latency can go before overloading the system.

There is no hard number for what is acceptable latency. In conversations with engineers from Yamaha, they have said that they want their percussion systems to have latencies of less than 4 ms. There is some dependence on the sound being used. Crisp attacks make latencies more apparent than smooth ones (a triangle vs. a gong, for example). Also players can learn to accept some latency. For example, the delay due to the speed of sound in air at a distance of 30 feet is 27 ms, not an unusual distance to be between a conductor and percussionists in an orchestra.

**PD control latency** Control data in PD is updated every 64 samples, giving up to 1.45 ms latency and jitter internal to the application. This could be improved by performing the calculations at audio rate, but at the expense of efficiency.

## 7.2 Observations and input from percussionists

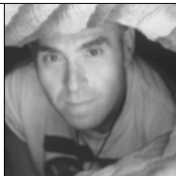



Percussionists overview		
Rakalam Bob Moses  Jazz Studies and Improvisation at New England Conservatory	Looking for sounds that were “more freaky-deaky” Used stick-slip rubber mallets, extended techniques. Transcends the physical world: objects are only a means for expressing the music.	
Jamey Haddad  Associate Professor of Percussion, Berklee College of Music	Created brush controller-as-snare / pickup for an acoustic frame drum. Added guiro effects, hand drumming technique. Favored shorter impulses, wanted more control over damping.	
Curt Newton  Jazz Percussionist	Highlighted the importance of damping, clear articulation. Used bass drum mesh artifact as interesting timbral element with metallic sounds.	
Dave Flaherty  Masters student, Jazz percussion	Brush-in-air + FM sounds, snare sound on the bass drum. Discovered how to tune head tension and damping for frame controller to best match different impulses. Jackhammer+ Bodhran, realistic Djembe mappings, observations about percussionists and technology.	

Figure 7-5:

### Rakalam Bob Moses

Rakalam Bob Moses, an extraordinary percussionist and teacher of Jazz Studies and Improvisation at the New England Conservatory, was kind enough to invite me (and an early version of the system) to his home. His biography is attached in appendix A.

His drum set is quite unusual. He uses two bass drums; one is a double-headed hand drum, the other is a djembe. They are tuned a fifth apart which enables him to set up drones almost like a bass is also playing. He also uses several metal Rototom frames stacked on top of his hi-hats, and he has a metal lamp cord that can be placed over a drum or cymbal to add a sizzle or snare-like effect. His snare and toms are adapted to enable hand drumming as well as playing with sticks by using hand-drum type rims that are below the playing surface. In the room surrounding his drum set was a huge assortment of instruments including a wind gong, a bent metal triangle that could bounce around on a snare while it is played, a

log drum, and a assortment of bells and other percussion elements.

He improvised using these objects for over an hour, stopping to explain a particular item or element periodically. The striking thing about how he worked was the complete seamlessness between all of the objects. A particular rhythm could move from the entire drumset to a single drum without losing its form, existing on only a snare, cymbal, or even the bass drum played by hand. It was as if he had transcended the divisions between the instruments as unique objects, and was instead using them collectively to channel his musical visions. This approach carried over to his interaction with my system, he began playing with mallets, eventually switching to a pair of 3-headed "crazy" mallets (figure 7-6) and quickly discovered that he could obtain extended "groaning" tones by drawing them across the drumhead. This was exactly the same technique that he had used on his acoustic drumset, but depending on the processed sound, it could have very different results since the stick-slip behavior of the mallets on the head was a function of the acoustics of the physical pad-and-mallet system only, while the resulting output was a combination of acoustic and processed sound.

He continued to play for at least two more hours, while I switched the system through several different modes. He wasn't confined to just using the system, but he played all surfaces within reach. When asked for his comments, his only request was that the sounds be more extreme, or in his words, "more freaky-deaky" As an example of what he was looking for, he played a cd track of a drummer playing through phaser and flanger effects.



Figure 7-6: "Crazy" mallets [59]

It took me several days to process and understand the value of this meeting. My expectation was that I would get lots of specific feedback about the instruments, so when it went differently, it took me a while to understand why. In the music technology context, things like flanger effects are trivial and well established, while making nuanced and controllable sounds is unusual and novel. In his context though, such nuance already exists in the acoustic instruments he uses all the time, and what is radical is something that sounds really different from the acoustic instruments. A second misconception I had going in to my visit was about the importance of specific instruments. I wanted for my system to be the center

of attention. Instead, as with all of his other instruments, it faded into the background – it was the music that took the foreground.

While I was there, Rakalam read several passages written by his spiritual teacher, the guitarist Tisziji Munoz on music, creativity, and spirituality. Later I found some excerpts of his writings, and one passage stood out as representative of my experience.

What is relatively creative  
binds one to consciousness  
or identification with things and beings.  
What is absolutely creative  
liberates one from self, thought and mind  
of everything and everyone. [48]

## **Jamey Haddad**

Jamey Haddad is a “genre-bending hand percussion specialist” [32] and associate professor at Berklee college of music, faculty at the Cleveland institute of music, and visiting professor at Oberlin conservatory (his bio is attached in appendix A). I had the opportunity to bring a version of my system (one practice pad controller and a wired brush) to one of his advanced hand drumming classes.

I was early, so I got to see the end of an introductory class where the students were using brushes on a frame drum. One brush acted as a snare and damper, and could also provide off-beats, while the other brush provided the dominant rhythm.

After the class dispersed, Mr. Haddad asked my what he was supposed to do with only one brush. We soon found out what that was. One of the first things he tried was to use the wired brush on an acoustic frame drum in the role of snare and damper, while playing the drum with a conventional brush. This had some surprising effects. the brush acted as a pickup for the drum itself, effectively coupling it to the convolver, but lifting the wired brush de-coupled the drum, allowing it to act purely acoustically. The ability to engage and disengage the processing using such a simple mechanism enabled a surprising level of depth.

He and his students took turns improvising with the elements of the system and other

instruments. and jammed for a couple of hours. When comparing the convolution technique to simple triggering, Haddad described the system as “way better than a [digital drum]pad” but not yet as expressive as a real drum. In particular, he also wanted more control over damping and the ability to fade between sounds. He also suggested adding textured elements to the pad for Guiro effects.

## **Curt Newton**

I met with drummer and percussionist Curt Newton (bio attached in appendix A) at our lab. He was the first percussionist I had shown the system to, so I was very interested in his feedback. I was able to show him early brush and pad controllers as well as the bass drum controller. After trying various settings, he found that the combination of the bass drum and a recording of a piano being struck while the damper pedal was held down was interesting, not for the way it responded to hitting, but for how it responded to scratching the mesh head of the bass drum with fingernails or brushes. The result was a distant and eerily resonant scratching sound.

Regarding the pad, he was concerned with damping and its effect on articulation, showing a strong preference for shorter stored impulses (at that time there was no control over damping).

## **Dave Flaherty**

Dave Flaherty is a Masters student studying percussion at the New England Conservatory, and he has been incredibly helpful in providing extensive feedback on the system over multiple extended sessions. Some specific design input included greatly decreasing the damping of the frame controller, and extending the nonlinear crash cymbal modeling to include more unusual sounds at the top end.

In general, he favored synthesizer samples I had rejected as being too weird or not controllable enough, and managed to extract a much greater range of sounds than I had. Like Rakalam, he was in favor of extreme sounds, but he was more interested in mappings that covered a continuum from normal to more unusual. For example, he found that shaking

the brush in the air could provide a rhythmic pulse that could be accented with hits and swipes. This worked best for bright, noisy, exponentially decaying FM sounds. He also favored more unusual pairings of timbres; one frame drum controller example convolved hits to the center of the head with a concert bass drum sample, while hits to the edge were processed with a jackhammer sample.

As the youngest percussionist in the group (in his mid twenties), his perspective on the relationship with technology that he and his contemporaries have differed from that of the older generation.

A lot of drummers use laptops right next to them when they're playing. They will play drums, do some electronic manipulation, sample themselves, mess with that. Hooking up one of these controllers would give another option for all of those things we do.

Other of Dave's contributions included increasing the tension on the bass drum controller head to allow for realistic snare drum sounds and technique, placing objects on the bass drum head to modify the drum sound (similar to the effects Rakalam had achieved with the lamp cord and triangle on his acoustic snare), and more generally recognizing head tension and physical damping as variables that can be optimized for different stored impulses and playing styles. For example, the frame controller benefitted from a tighter head when playing djembe sounds, and a looser one for driving bodhran timbres.

### 7.3 My own observations

Some features of the instruments are apparent. The instruments are highly controllable and entertaining to play, and with the right impulse responses, players have found that they begin to believe that the physical object is responsible for the response – they (and myself as well) are even surprised when the system is turned off or when the impulse response is substituted for another. These instruments integrate well with their acoustic counterparts and are uniformly favored over simple triggering for the same sounds and physical interfaces.

Having some secondary gestural control over the convolution algorithm allowed more realistic control for parameters such as damping, cross fades, and pitch shifts. For damping

using the cymbal controller, it was much more important for the entire surface of the cymbal to respond to a choke on or off than it was to have a range of pressures in one location.

Musically, I found it to more than hold my interest to hear the various controllers played by these players. I had good luck that all of them had strong timbral sense, and worked to extract a range of tones from the instruments. There were some surprises: I had not expected that realistic pitch control could come from the drum head of the frame drum, and the range of rim shot timbres possible when playing the bass drum with sticks and a snare sound was beyond what I would have thought would be possible. Tuning the heads also had a greater effect on the sound than I had anticipated. In each case, it was easy to forget that there was only one (or in some cases two) stored samples, around three seconds of audio responsible for such a broad range of timbres.

One of the strengths of the system is how it handles more unusual timbres. For example, the purely synthetic FM sounds, when connected to the brush controller, took on a much more acoustic quality, and responded as if some impossible structure was being played with brushes. All of the articulation and the variation one would hear in a drum played with a brush was mapped instead onto this totally different and acoustically unlikely timbral space. Because of this, the stored impulses can depart quite radically from real sounds, but still maintain a level of understandability due to their imposed acoustic grounding.

Some other observations:

- All of the percussionists noticed and remarked on latencies above about 15 ms, though all were able to adapt and play with latencies as high as 40 ms without difficulty.
- The system can only put energy into the mode structure represented in the impulse response sample – it is not possible to excite modes that were not present in the impulse response. For some sounds, this necessitates running multiple convolutions using samples taken by hitting different parts of the target object to excite different modes, or dynamically switching between impulse responses.
- The technique of sending audio through FSRs and bend sensors to measure their resistance is susceptible to capacitive coupling. This means that sometimes I am not actually measuring the bend or pressure, but am instead picking up the presence of a



hand or conductive object near the sensors.

Further discussion and future direction continues in the next chapter.



## Chapter 8

# Discussion / future directions

### 8.1 Applications

Applications of this work range from the immediate and obvious (digital drumsets) to the more far-reaching such as product design and toys.

**Digital drumsets** Incorporating the techniques presented in this thesis would add a level of realism to digital drum sets, particularly for cymbals. This points to a new genre of instrument, the acoustic / electronic drumset. This could be combined with the existing highly-refined triggering systems and direct audio processing in digital drum sets to strike a middle ground between the controllability and realism of the techniques presented in this thesis and the interchangeability, resistance to feedback, and heightened articulation that are present in the best digital drum systems.

**Desktop percussion** Another application would be for desktop percussion systems, literally turning the top of a desk, or the surface of a laptop into a percussion instrument. This could be simply for recreation, or as a way to program drum tracks with a more human groove. will.i.am of the Black Eyed Peas described his technique of laying down beats as literally drumming on the mixing desk, with a microphone pointed at it. He would then go back and insert samples at each hit [97]. A desk-top percussion system could provide

processed feedback through various resonances, while still recording the raw signals to be manipulated later as necessary, either by his current technique of substitution, or through convolution with different impulses.

**Laptop percussion** A contact microphone could be attached to the lid or surface of a laptop, making a capable portable percussion composing system. Even just using the internal microphone is sufficient to get some interesting output. Brushes and sticks with audio sensors could also be integrated, possibly as a performance system to provide more expression and visual interest to the typical laptop performance. One could "brush out" pre-synchronized loops and phrases as part of a live set, or overlay additional rhythm and texture layers in real time.

**Toys** Having seen some of the extreme cost and market constraints of the toy industry in a previous attempt to commercialize a percussion instrument [2], I am very aware of the barriers to using a system like this in a toy context. While the processing requirements are way beyond what is currently possible to integrate into a toy, any current computer would be sufficient. This points toward peripherals that connect to a computer. One idea would be a flash audio recorder "sound sucker" that kids could gather and play back sounds with, and then connect to their computer to construct drumsets out of sounds from their environment. As mobile phones become more capable, they could also provide an excellent platform for mobile sound-transforming and gathering applications.

**Foley** Foley artists, already adept at creating their own rigs to produce sounds, could add these techniques to their arsenal; rather than manipulating the Foley sounds after recording, they could be processed with resonances of preexisting sound effects or custom-gathered sounds for a particular application. The acoustics of objects that are too ungainly to be brought into the studio could be sample and manipulated via contact microphones and other physical objects.

**Beyond percussion, modifying apparent acoustics for design purposes** Real objects have their own unique acoustic response: using the techniques in this thesis, the

apparent auditory and vibratory responses of those objects can be modified. This could be used to express the state of the system, convey ambient information, or to emulate more expensive materials; plastic that sounds like stone when you touch it, or a membrane switch that has a satisfying “thunk”. Automotive applications could include car doors, trunks, or the exhaust note. Anything that has an apparent idiophonic acoustic quality is a candidate for augmentation and transformation. This could be a subtle enhancement, or a radical departure from the expected acoustics.

**Installations** These techniques could be used in interactive installations to let people strike, scrape, and handle objects with unlikely or impossible acoustics; for example, they could freely reassign the acoustic properties of objects from a set of choices. The sounds could range from the conventionally musical to nonmusical sounds like gravel and dirt. The same techniques could apply to site-specific installations that turn otherwise dead surfaces into responsive sonic surfaces. Most of the work in this thesis has been at the smaller-than-human scale. Moving up to a more architectural scale suggests different acoustic possibilities and the response of objects starts to overlap the responses of rooms and spaces.

## 8.2 Future work

On the technical side, there is still a lot more that should be done.

**Efficient implementation** The current implementations in Max/MSP and Pd exchanged efficiency for easy experimentation and reconfiguration. Getting the system to run using fewer resources or to process more input simultaneously would make it useful to more people.

**Spectral morphing** The full range of available spectral transformations is possible, and should be implemented. For example, to morph between two samples at different pitches, it would be necessary to identify which spectral peaks of one sound correspond to the peaks in the second.

**Embedded development** For this work to be easily integrated into commercial products, it would need to be embedded, possibly in lower-cost DSPs. As general purpose computers become more capable and less expensive, they may also represent an alternative to DSPs for embedded applications. In either case, minimizing system latency will be important for any percussion application.

**Physical controller design** The controllers presented in this thesis are just a glimpse of what is possible, but it will take the same dedicated development and continued optimization that has been necessary for all existing instruments.

**Better sensing** Similarly, the sensing systems described in this thesis are not optimal, and there are likely many sensing modalities that could enable other physical interactions and extend the ergonomic or esthetic range of the instruments.

**Broader study of the role of acoustics in the identity of objects** There is more to understand about how we perceive objects, and what it means to be able to alter their apparent acoustics is an open question for scientific and artistic investigation.

## 8.3 Conclusion

Specific contributions of this thesis are:

- A novel system architecture that allows players to apply their intuitions and expectations about real acoustic objects to new percussion instruments that are grounded in real acoustics, but can extend beyond what is possible in the purely physical domain.
- Extensions to the functionality of convolution algorithms to accommodate muting, pitch shifts, approximation of nonlinear effects, and inverse filtering.
- A range of semi-acoustic physical controllers designed to integrate with the system architecture and that illustrate design principles for future instruments.

- An implementation of these algorithms that can serve as a platform for future development and allow customization to meet future creative goals.
- Applications to the areas of human-computer interface and product design, exploring apparent acoustic properties as a design parameter and for information display.

**First there was audio...** One of the threads of this thesis is really about using realtime audio to control synthesis. In the context of 23 years of MIDI as a dominant paradigm in digital music, this might seem new, but it actually has more in common with the early modular synthesizers. Audio input from percussion pads, microphones, or anything else could be plugged into envelope followers, banks of filters, or used to modulate other aspects of synthesis. Since everything was represented as an analog voltage, there was no technical distinction between audio signal and control signal.

Over time most synthesizer companies moved away from the modular design to a more keyboard-centric paradigm, finally resulting in the MIDI specification. Even the idea of what a “controller” or synthesizer can be is strongly shaped by the framework of MIDI. Digital percussion has largely worked within this framework not because it is the ideal way to represent percussion, but because it represents the dominant paradigm.

Applications such as sequencers were only possible using the early personal computers because of the simplicity of the MIDI representation. It perfectly matched the computational capabilities of the time. Even the concept of digital processing had to do with a human-scale number of bits, a logical representation. As computational power has advanced though, personal computers are now capable of sophisticated audio processing and synthesis on their own. We have pushed so far past the threshold where a person could still keep track of all the bits that the fact that something is digital means that it can actually behave a lot like something analog. At the same time, ever-more inexpensive multichannel audio interfaces can allow a much more sophisticated connection between the real sonic world and the computer. Without these two elements, fast computers and cheap i/o, this thesis work would not have been possible, and the continued developments along these two axes suggests that there are more opportunities for music and instruments that further blur the distinctions between what is digital and what is everything else.

**Surprise the designer** Another of the key contributions of this thesis is to provide a method for negotiating the gap between the digital and physical world through the shared ground of acoustics and physical laws. While typically a musical interface designer has to anticipate every possible gesture that that a player may want to perform, there is some continuity that is achieved for free. Accidents, or at least events unexpected by the designer, have been commonplace in my interactions with percussionists, and the system architecture supports a broad range of uses and abuses.

For example, Rakalam Bob Moses used the rubber multi-superball mallets to rub the head of the pad, getting stick-slip response that created sustained tones. Jamey Haddad used the wired brush controller as a snare element by holding it on the head of an acoustic frame drum, letting it also act as a pickup for the drum. This also allowed him to revert to purely acoustic playing just by lifting the brush from the drum head. Curt Newton found that scraping the mesh head of the bass drum controller with his fingernails was an interesting interaction, creating a distant resonant scratching sound when convolved with a bass drum or an undamped piano resonance, and Dave Flaherty turned a bass drum into a snare, and used brushed in the air to control FM sounds.

Each of these examples would have been totally impossible in a conventional triggering system, and the instruments encourage that kind of exploration. If someone tried to scrape a pad and heard nothing, they'd just stop.

One of the most humbling aspects of these interactions with percussionists is that they are so good at incorporating any object and getting a range of sounds out of it that it is important to remember that the systems in this thesis are only some tools out of many that are available. The point isn't the tool, but the use of it. I remember when Rakalam Bob Moses was playing with an early pad controller, he began playing off beats on the chair it was resting on, and he even transitioned to just playing the chair for a while. My initial reaction was one of dismay – my instrument was being ignored, and for a chair no less! But I later realized that I was viewing the event as a technologist, and that, in fact, creating an instrument that can seamlessly meld with other objects in the percussionist's environment is actually the best one can hope for. A chair may be good for sitting, but in the hands of a great percussionist it becomes something else entirely.



# Appendix A

## Percussionist biographies

### Rakalam Bob Moses

From his New England Conservatory Bio:

Drummer, composer, artist, poet, dancer, visionary, nature mystic: Bob Moses's life has been a continuous quest for vision, spirit, compassion, growth, and mastery in a multiplicity of art forms. A partial list of stellar musicians Bob Moses has worked and/or recorded with both as a leader and a sideman includes:

Charles Mingus, Rahsaan Roland Kirk, Jim Pepper, Larry Coryell, The Free Spirits, Dave Liebman, Gary Burton, Keith Jarrett, Steve Swallow, Michael Gibbs, Pat Metheny, Lyle Mays, Paul Bley, Herbie Hancock, Jack DeJohnette, Jaco Pastorius, Hal Galper, Michael Brecker, Randy Brecker, Steve Kuhn, Sheila Jordan, Bobby McFerrin, Dave Sanborn, Bill Frisell, Eddie Gomez, Don Alias, John Scofield, Terumasa Hino, Dave Holland, Charlie Haden, Hermeto Pascoal, Jovinos Santos Neto, Danilo Perez, David Sanchez, Chucho Valdes, Jimmy Slyde, Savion Glover, Gregory Hines, Stan Strickland, Tiger Okoshi, Nana Vasconcelos, Obo Andy, John Medeski, Vernon Reid, DJ Logic, Badal Roy, Ravikiran, Master Anand, Raqib Hassan, Sam Rivers, Pharoah Sanders, Tisziji Munoz, and many others.

Moses's first recording as a leader was the 1968 album *Love Animal*, with Jim Pepper, Larry Coryell, Steve Swallow, and Keith Jarrett. The albums *Bittersuite in the Ozone* and *Love Everlasting*, the later recorded with spiritual master and guitar genius Tiszij, both received five stars in *Downbeat*. The album *Munoz*, released on Amulet Records in 1987, features Jerry Bergonzi, George Garzone, John Medeski, and John Lockwood, while the album *Drumming Birds*, also 1987, features percussion duets with Billy Martin. [51]

## Jamey Haddad

Official biography:

Born in Cleveland Ohio, Percussionist/Drummer Jamey Haddad holds a unique position in the world of Jazz and Contemporary Music. Since 1991 Jamey has performed in the working bands of Dave Liebman, Joe Lovano, Alan Farnham, The Paul Winter Consort, Carly Simon and Betty Buckley. Performing with the great oud players/composers, Rabih Abou Khalil and Simon Shaheen in the Mid East.

Haddad's musical voice transcends styles and trends, and the universal quality of his playing has attracted many international collaborations.

Most recently Haddad performed with long time musical associate saxophonist Joe Lovano for a duet concert in New York city, followed by a trip to the Mideast with oud and violinist Simon Shaheen. For over ten years fellow percussionist / composer Steve Shehan has invited Haddad to collaborate on numerous projects most recently to Paris and Caracas with the great Touareg musician Baly Othmani.

In 1992 Haddad was invited by composer Richard Horowitz and the Moroccan Government to help develop and perform compositions with 10 different Berber and Ganawan groups for a Crown Performance at the 1992 Worlds Fair in Seville, Spain. Haddad was also one of two Americans to perform in "World Drums" at the 1988 Olympics in Calgary, Canada and again at the 1988 Worlds Fair in Brisbane Australia, over 250 Percussionist from 25 Countries were invited. Other recent performances include touring Austria with bassist Peter Herbert.

Haddad is the Recipient of a Fulbright Fellowship to South India, four National Endowment for the Arts Fellowships, two in Jazz Performance and two in International Music Studies/Collaborations. The Ohio and Pennsylvania Council on the Arts also awarded Haddad jazz performance grants.

Internationally in demand for his seminars, master classes, and written essays on music, Haddad chooses to talk about the more eternal and universal qualities of a musical life. He has developed two extremely popular courses in "World Music" and teaches at Berklee College of Music in Boston and the New School in New York City. Recent seminar / performances include the 1996 - 97 and 98 Percussive Arts Society Conventions, performing with hand drumming virtuoso, Glen Velez; The Berklee Percussion Seminar, performing with South Indian Master Drummer Trichy Sankaran and most recently returned from a performance and seminar at a World Percussion Conference in Mexico City.

Haddad has been the co-creator of many musical instruments and playing techniques that are finding their way into the hands of percussionist worldwide. The are the Hadgini, the Hadjira, the Hadjenga, and the Kohabata Drums.

In the jazz and contemporary music scene Haddad has appearing on over 75 recordings in addition to hundreds of performance credits as leader and sideman world wide.

Haddad has recently completed a Book and supporting Video on the subject of internalizing your personal rhythm. The observations and lessons found in "Global Standard Time" are addressed to any musician looking to strengthen their perceptions of levels of time and rhythms, and the grooves they dance in. [31]

## Curt Newton

Biography from <http://www.curtnewton.com>:

Boston-based drummer and percussionist Curt Newton has been playing uncompromising original music since 1986. A highly versatile ensemble player and soloist, he integrates innovative sounds and extended techniques with the jazz and rock drumset traditions. He has worked with Ken Vandermark, Joe Morris, Nate McBride, Pandelis Karayorgis, Charlie Kohlhase, and Debris; performed across the United States, Canada, and Europe; and appears on CDs released by leading independent labels like Okkadisk, Boxholder, Soul Note, Music and Arts, Buzz, and Rastascan.

Curt currently performs with the following groups:

- Nate McBride Quartet (w/ Charlie Kohlhase, saxophones; Taylor Ho Bynum, brass; Nate McBride, bass)
- the mi3(w/ Pandelis Karayorgis, fender rhodes; Nate McBride, bass)
- The Chris Allen Estate
- Trio Ex Nihilo (w/ Jeff Song, cello; Taylor Ho Bynum, brass)
- on rare occasions, The Poppies

He has worked since 1986 with Chicago saxophonist and composer Ken Vandermark, releasing three acclaimed CDs; their most recent project is the group Tripleplay with bassist Nate McBride. He also plays with leading musicians including Hans Poppel, Joe Morris, James Rohr, Greg Kelley, and Jorrit Dijkstra.

**Percussion Performances** Curt has been developing a solo drumset repertoire for several years, with original compositions, improvisations, and occasional works for drumset by other composers. Select performances include:

- an all-drumset concert with master drummers Alan Dawson and Bob Gullotti (1992)
- world premiere of John Zorn's Hwang Chin Ee for two drumsets and narrator (New England Conservatory, 1996)
- original transcription of Lutoslawski String Quartet for solo drumset (1997)
- activating Chen Zhen's sculpture Jue Chang (50 Strokes to Each) at Boston's Institute of Contemporary Art (2002)

**Education and Influences** Curt studied drum set with Bob Gullotti and has a Master of Music degree in Jazz Performance from the New England Conservatory of Music. His musical influences and interests, in no particular order, have included Max Roach, Andrew Cyrille, Steve Shelly, Gerry Hemingway, Joey Baron, Paul Lovens, Bob Gullotti, Ornette Coleman, Charles Mingus, Eric Dolphy, Cecil Taylor, Son House, Captain Beefheart, Igor Stravinsky, Iannis Xenakis, Witold Lutoslawski, and the traditional music of Morocco, Ghana, Egypt, Turkey and Greece. [52]



## Appendix B

# Offset deconvolution in Octave

offsetdeconv.m

---

```
function [x]=offsetfdeconv(y, h)
%   OFFSETFDECONV Fast deconvolution with an offset
%   [x] = OFFSETFDECONV(y, h) deconvolves h out of y, and noramalizes the
%       output to +-1.
%
%       y = input vector
%       h = input vector
%
%       See also DECONV
%
% based on fdeconv 1.0 by: Stephen G. McGovern
%

Lx=length(y)+1;                                % assuming y is longer
Lx2=pow2(nextpow2(Lx));                         % Find smallest power of 2 that is > Lx
Y=fft(y, Lx2);                                 % Fast Fourier transform
H=fft(h, Lx2);                                 % Fast Fourier transform
magY=abs(Y);                                    % magnitude of Y
phaseY=angle(Y);                               % phase of Y
magH=abs(H);                                    % magnitude of H
phaseH=angle(H);                               % phase of H
magX=(magY)./(magH+.1*(max(magH)));             % add an offset
phaseX=(phaseY).-(phaseH);                     % subtract phases
X= magX.*exp(i*phaseX);                        % convert back to complex
x=real(ifft(X, Lx2));                          % IFFT
x=x(1:1:Lx);                                   % first N elements
x=x/max(abs(x));                              % Normalize the output
```



# Appendix C

## Pin outs

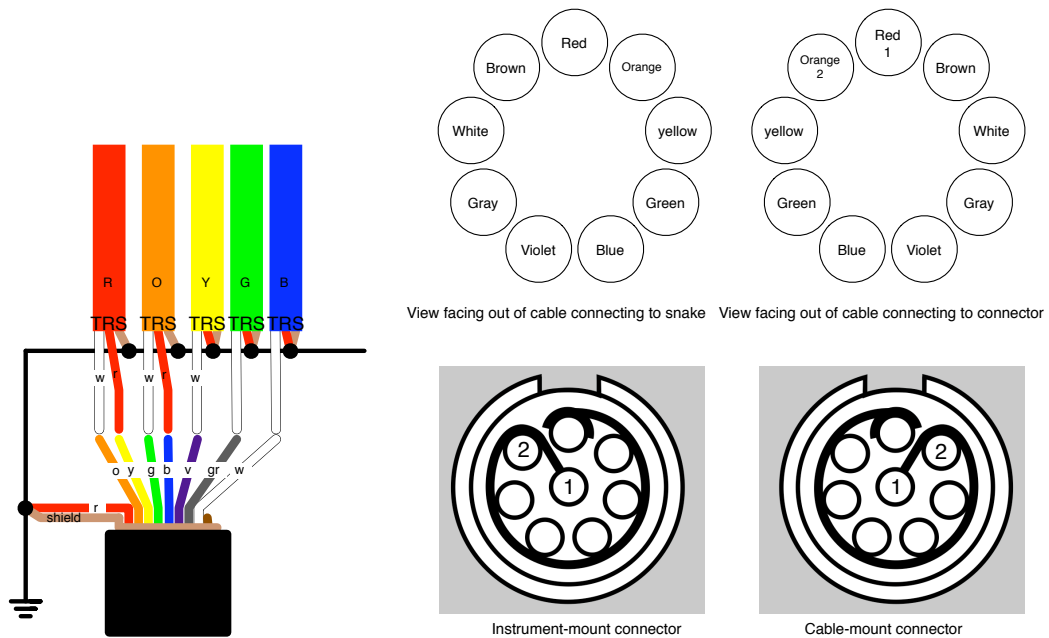


Figure C-1: Connector and cable pinouts





# Bibliography

- [1] Abrams/Gentile Entertainment, Inc. New York, NY. <http://www.ageinc.com>.
- [2] Roberto Aimi. New expressive percussion instruments. Master's thesis, Massachusetts Institute of Technology, 2002.
- [3] Craig Anderton. 20 great achievements in 20 years of musical electronics, 1968-1988. *Electronic Musician*, 4(7):28–97, July 1988.
- [4] Audacity. <http://audacity.sourceforge.net/>.
- [5] Tony Bacon and Paul Day. *The Rickenbacker Book: a Complete History of Rickenbacker Electric Guitars*. GPI-books, San Francisco, 1994.
- [6] G. D. Beauchamp. Electrical stringed musical instrument. US Patent # 2089171, 1937.
- [7] Thomas Bloch. The ondes martenot. *Experimental Musical Instruments*, 8(1):8–12, 1992.
- [8] Robert Boie. Radio drum. In S. Arnold and G. Hair, editors, *Proc. 1990 International Computer Music Conference*, pages 398–400, San Francisco, CA, 1990. International Computer Music Association.
- [9] Richard Boulanger. Two new digital techniques for the transfiguration of speech into musical sound. *Ex Tempore*, 4(1):21–55, Spring 1986.
- [10] Donald Brosnac. *The Electric Guitar: its History and Construction*. Panjandrum Press, San Francisco, CA, 1975.
- [11] Bob Brozman. *The History and Artistry of National Resophonic Instruments*. Centerstream Publishing, Fullerton, CA, 1993.
- [12] Joel Chadabe. *Electric Sound: The Past and Promise of Electronic Music*. Prentice-Hall, New Jersey, 1997.
- [13] Antoine Chaigne, Mathieu Fontaine, Olivier Thomas, Michel Ferre, and Cyril Touzé. Vibrations of shallow spherical shells and gongs: a comparative study. *Forum Acusticum*, 2002.
- [14] Cary Clements. Augustus Stroh and the famous Stroh violin. *Experimental Musical Instruments*, 10(4):8–15, 1995.

- [15] Julian Colbeck. The wave of the future? *Sound on Sound*, November 1994.
- [16] Cycling '74. *Max/MSP*. <http://www.cycling74.com/>.
- [17] A. Darvishi, E. Munteanu, V. Guggiana, M. Motavalli, M. Rauterberg, and H. Schauer. Automatic impact sound generation for using in non-visual interfaces. *Proceedings of 1st Annual International ACM/SIGCAPH Conference on Assistive Technologies*, pages 100–106, 1994.
- [18] Hugh Davies. Electronic percussion. In S. Sadie, editor, *The new grove dictionary of musical instruments*. Macmillan, London, second edition, 1984.
- [19] John Dopyera. Stringed musical instrument. US Patent # 1,741,453, 1929.
- [20] John Duesenberry. Convolution number nine. *Electronic Musician*, 15(6):118–124, June 1999.
- [21] E-MU systems, Inc. *Emax SE update manual*, 1988.
- [22] Edison national historic site archive. National Park Service, US Department of the Interior, West Orange, New Jersey <http://www.nps.gov/archive/edis>.
- [23] Tom Erbe. *SoundHack user's manual*. CalArts School of music, <http://www.soundhack.com/SndHckDoc/>.
- [24] Vic Firth. *Snare drum method : book 1 - elementary*. Carl Fischer, New York, N.Y., 1967.
- [25] Neville H. Fletcher. *Complex Systems: From Biology to Computation*, chapter Non-linear Dynamics and Chaos in Musical Instruments, pages 106–117. IOS Press, Amsterdam, 1993.
- [26] Gillian Gaar. Liner notes in *Talk normal: the Laurie Anderson anthology*.
- [27] Lynne Gagné. A short history of orchestral percussion. *La Scena Musicale*, 9(3), November 5 2003.
- [28] William Gardner. Efficient convolution without input-output delay. *Journal of the Audio Engineering Society*, 43(3):127, 1995.
- [29] Roland Gelatt. *The fabulous phonograph, 1877-1977*. Macmillan, New York, 1977.
- [30] Albert Glinsky. *Theremin: ether music and espionage*. Music in American life. University of Illinois Press, Urbana and Chicago, 2000.
- [31] Jamey Haddad. *About Jamey Haddad*. <http://www.jameyhaddadmusic.com>, 2000.
- [32] Jamey Haddad. Coda: The spirit of world music. *Berklee Today*, 16(2), Fall 2004.
- [33] A. Hunt, M. M. Wanderley, and R. Kirk. Towards a model for instrumental mapping in expert musical interaction. In *Proceedings of the 2000 International Computer Music Conference*, pages 209–211. International Computer Music Association, 2000.

- [34] A. Hunt, M. Wanderly, and M. Paradis. The importance of parameter mapping in electronic instrument design. In *New Interfaces for Musical Expression proceedings*, Dublin, 2002.
- [35] Interlink Electronics, Inc. Camarillo, CA, <http://www.interlinksensors.com>.
- [36] Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, 1997. SIGCHI: ACM Special Interest Group on Computer-Human Interaction, ACM Press.
- [37] J. M. A. Stroh. Violin or other stringed instrument. US Patent # 644695, 1900.
- [38] Ikutaro Kakehashi and Robert Olsen. *I Believe in Music: Life Experiences and Thoughts on the Future of Electronic Music by the Founder of the Roland Corporation*. Hal Leonard Corp, 2002.
- [39] Ajay Kapur, Philip Davidson, Perry R. Cook, Peter F. Driessen, and W. Andrew Schloss. Digitizing north indian performance. In *Proceedings of the International Computer Music Conference, 2004*, 2004.
- [40] Korg corporation. Tokyo, Japan <http://www.korg.com>.
- [41] Korg corporation. *Wavedrum user's guide*. Tokyo, Japan, 1994.
- [42] John Krogh. Keyboard reports - roland hpd-15 handsonic: Electronic hand percussion instrument. *Keyboard Magazine*, 27(2):108 (6 pages), 2001.
- [43] David Stanley McGrath. Efficient convolution method and apparatus. United States Patent 6,574,649, June 22 2001.
- [44] Measurement Specialties, Inc. Hampton, VA. <http://www.meas-spec.com>.
- [45] Measurement Specialties, Inc. *Piezo Film Sensors Technical Manual*, 1999.
- [46] Liz Mellish and Nick Green. *The Eliznik Romania pages*. <http://www.eliznik.org.uk>.
- [47] MIDI Manufacturers Association. *The Complete MIDI 1.0 Detailed Specification, Version 96.1*. Los Angeles, CA. <http://www.midi.org>, 2nd edition, November 2001.
- [48] Tisziji Munoz. *The Great Equation*. The Illumination Society Publications, 2001.
- [49] Nady systems, inc. Emeryville, CA, <http://www.nady.com>.
- [50] National Reso-Phonic Guitars, Inc. San Luis Obispo, CA <http://www.nationalguitars.com>.
- [51] New England Conservatory. *NEC faculty: Rakalam Bob Moses*. <http://www.newenglandconservatory.edu/faculty/mosesr.html>.
- [52] Curt Newton. *Curt Newton – Drumset/Percussion – Jazz and Improvised Music*. <http://www.curtnewton.com>.
- [53] PaiA, inc. Touch switches. <http://www.paia.com/touchsw.htm>.

- [54] Jacques Paillard. *The Use of Tools by Human and Non-human Primates*, chapter The hand and the tool: the functional architecture of human technical skills, pages 36–50. Clarendon Press, Oxford, 1993.
- [55] Joseph Paradiso. American innovations in electronic musical instruments. *New Music Box*, 6, October 1999. Available at [http://www.newmusicbox.org/third-person/index\\_oct99.html](http://www.newmusicbox.org/third-person/index_oct99.html).
- [56] James Patten. *Mechanical Constraints as Common Ground between People and Computers*. PhD thesis, Massachusetts Institute of Technology, 2006.
- [57] Photograph of Charles D’almaine. Collection of Glenn Sage, Portland, OR <http://www.tinfoil.com>.
- [58] Pintech USA, inc. Greenville, SC <http://www.pintechworld.com>.
- [59] Plugs Percussion. [plugspercsm@aol.com](mailto:plugspercsm@aol.com).
- [60] Poșta Română. <http://www.posta-romana.ro>.
- [61] William Preece. The phonograph. *Journal of the Society of Telegraph Engineers*, 7, 1878.
- [62] Miller Puckette. "pure data.". In *Proceedings, International Computer Music Conference*, pages 269–272, San Francisco, 1996. International Computer Music Association.
- [63] Speranta Radulescu. Item descriptions, Museum of the Romanian Peasant. [http://www.itcnet.ro/mtr/index\\_en.html](http://www.itcnet.ro/mtr/index_en.html), 2003.
- [64] Gordon Reid. 40 years of gear: The history of korg: Part 2. *Sound on Sound*, November 2002.
- [65] Gordon Reid. French connection: Analogue systems french connection analogue keyboard controller. *Sound on Sound*, February 2002.
- [66] Robert Rich. Buchla thunder. *Electronic Musician*, page 84, August 1990.
- [67] Marc Richardson and Ron Potter. Viscous vs. structural damping in modal analysis. In *46th. Shock and Vibration Symposium*, 1975.
- [68] Eugene I. Rivin. *Stiffness and Damping in Mechanical Design*. Marcel Dekker Inc, 1999.
- [69] Curtis Roads. Musical sound transformation by convolution. In *proceedings of the 1993 international computer music conference*, pages 102–109, Tokyo, 1993. international computer music association.
- [70] Curtis Roads. *The Computer Music Tutorial*. MIT Press, Cambridge, MA, USA, 1996.
- [71] Gino Robair. Roland v-pro td-20s. *Electronic Musician*, 21(4):92, 2005.
- [72] Roland Corporation. *Handsonic HPD-15 Owner’s Manual*, 2000.
- [73] Roland corporation. *percussion sound module TD-20 owner’s manual*, 2004.

- [74] Thomas D. Rossing. *Science of Percussion Instruments*, volume 3 of *Series in Popular Science*. World Scientific, Singapore, 2000.
- [75] Thomas D. Rossing. Acoustics of percussion instruments: recent progress. *Acoustical Science and Technology*, 22(3):177–188, 2001.
- [76] J. B. Rován, M. M. Wanderley, S. Dubnov, and P. Depalle. Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In A. Camurri, editor, *KANSEI - The Technology of Emotion, AIMI International Workshop*, pages 68–73, Genova, October 3-4 1997. Associazione di Informatica Musicale Italiana.
- [77] Greg Rule. Keyboard reports: Korg wavedrum. *Keyboard*, 21(3):72–78, March 1995.
- [78] Greg Rule. Roland v-drums - a new standard in positional head sensing, plus a megaton of punchy percussion samples. *Keyboard*, 24(4):77–84, April 1998.
- [79] Greg Rule and Steve Fisher. V-drums history: How Roland became the biggest drum company in the world. *Roland users group insider* [http://www.rolandus.com/community/insider/artists\\\_articles.aspx?ArticleId=20](http://www.rolandus.com/community/insider/artists\_articles.aspx?ArticleId=20), 2005.
- [80] Ben Saylor. `partconv~0.2`. <https://puredata.org/Members/bensaylor/partconv~-0.2.tar.gz/view>, 2005.
- [81] Kenji Senda and Yasunobu Miyamoto. Apparatus and method for detecting and processing impacts to an electronic percussion instrument. US Patent # 6601436, Filed January 2001. Assignee: Roland Corporation, Osaka, Japan.
- [82] Zack Settel and Cort Lippe. Real-time timbral transformation: FFT-based resynthesis. *Contemporary Music Review*, 10:171–179, 1994.
- [83] James R. Smart and Jon W. Newsom. *A wonderful invention : a brief history of the phonograph from tinfoil to the LP*. Library of Congress, 1977.
- [84] Julius O. Smith. *Mathematics of the Discrete Fourier Transform (DFT)*. W3K Publishing, <http://www.w3k.org/books/>, 2003.
- [85] Julius O. Smith III. Physical modeling using digital waveguides. *Computer Music Journal*, Volume 16(4):74–91, Winter 1992. Special issue: Physical Modeling of Musical Instruments, Part I.
- [86] Julius O. Smith III. *Physical Audio Signal Processing: Digital Waveguide Modeling of Musical Instruments and Audio Effects, August 2004 Draft*. <http://ccrma.stanford.edu/~jos/pasp/pasp-citation.html>, August 2004.
- [87] Sound Ideas. Sound ideas sound effects library: series 1000, 1983.
- [88] Spirit Canyon Audio. *Spectral Relativity*. <http://www.spiritcanyonaudio.com/>, 2004.
- [89] Thomas G. Stockham, Jr. High-speed convolution and correlation. In *AFIPS Proceedings, 1966 Spring Joint Computer Conference*, volume 28, pages 229–233, 1966.

- [90] Larry the O. Nearfield multimedia marimba lumina: This is not your mother's midi controller. *Electronic Musician*, 16(6):138+, June 2000.
- [91] Cyril Touzé and Antoine Chaigne. Lyapunov exponents from experimental time series: application to cymbal vibrations. *Acta Acustica*, 86:557–567, 2000.
- [92] Mark Vail. Physical modeling synthesizer Yamaha VL1. *Keyboard*, 32(8):84, August 2006.
- [93] K. Van Den Doel and D. Pai. *Audio anecdotes*, chapter Modal synthesis for vibrating objects. A.K. Peter, Natick, MA, 2003.
- [94] Marcelo Mortensen Wanderley and Nicola Orio. Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*, 26(3):62–76, Fall 2002.
- [95] Norman Weinberg. Electronic percussion. In John S. Beck, editor, *Encyclopedia of percussion*, number 947 in Garland reference library of the humanities, pages 191–193. Garland publishing, New York and London, first edition, 1995.
- [96] Walter L. Welch, Leah Brodbeck, and Stenzel Burt. *From tinfoil to stereo : the acoustic years of the recording industry, 1877-1929*. University Press of Florida, Gainesville, FL, 1994.
- [97] will.i.am. Personal communication, August 6th 2003.
- [98] Matthew Wright. <http://ccrma.stanford.edu/matt/latencytest/>.
- [99] Matthew Wright. Problems and prospects for intimate and satisfying sensor-based control of computer sound. In *Sensors and Input for Multimedia Systems (SIMS) '02*, Santa Barbara, CA, June 20-21 2002.
- [100] Hiroyuki Yanase. Electronic pad. US Patent # 6822148, Filed December 2001. Assignee: Roland Corporation, Osaka, Japan.
- [101] Kiyoshi Yoshino, Kenji Hirano, and Masato Katsuda. Electronic percussion instrument with a net-like material to minimize noise. US Patent # 5920026, Filed July 1997. Assignee: Roland Kabsuhiki Kaisha, Osaka, Japan.