uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Emile Pelletier

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Mathematics)

GRADE / DEGREE

Department of Mathematics and Statistics

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Instrument De-synthesis Using Wavelets

TITRE DE LA THÈSE / TITLE OF THESIS

R. Vaillancourt

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

T. Giordano

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

M. Ashino

Y. Billig

Gary W. Slater

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

INSTRUMENT DE-SYNTHESIS USING WAVELETS

By
Emile Pelletier, B.A.
March 2005

A Thesis
submitted to the School of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of
Master of Science in Mathematics[1]

[1]The M.Sc. Program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Your file Votre référence
ISBN: 0-494-11380-4
Our file Notre référence
ISBN: 0-494-11380-4

Abstract

Our point of departure is the concept of 'additive synthesis', which is the traditional explanation for the individual of 'timbre' or 'colour' of the sound of the various musical instruments.

When an instrument sounds a note, one hears the note as if by itself, but this is not what is physically happening. What is in fact occurring is a complex waveform featuring a collection of harmonic frequencies, referred to as the spectrum.

A synthesizer attempts to imitate the sound of a particular instrument by replicating the amplitudes of its harmonics.

We use the term 'de-synthesis' to refer to the inverse procedure, computerized instrument identification.

We describe an experiment that we designed and executed with MATLAB to explore the hypothesis that a computer will be able to recognize an instrument by its characteristic timbre.

The idea of applying wavelets to analyze music comes naturally since music consists of sound waves, and wavelets are wave shaped functions.

We propose a mathematical model that can take certain musical instrument's attack and decay features into account that utilizes Malvar wavelets: Super Malvar wavelets. Super wavelets are superpositions of ordinary wavelets in some linear combination that can be treated as a wavelet in itself.

This provides us with a way to both represent the harmonics that are present in the notes and to deal with their individual attack and decay envelopes.

Instrument recognition is a problem that has had some research done by researchers in varying fields such as musicology, but also mathematics and computer science. We present a method that uses Super Malvar wavelets to identify the four musical instruments: the clarinet, the oboe, the trumpet and the violin. By using Super Malvar wavelets in this regard we hypothesize that we will have improved accuracy in the identification process because we will be considering the instrument's harmonic attack and the decay non-synchrony characteristics. Non-synchrony of the harmonics is a term that refers to the physical property that the higher numbered harmonics (≥ 2) appear later in the attack portion of a note and disappear sooner during the decay than the fundamental or first harmonic.

Acknowledgements

I would like to thank my supervisors, Thierry Giordano and Remi Vaillancourt. Also, I would like to thank Mackenzie Ells for her moral support and thanks to my friend Tim Prendergast who helped me with MusiXTEX.

Dedication

I dedicate this work to my father and mother, Donald Pelletier and Joan Wick Pelletier.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Every civilization that has made music has used ingenuity for the creation of musical instruments. People like playing instruments to go along with their singing and dancing traditions whether it be by drumming rhythms on home-made percussive instruments to create music or making pitch definite sounds by blowing into carved out bones [12, pp. 27, 36a] as early as 82000 years ago. Bob Fink's Theory of Harmony explains why different cultures seem to have come up with similarities in the instruments they made. He explains this coincidence in terms of what it is that makes humans tend towards instrumentalizing. Something which we will define in Chapter 2, the harmonic overtone series, an infinite series of rising notes, is the common starting point for these cultures. This is what binds all styles of music from every culture [11].

Fink's argument entails the first 7 harmonics; within these, there are four 'prime' notes for melodic construction. His claim is that these 'prime' notes appear in the popular tunes across many cultures from many time periods and that the common links between them are not mere coincidence.

A hypothesis stemming from this idea is that the first 7 notes in the harmonic overtone series could be made into pleasant sounding music in, and of, themselves. As an example of what the first 7 harmonic overtones sound like collectively—only 4 of them are distinct notes while some are repeated but in higher octaves —a guitar can be tuned in a manner that the distinct notes within the first 7 harmonics in the overtone series having as base frequency the lowest note regularly playable on the guitar, are on the 6 strings of the guitar, but with some of them lowered by one or two octaves. When all 6 strings are strummed with the fingering such that the notes lie in the correct octaves in the harmonic overtone series their sound is pleasant and not offensive. However, when all six strings are strummed without fingering, the sound is dissonant, almost appalling to the ear, even though they are the same notes transposed one or two octaves and this means they are supposed to sound the same.

One idea that exemplifies what type of differences arise when sounds are of different harmonic breakdown, something that is the subject matter in Section 2.2, is to compare and contrast the different timbres attainable by the organ. The organ is made from pipes of varying materials, usually metal or wood, that can affect the timbre of the instrument; it also permits any combination of musically related pipes to sound with a single key strike, which is a true form of what is known and which is explained below as additive synthesis. As can be seen with the organ, instrument designers utilize the relationships of various harmonics to affect the sound of the instruments they make. The term timbre or colour is used to describe the sound of a single musical tone and the emotion it elicits. Note that instruments tend to have characteristic timbres, recognizable to humans. For example, the oboe has a sweet timbre and the cello has a warm colour.

Some instruments can be synthesized by measuring the intensities of each individual harmonic in the overtone series in notes played by the instruments and recreating these using a computer. This is known as *additive synthesis*: adding various frequencies in various intensities or amplitudes together in an attempt to synthesize an instrument's sound. The first section of this thesis sets out to understand how this principle can be used to analyze sounds. It is hypothesized that by analyzing the harmonics in a given note, a computer will be able to recognize the instrument by its characteristic timbre.

1.1 Preamble

To aid in achieving the goal of Artificial Intelligence (AI), dissecting human hearing and seeing processes automatically into all of their components poses a problem mathematically, psychologically, and sociologically to us because of the complexity of those processes. However, the more we learn about the psychological and sociological aspects of sight and sound, the more capable we become of programming what we know into computers.

Science, in its attempt to explain our world or create technology, does so, in fact, while remaining within the boundaries of our current understanding of the world and state of the art technology. The boundaries now at hand recur in what are not the most difficult problems to understand; it is often rather what is most natural to us that poses the biggest problems in our understanding. Diverse research areas focus on de-constructing what for humans are reflexes and re-synthesizing these reflexes artificially.

> Despite the progress in speech recognition systems, most recognizers are defeated by interfering background noise of a kind that would present little challenge to people [18].

For example, when a computer was used to try to separate two musical parts in a vocal piece accompanied by guitar, which would effectively constitute a form of recognizing one instrument from another, it was able to obtain the voice part alone but not the guitar part; traces of the voice remained [18].

1.2 Purpose

For the purpose of this thesis, suppose we have a music signal which is a duet between the clarinet and violin but we are interested only in the clarinet part. As listeners, when these two instruments exchange in a musical dialog, we are quite aware of the clarinet playing and the violin playing as separate lines. Over the last fifteen years there has been much progress in the development of mathematical tools for the analysis of digital musical signals, useful for a wide range of problems and with specific tasks in mind [8]. The short-time Fourier transform has permitted us to transform a signal from its original form into a representation of it in terms of its frequency components. The use of such a tool is widespread in many areas such as engineering, physics, biology, mathematics, but it is especially important for musical purposes due to the role that frequency plays in music. We present ideas for new tools.

In Chapter 4, we experiment with four instruments: the clarinet, the oboe, the trumpet and the violin, with particular emphasis on the violin and the oboe. The reason for this is that, in an experiment [19] on human identification of musical instruments, the two instruments most commonly confused for one another were the violin and the oboe (see Tables 1 and 2). From Table 1, one can also note that there is a high rate of misidentification between the flute and saxophone pair. In Section 3.3, we propose to use Malvar wavelets, or lapped orthogonal transforms (LOT), for the analysis of musical instruments. This method is chosen for its possibility of analyzing instrument features during the attacks and decays.

1.3 Experiments with Synthesizers

> Timbre has been defined by the American Standards Association as "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar ... Timbre depends primarily upon the spectrum of the stimulus, but it also depends upon the waveform, the sound pressure and the frequency location of the spectrum, and temporal characteristics of the stimulus"[1] [21, p. 14].

In this section we present some computer experiments dealing with instrument recognition conducted by the author. Two experiments were conducted. The experiments were based on a project that was presented in a seminar course on Fourier Series and Wavelets at Carleton University. The idea for the project was conceived

[1]Grey, John M. *An Exploration of Musical Timbre.* Ph.D. Thesis, Stanford University, 1975. Distributed as Dept. of Music Report No. Stan-M-2. P. 1

after experimenting with the synthesis of an instrument's sound using the harmonic overtones with a connection to the Fourier sine series. Fourier series and the Fourier transform are topics in Chapter 2. "Traditionally, musical tones have been divided into three major sections: [attack, sustain and decay]. The first major methodical research into timbre was carried out by Helmholtz, who concluded that timbre is determined by the relative strengths of the harmonics during the steady-state portion of a tone" [21, p. 14].

We outline here a method to synthesize a piano playing a note with fundamental frequency F. The method consists of defining a function f,

$$f(t) = b_1 \sin(Ft) + b_2 \sin(2Ft) + \cdots + b_7 \sin(7Ft), \tag{1.3.1}$$

where the coefficients, $b_1, \ldots, b_7$ are chosen in such a way as to model the sound of the piano, from the relative strengths of the harmonics from the analysis of a piano tone.

> Helmholtz's work was based on the discovery by a young physicist named Fourier that a *periodic* waveform can be analyzed and represented as a sum of simpler waveforms. This "Fourier analysis" assumes that the constituent frequencies of a complex spectrum are all whole-number multiples of a "fundamental frequency" of the complex waveform. Mathematically, this can be expressed as
>
> $$e = C_0 + C_1 \sin(f + \phi_1) + C_2 \sin(2f + \phi_2) + \cdots + C_i \sin(if + \phi_i) \tag{1.3.2}$$
>
> or
>
> $$e = C_0 + \sum_i C_i \sin(if + \phi_i)$$
>
> where
>
> $$\begin{aligned} f &= \text{fundamental frequency} \\ 2f, 3f \ldots &= \text{frequencies of harmonics} \\ \phi_1, \phi_2 \ldots &= \text{phases of harmonics} \\ C_1, C_2 \ldots &= \text{amplitudes of harmonics.} \end{aligned}$$

> [...] The question of the role of phase (ϕ in Eq. [(1.3.2)], above) raged for a long time. [...] It is now agreed that phase plays a relatively minor role in determining timbre, "and that the maximal difference due to phase is about equivalent to the changes in quality between very closely related vowels"[2] [21, p.14].

Thus, Equation (1.3.1) comes from this theory without phase, $C_0 = 0$ and $b_1, b_2, \ldots, b_7$ are the amplitudes of the first seven harmonics during steady-state portion of the piano tone.

1.3.1 Motivation

We have seen that every instrument has a specific timbre that makes it recognizable as having a particular sound. It helps to think of the analogy with vowels in the English language. Although one vowel can have several different sounds in different contexts, we learn to recognize vowels mainly because the list of vowels is finite and we have a substantial amount of practice. Provided we are familiar with an instrument that is playing, recognizing a musical instrument by its sound can be like the 'recognizing' of vowels in a familiar language.

Due to the difficulty of the problem when the identification is to be made by computer program—Gerhard, 2003, at the University of Regina, has shown that computers may be unable to distinguish spoken words from sung words while even babies of a very young age learn this difference—we have chosen to use synthesizer instruments for the computer's training set to simplify the experiment. In one experiment, real instruments were used to test the hypothesis.

1.3.2 Hypothesis

It is hypothesized that computers will be capable of distinguishing the sounds of synthesizer instruments with a certain degree of accuracy by comparing their harmonic spectrums. Moreover, we conjecture that the accuracy of the identification process will be further improved by characterizing the pitches and dynamics of the test notes.

[2] Grey, 1975, p. 6

1.3.3 Purpose

The purpose of the experiments was to see whether a computer could distinguish an instrument from other instruments given a training set based on the amplitude spectrum.

Such knowledge would be progress toward the development of a computer program that takes a musical recording and produces a written score where the distinct instrument parts are written on separate lines, as it would appear in the original. Due to the differing natures of the sound of an instrument in its various registers and articulation styles, we wanted to see whether the effects of change in dynamic and pitch, within an instrument's range, were factors in the identification accuracy.

1.3.4 Research

In a 1975 experiment [19] conducted on graduate music students, the subjects were made to identify, from a recorded note, the instrument, given a list of possibilities. The purpose was to identify the role that attack and decay portions of a note had in human identification of musical instruments. This was done with real instruments recorded playing the note $E3$ four times in two different durations: twice for 7 s on tape A and twice for 6 s on tape B. All the tones on tape A had the first and last 0.5 s removed which resulted in 6 s long tones on both tapes. Those on tape B were left unaltered. There were 9 instruments for a total of 18 test tones in random order on each tape and 57 subjects in this experiment. The names of the 9 instruments involved were given to the subjects along with two answer sheets each where 18 blanks were to be filled with their guess as to which instrument was playing each of the tones. The test was conducted in two stages, in which the subjects listened to the two tapes, first tape A then tape B. All tones were 6 s long in duration and followed by 8 s of silence. The cumulative results from all 57 subjects for the altered tones as well as the unaltered tones are shown in tables in Tables 1 and 2.

One conclusion to be drawn from this is that a computer program that identifies instruments should take into account attack and decay features that various instruments possess. It is believed that successful part separation requires that the

Table 1: Responses of Subjects by Instrument for Part A (Altered Tones) [19]

Instrument	Flute	Clarinet	Oboe	Bassoon	Saxophone	Trumpet	Violin	Cello	Trombone	No Response	Totals
Flute	33	7	4	5	30	7	2	6	15	5	114
Clarinet	1	102	4	1	3	1	–	–	–	2	114
Oboe	–	3	85	11	3	8	–	–	2	2	114
Bassoon	1	5	19	51	28	5	–	1	2	2	114
Saxophone	18	4	2	10	39	7	2	7	17	8	114
Trumpet	2	8	2	5	11	77	–	–	5	4	114
Violin	–	3	63	5	14	10	8	3	1	7	114
Cello	1	5	57	2	8	9	18	10	–	4	114
Trombone	3	2	–	36	8	6	1	3	51	4	114
Totals	59	139	236	126	144	130	31	30	93	38	

Table 2: Responses of Subjects by Instrument for Part B (Unaltered Tones) [19]

Instrument	Flute	Clarinet	Oboe	Bassoon	Saxophone	Trumpet	Violin	Cello	Trombone	No Response	Totals
Flute	90	4	–	1	6	–	2	9	–	2	114
Clarinet	2	105	–	2	3	–	–	1	1	–	114
Oboe	–	3	103	1	1	3	2	1	–	–	114
Bassoon	1	–	5	90	12	–	2	4	–	–	114
Saxophone	14	22	1	5	65	–	–	6	–	1	114
Trumpet	–	1	–	–	3	107	–	–	3	–	114
Violin	–	–	7	–	1	5	70	31	–	–	114
Cello	–	–	22	1	3	2	53	33	–	–	114
Trombone	1	–	–	2	–	5	–	2	103	1	114
Totals	108	135	138	102	94	122	129	87	107	4	

computer can recognize the instruments playing.

1.3.5 Procedure

All the tests were done using computer software, MATLAB. In both experiments, test tones coming from unknown instruments were compared, one at a time, to tones coming from known instruments that comprised the sample set. Using the discrete Fourier transform (DFT) for all of the tones, each of which were recorded digitally, the amplitudes of the first 7 harmonics could be estimated from the graphs of the magnitudes of the DFTs and were stored in vectors, $(b_1, b_2, b_3, b_4, b_5, b_6, b_7) \in \mathbb{R}^7$. For two specific tones, we defined as the cosine of these two corresponding harmonic amplitude vectors the similarity between the two tones. For instance, if x and y were the two vectors in $\mathbb{R}^7$ containing the amplitudes of the first 7 harmonics for two tones, then the similarity between them was given by

$$\cos(\theta) = \frac{x \cdot y}{\|x\| \|y\|}.$$

The closer this number was to 1, the more similar the two tones were deemed.

1.3.6 Yamaha Synthesizer Experiment

One experiment dealt with 10 different instrument sounds on an inexpensive synthesizer. Two samples tones (differing by an interval of a sixth) were taken for each sound. A third tone (at random pitch) was taken and used as a test tone for each instrument.

For every pair of test tone and sample tone in the set of data, the similarity quotient was computed and the results are shown in the 10×10 table found in Appendix A. The rows represent the test tone for each instrument, while the column represents the sample tones. The first and second numbers in each entry correspond to the test tone's similarity with the first and second sample tones, respectively. Using the average of the two numbers in each entry, if the highest number in the row falls on the main diagonal, then the identification was successful for that test.

This method could be altered for an experiment using real instruments. Every instrument has natural limits to the pitches that it can play, and this could be used to help identify an instrument: the pitch of the test tone could be identified before attempting to find the closest match from a set of sample tones. For example, if the test instrument was a flute, then the lowest possible note it could play is middle C on a piano, which would automatically rule out the bass, bassoon, and the tuba. If the tone was in a high register of the flute's range, then the list of possible choices would be greatly reduced, say, to only the violin, flute, piccolo, and trumpet as possible choices.

1.3.7 Technics KN6500 Experiment

In another experiment conducted by the Author, a state of the art synthesizer, Technics KN6000, was employed. A database consisting of twelve instrument sounds, with nine sample tones for each instrument, having different registers and dynamics (low/soft, low/medium, low/loud, middle/soft, middle/medium, middle/loud, high/soft, high/medium, high/loud) was made. Using the usual markings for pianissimo, piano, mezzo-piano, mezzo-forte, forte, and fortissimo in music (pp, p, mp, mf, f, ff), for each of the twelve instruments in the set, the nine different register/dynamic combinations can be seen in musical notation in Figures 12–14 in Appendix B. For example, for the nine violin tones, we have: low/pp, low/mf, low/ff, med/pp, med/mf, med/ff, high/pp, high/mf, high/ff.

Tests were conducted on the violin, oboe, bassoon, flute, clarinet, cello and bass. For each test, excerpts of solo passages for one instrument from classical music were used. Only one note coming from one of the instruments was taken as the test tone in each case. Results can be found in Appendix B.

In the tables, each column represents one instrument in the sample set while each row represents one of the nine possible combinations of pitch and dynamic. Every entry in the 9×12 sub-matrix constitutes the similarity quotient of the test tone to the sample tone corresponding to the register/dynamic and instrument, row and column, respectively. The last row gives the mean of all the entries in the column;

the entries are a general similarity between the test tone to the instrument in that column.

1.3.8 Conclusion

The **Yamaha Synthesizer Experiment** was successful 9 times out of 10, and the conclusion can be made that the method of comparing harmonic spectrums to identify instruments was successful for synthesizer instruments. The one misidentification occurred with the oboe. The reason for this error is not known; it could be attributed to a faulty test. A natural question to ask is whether similar experiments that use larger lists of real instruments instead of synthesizer instruments would yield as good results? There are some advantages to having used an inexpensive synthesizer for the experiment: first, a synthesizer relies on additive synthesis as its method of approximating an instrument's sound so the method used in the experiment is particularly well suited for synthesizer tones; second, the less expensive a synthesizer, the less variations within one patch (instrument) are produced by that synthesizer than for a more expensive synthesizer due to the lower amount of computational power from one synthesizer to the next. This will have had an effect on the results because the two notes, the test note and one of the sample tone, played on a cheaper synthesizer will not differ the way they might have if they were played on a more expensive one. In this regard, it is easier for the computer to tell the difference between patches of an inexpensive synthesizer than between those of a more expensive one.

These conclusions can be further supported by the results from the **Technics KN6500 Experiment**. The Technics synthesizer was a much higher quality instrument than the Yamaha synthesizer. Using the means of the entries in each column corresponding to the nine sample tones, shown in the bottom rows of Tables 7–14 in Appendix B, for each instrument in the sample set, the computer identifications were correct for two of the seven experiments, namely in the Bassoon test tone 2, and the Clarinet test tone. In these two cases, although the means of the nine sample tones were highest for the correct instruments, for each individual row corresponding to a particular register/dynamic combination, incorrect instruments had higher similarities than the correct instrument. This meant that, although the most similar instruments from the sample set to the test tone were the correct instruments in these two cases, in one particular register/dynamic combination, it was closer to one of the other twelve instruments. Although there were few correct identifications, it does not undermine the validity of the hypothesis because the variety of the test tones could have accounted in part for the fewer correct identifications: the seven test tones were taken from excerpts of classical symphony music recordings. For the test tones, we attempted to isolate one instrument playing one note during quiet passages however they may not have been completely isolated from other instruments such as background instrument noises.

Chapter 2

The Harmonic Overtone and Fourier Sine Series

Dating as far back as known, every musical system is based on dividing the octave, the most natural sounding interval, "[which] has the mathematical ratio of 2 to 1, "1" being a symbol-number to represent the number of vibrations of any original note" [12, p. 4]. More recently, mathematics has been applied to and used to explain music composition [12, p. 6]. In this chapter, we explore some of the mathematical properties in music, beginning with some history.

In Western music, since the invention of equal temperament tuning of instruments, there are 12 equal steps that make up one octave, each of the steps called semi-tones which constitute a factor of $2^{\pm 1/12}$ in frequency, the plus or minus depending on whether the semi-tone is rising or falling. Rising by one semi-tone twelve times produces the note one octave above the starting note, since

$$2^{1/12}2^{1/12}\cdots 2^{1/12} = (2^{1/12})^{12} = 2.$$

From any starting note, displacing upwards or downwards by the same number of semi-tones constitutes equal factors in frequency, hence the term equal temperament. For example, if you start at middle C on a keyboard that has frequency 264 Hz and displace yourself upwards or downwards three semi-tones, the higher note has a frequency of $2^{3/12}264$ Hz whereas the lower note has a frequency of $2^{-3/12}264$ Hz. It is thus a completely symmetric musical system with the benefit of maximum flexibility for manipulating harmony and melody. The word temperament coming from the Latin word *temper* has the meaning to tamper with the pure tuning which relies completely on rational numbers. For every whole number ratio in frequency there is an acoustical effect, often pleasing, consonant sounds. In older tunings, without the symmetry of equal temperament, the octave would still have constituted a factor of two in frequency, but other intervals within the octave, such as the interval of a fifth, would have been 'perfect' or 'pure' ratios of frequency, in this case, $\frac{3}{2} = 1.5$, whereas the corresponding ratio of frequency for the fifth in equal temperament is $2^{7/12} \approx 1.4983$ [24].

Although better sounding in one key than equal temperament would sound in any key—because of its symmetrical properties, equal temperament sounds the same no matter which key the music is in— pure tuning for the keyboard instruments sounded out of tune in many other keys. As a result, the more complex music became—modulating from key to key throughout the same piece of music—the harder it became to use pure tuning and equal temperament eventually became common practice. The reason that pure tuning on keyboard instruments was problematic is due to an internal conflict which can be attributed to the incommensurability of the numbers 2 and 3, that no power of two can equal a power of three.

The Western music system relies on the fact that consecutively rising by the interval of a fifth twelve times should bring you back to the same note several octaves higher [24]. Therefore, this frequency should be a power of two times the starting frequency. However, consecutively rising by a fifth in pure tuning constitutes a factor of $(3/2)^{12} \approx 129.75$ in frequency whereas it should equal $2^7 = 128$ which corresponds to the factor of frequency for a note seven octaves higher than the starting note. Comparatively, rising by a fifth twelve times in equal temperament constitutes a factor of $(2^{7/12})^{12} = 2^7 = 128$ in frequency. This approximate result entails the problem of incommensurability.

Table 3: Twelve harmonics starting on C1.

Fundamental C1 33 Hz				
2nd harmonic C2 66 Hz	3rd harmonic G2 99 Hz			
4th harmonic C3 132 Hz	5th harmonic E3 165 Hz	6th harmonic G3 198 Hz	7th harmonic B♭3 231 Hz	
8th harmonic C4 264 Hz	9th harmonic D4 297 Hz	10th harmonic E4 330 Hz	11th harmonic F♯4 363 Hz	12th harmonic G4 396 Hz

2.1 Theoretical Aspects

For a given starting note, whose frequency we will call the fundamental or base frequency, the *harmonic overtone series* is the infinite series of notes whose frequencies are integer multiples of the base frequency. The first of these is the starting note itself and is referred to as the fundamental. Each subsequent harmonic is called by its placement in the series: therefore we have the series consisting of the fundamental, then the second harmonic with frequency twice that of the base frequency, then the third with frequency three times the base, etc. For an example of a harmonic series with base note C1 on a keyboard that we will assume has a frequency of precisely 33 Hz, based on A4 having frequency 440 Hz and using pure tuning to find a frequency of C1 (the frequency ratio of a major sixth as $5:3$), the first 12 notes and their corresponding frequencies in the harmonic series are shown in Table 3.

Remark. In Table 3, all the harmonics in the same row lie in the same octave beginning on a C, and the n^{th} row corresponds to $(n-1)$ octaves above C1. For instance, in the third row, we have the fourth, fifth, sixth and seventh harmonics, each of which lies within the octave between C3, two octaves above C1, and C4.

This naturally occurring phenomenon is not consistent with equal temperament tuning: although equal temperament tuning was chosen for its symmetric properties, it was in fact accepted only because it was a close enough approximation to the authentic, pure tuning. We neglect the inconsistencies between harmonics and equal temperament tuning throughout the remainder of this thesis.

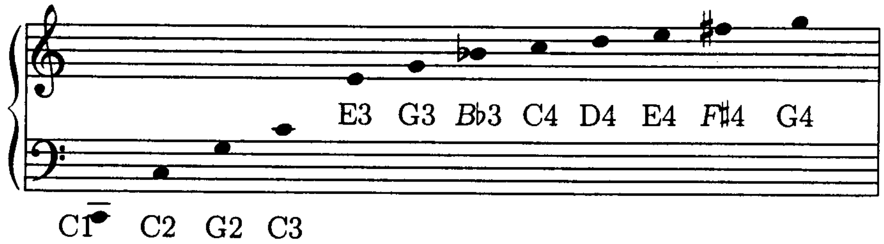

Figure 1: Twelve harmonics starting on C1.

The first 6 harmonics match up with notes that outline the first, third and fifth (a major chord) in the key of the fundamental. Here, we use the conventional terminology: any note equal to the fundamental or any number of octaves above or beneath the fundamental is called the first in the key of the fundamental; the second, third, fourth, fifth, sixth and seventh are then the notes of major scales starting on the first. Table 4 describes the first 6 harmonics in these terms and also give the frequency ratios with respect to the nearest first lower in the series. For example, the third harmonic is in a ratio of $3:2$ with the second harmonic which is the nearest harmonic lower in the series that is a first.

As the harmonic series progresses, the notes become closer and closer together to the point that the notes produced by the series fall in between the notes in Western tonality. This means that some of the notes in the harmonic overtone series do not have names other than their number in the series. Harmonics with numbers less than twelve almost all lie within the tonal system with the exception of the seventh and eleventh harmonics which are close in pitch to notes in the system. In this example, they are approximately the notes B♭3 and F♯4, respectively, which is why they appear in slanted type in Table 3 and Figure 1.

Table 4: Six harmonics outlining a major triad.

1	2	3	4	5	6
First	First	Fifth	First	Third	Fifth
1 : 1	2 : 1	3 : 2	2 : 1	5 : 4	3 : 2

A useful bit of information to help with the letter names of the notes in the series is that harmonics follow a geometric pattern: each of the harmonics whose numbers follow a geometric progression of powers of two times an odd number is the same note. In our example with C1 as the starting point, we have that the harmonics with numbers $\{1, 2, 4, 8, 16, \ldots\}$ are all C; harmonics with numbers $\{3, 6, 12, 24, \ldots\}$ are all G, etc.

2.2 Fourier Sine Series

Given a note with frequency F, then representing it by a simple sine wave as a function of time,

$$\sin 2\pi F t$$

we can then easily represent the notes in the harmonic series starting on that note as the terms in the Fourier sine series,

$$\sum_{n=1}^{\infty} b_n \sin 2\pi n F t$$

since for any $n \geq 1$ the frequency of the nth term in the series, $\sin 2\pi n F t$, is

$$\frac{2\pi n F}{2\pi} = nF,$$

which is n times the fundamental frequency, F, and therefore the nth harmonic.

When an instrument sounds a note, one hears the note as if by itself, but this is not what is physically happening. What is in fact occurring is a complex waveform featuring a collection of harmonic frequencies, referred to as the spectrum. Although not readily apparent, the amplitudes of the harmonic frequencies b_n, $n = 1, 2, 3, \ldots$ are mostly all non-zero up to some number, with a general decreasing tendency of amplitudes corresponding to higher harmonics in the series and this affects the timbre of the note. Even though the pitch heard is associated with the frequency corresponding with the first term, some experiments have shown that this term may not have the greatest amplitude in the spectrum. Indeed some instruments tended to portray this surprising feature, such as the synth-bass on the **Yamaha** synthesizer.

This is similar to the way in which synthesizers work in trying to replicate the amplitudes of the harmonics, a method called additive synthesis. In fact the **Hammond** organ is an instrument which allows the user to set the amplitudes of each harmonic independently in order to come up with different sounds. There, it is commonly the first 9 harmonics that are controllable.

2.3 Fourier Series and Fourier Transform

In this section we define the Fourier series representation of a 2π-periodic function. If $f : \mathbb{R} \to \mathbb{C}$ is 2π-periodic, $f \in L^1(0, 2\pi)$, then we can write the trigonometric and complex Fourier series representations or expansions of f,

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \tag{2.3.3}$$

and

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}, \tag{2.3.4}$$

respectively, where

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\, dx, \qquad n = 0, 1, 2, \dots, \\ b_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\, dx, \qquad n = 1, 2, 3, \dots, \end{aligned}$$

and

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ikx}\, dx, \qquad k = 0, \pm 1, \pm 2, \dots.$$

If $f \in L^2(0, 2\pi)$, then the Fourier series converges to f in the following sense,

$$\int_0^{2\pi} \left| f(x) - \sum_{-N}^{M} c_k e^{ikx} \right|^2 dx \underset{\substack{N \to \infty \\ M \to \infty}}{\longrightarrow} 0.$$

Remark. This convergence does not imply that the Fourier series expansion of f converges to the value of $f(x)$ at all points $x \in [0, 2\pi]$, rather it may fail to converge to $f(x)$ at infinitely many values of x in the interval [2, p. 90].

When speaking of Fourier series, the following definitions from [2, pp. 26, 93] are common and necessary for the Theorem 2.3.3.

Definition 2.3.1 *A function f is said to be piecewise continuous in an interval $a \leqq x \leqq b$ if there exist n points $a = x_1 < x_2 < x_3 < \cdots < x_n = b$ such that f is continuous in each interval $x_i < x < x_{i+1}$ and has finite one-sided limits $f(x_i+)$ and $f(x_{i+1}-)$ at the endpoints of each such interval $(i = 1, 2, \ldots, n-1)$.*

Definition 2.3.2 *A function f is said to be piecewise smooth if it is piecewise continuous, and, in every interval (a, b) of finite length, it has a derivative (except perhaps at a finite number of points) that is piecewise continuous.*

We now state the following theorem from [2, p. 122].

Theorem 2.3.3 *The Fourier series representing a piecewise smooth function $f(x)$ in $0 \leqq x \leqq 2\pi$ converges to $[f(x+) + f(x-)]/2$ whenever $0 < x < 2\pi$ and to $[f(0+) + f(2\pi-)]/2$ when $x = 0$ or $x = 2\pi$.*

These results hold for the trigonometric as well as the complex Fourier series expansions of f.

For a function $f \in L^2(\mathbb{R})$, we define the Fourier transform (FT):

$$\begin{aligned} FT: \ & L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R}), \\ FT: \ & f \longrightarrow \hat{f}, \end{aligned}$$

where

$$\hat{f}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\, e^{-i\omega x}\, dx. \tag{2.3.5}$$

We can recover f from the transform by

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega x}\, d\omega,$$

which is called the inverse Fourier transform (IFT).

In audio signal analysis, signals are discrete with a given sample rate, say 22050 samples per second. A discrete audio signal is a sequence of numbers where the distance between each number in time is $h = \frac{1}{22050}$ second. The discrete Fourier transform of a signal f is:

$$\{f_1, f_2, f_3, ..., f_N\} \longrightarrow \{\hat{f}_1, \hat{f}_2, \hat{f}_3, ..., \hat{f}_N\}.$$

where

$$\hat{f}_k = \sum_{m=1}^{N} f_m e^{-2i\pi mk/N}, \qquad \text{with corresponding frequency } \omega_k = \frac{k}{T}.$$

Note that the signal is considered periodic with period, $T = Nh$.

Definition 2.3.4 *The Nyquist frequency of a signal is half the sampling rate.*

A theorem called the sampling theorem, which for its relevance to communication theory and signal processing is attributed to Claude Shannon (1949), proves that a function f can be completely and accurately represented with a discrete signal with a given sample rate if that function contains no frequencies higher than the Nyquist frequency. Equivalent forms appeared in mathematical literature much earlier (Whittaker, 1915; Kotel'nikov, 1933) [10].

The next chapter introduces briefly the definition of wavelets citing results from various different sources and we mention some of the applications of wavelets in 1-D discrete signal analysis. In Section 3.6, we introduce the concept of Super Malvar wavelets, superpositions of Malvar wavelets which are the topic of Section 3.4. Introduced by Yves Meyer, Malvar wavelets begin by a segmentation of a signal–the signal is 'cut into pieces'–and perform a Fourier analysis on the pieces [13, p.102], which in turn requires the computation of the coefficients or inner products.

Chapter 3

Super Malvar Wavelets

A wavelet is a time-frequency atom [7]. One of the main uses of wavelets is discrete signal compression by means of the discrete wavelet transform (DWT). A signal can be measured in size by the number of bits used to represent it. Depending on the wavelets, upon transformation, the new form of the signal could have some desired properties such as many near zero terms, whereas the original signal did not. The many near zero terms can be set to zero; therefore the transformed signal can be coded with a few bits. Although the transformed signal is represented differently, the original signal can be recovered from the new representation with little loss, when the wavelet used is known, by means of the inverse discrete wavelet transform (IDWT). Music signals can be compressed by using the DWT and IDWT to reduce the number of significant bits: this is done using an algorithm that performs the following steps.

(i) Discretize the signal (if it is not already in digital form).

(ii) Transform the signal using the DWT with a chosen wavelet.

(iii) Remove the strings of near-zero terms in the transformed signal at the time of compression replacing them with a few bits that encode their numbers and positions.

(iv) Replace them with 'true' *zeros* in the correct locations.

(v) Transform the modified signal using the inverse discrete wavelet transform (IDWT) with the same wavelet to recover the signal.

Assuming that the number of bits used to represent strings of near-zero terms is negligible compared to their original numbers in the transformed signal, then the compression ratio is defined to be the number of significant bits (bits that are non-zero) in the original signal divided by the number of significant bits in the modified signal. This number can sometimes be as large as ten or higher because the transformed signal possesses many near zero-terms. At the end of the entire process, although the recovered signal had been transformed using the DWT in step **(ii)**, modified slightly by steps **(iii)** and **(iv)** changing near-zero terms into 'true' *zeros* and then transformed again by the IDWT in step **(v)**, the new music signal and the old music signal are practically the same.

3.1 Overview on Wavelets

For a family of functions $\psi_{b,a}(t)$ to be called wavelets, they must be the translations and dilations of a 'basic wavelet' $\psi(t)$, a real-valued function of time, t, that exhibits an oscillatory characteristic like all waves must, and is localized to one segment of the time axis. The oscillatory condition is that the first m moments vanish. This requires its integral over the entire time axis, as well as the remaining $m-1$ moments equal zero,

$$0 = \int_{-\infty}^{\infty} \psi(t)\,dt = \cdots = \int_{-\infty}^{\infty} t^{m-1}\psi(t)\,dt.$$

The family of scaled and shifted functions are then defined as

$$\psi_{b,a}(t) = \frac{1}{\sqrt{a}}\,\psi\left(\frac{t-b}{a}\right), \quad b \in \mathbb{R}, \quad a > 0,$$

where a and b are the scaling and shifting parameters, respectively, [13, p. 37].

For functions $f \in L^2(\mathbb{R})$ of one variable, the continuous wavelet transform (CWT) is a function of two variables (a, b) defined for $b \in \mathbb{R}$, $a > 0$, by the integral [14, p. 5]

$$W_\psi f(b,a) = \int_{-\infty}^{\infty} f(t)\,\psi_{b,a}(t)\,dt.$$

It is seen that the continuous wavelet transform is an operation that takes an $L^2(\mathbb{R})$-function, f, and computes the integral of f times the translations and dilations of the mother wavelet, $\psi_{b,a}$. This transform uses the wavelets in varying scales (dilation) and translations to analyze the various scales of a signal in its entire domain [7].

A mother wavelet, $\psi \in L^2(\mathbb{R})$, whose Fourier transform satisfies the 'admissibility condition':

$$C_\psi := \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|}\, d\omega < \infty,$$

is called a 'basic wavelet' [14, p.9]. If $\psi \in L^1(\mathbb{R})$, then $|\hat{\psi}(\omega)|$ is continuous at $\omega = 0$ and, in turn, this requires that

$$\hat{\psi}(0) = 0,$$

and from the definition of the Fourier transform (2.3.5) from Section 2.3, this implies that

$$\int_{-\infty}^{\infty} e^{i0t}\psi(t)\, dt = \int_{-\infty}^{\infty} \psi(t)\, dt = 0,$$

which says that the first moment of ψ vanishes [9, p. 24].

For $f \in L^2(\mathbb{R})$, we can recover f from the CWT with the inverse continuous wavelet transform (ICWT) defined by the double integral [14, p.10]:

$$f(x) = \frac{2}{C_\psi} \int_0^{\infty} \left[\int_{-\infty}^{\infty} \{(W_\psi f)(b,a)\} \left\{ \frac{1}{\sqrt{a}}\, \psi\left(\frac{x-b}{a} \right) \right\} db \right] \frac{da}{a^2}.$$

One sees that C_ψ must not be zero nor infinity.

The idea of applying wavelets to music comes naturally since music consists of sound waves, and wavelets are wave shaped functions. We will define Malvar wavelets in Section 3.3. Figure 2 is a graph of a Malvar wavelet. They resemble musical notes in two ways. First, the shape of the envelope, the outer line, resembles the gradual beginning and end of a sound. Second, the oscillatory motion—the fluctuations inside—determines the *specific frequency* of the note. In Figure 2, the outer line is called the envelope and can be filled with either sines or cosines but not both [17]. One specific feature that Malvar wavelets possess that other wavelets do not is a physical meaning. According to Yves Meyer [17, p. 78], one of the mathematicians who introduced Malvar wavelets, it is the concept of a sound having a *specific duration*,

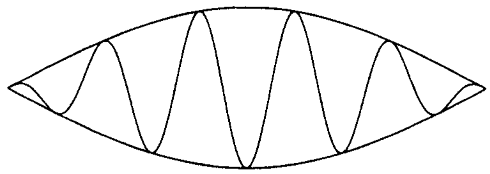

Figure 2: Malvar wavelet and envelope.

with a beginning and an end. Malvar wavelets, as opposed to other wavelets, show promise as a tool in speech and acoustics. The family of Malvar wavelets allows a wavelet representation of music signal with reasonable time-frequency localization [3], "they are wavelets that are very well localized in frequency, very different from the wavelets used in image processing," says Stéphane Mallat [17, p. 204]. As Meyer wrote:

> Tout signal admet une infinité de représentations et ceci nous amène à choisir la meilleure d'entre elles, selon un critère [13, p. 87].

In this thesis, we expand on the idea of a wavelet having specific frequency and duration to wavelets that also have a specific timbre. We call them Super Malvar wavelets or SMW's. We will try to use them to identify which of four instruments (the clarinet, the oboe, the trumpet and the violin) is playing. For this, we will use music test signals in which one note is played at a time and in which each note is separated by some small amount of silence. In the following sections of this chapter, we define Malvar wavelets and introduce Super Malvar wavelets while following mainly the work in the two texts [13] and [1].

3.2 Orthogonal Projections

In this section we recall the notion of orthogonal projections, mutually orthogonal projections and develop mutually orthogonal projections on half-lines which are necessary in the development of the theory for Malvar wavelets.

Definition 3.2.1 *Let $\mathcal{H}$ be a complex Hilbert space and W an arbitrary closed subspace. An orthogonal projection, or simply a projection, $P : \mathcal{H} \longrightarrow W$, is a continuous* *linear operator which is idempotent and self-adjoint, $P^2 = P$ and $P^* = P$, respectively.*

We state the following well-known lemma.

Lemma 3.2.2 *Let P and Q be orthogonal projections such that $PQ = QP$, then QP is an orthogonal projection.*

Recall that two projections P_1 and P_2 are mutually orthogonal if

$$P_1 P_2 = 0.$$

Definition 3.2.3 *An increasing profile function β associated to the interval $(-1,1)$ will denote a smooth monotone increasing function β such that*

$$\begin{aligned} \beta(t) &= 0, & t &< -1, \\ \beta^2(t) + \beta^2(-t) &= 1, & -1 \le t &\le 1, \\ \beta(t) &= 1, & t &> 1, \end{aligned} \tag{3.2.6}$$

and the associated decreasing profile function, $\widetilde{\beta}$, will be defined by $\widetilde{\beta}(t) = \beta(-t)$ ($\widetilde{\beta}$ is the reflection of β about zero).

Definition 3.2.4 *Given any monotone increasing profile function β associated to the interval $(-1,1)$, with any function f we associate the function $P^+ f$ (resp. $P^- f$) given by* [1]

$$P^+ f(t) = \beta(t)\left[\beta(t)f(t) + \widetilde{\beta}(t)f(-t)\right] := \beta(t)p(t), \tag{3.2.7}$$

and

$$P^- f(t) = \widetilde{\beta}(t)\left[\widetilde{\beta}(t)f(t) - \beta(t)f(-t)\right] := \widetilde{\beta}(t)q(t), \tag{3.2.8}$$

where $p(t)$ and $q(t)$ are even and odd functions defined by the middle terms in (3.2.7) and (3.2.8), respectively.

Lemma 3.2.5 *The two operators, P^+ and P^-, are mutually orthogonal projections, whose sum is the identity.*

Proof.

We first need to show that P^+ and P^- are in fact projections, meaning that they are idempotent and self-adjoint.

To show that P^+ is idempotent, it suffices to show that $P^+(P^+f) = P^+f$. Indeed, by the definitions of β and P^+f, if $t \notin [-1,1]$,

$$P^+(P^+f(t)) = \begin{cases} 0 = P^+f(t) & \text{if } t < -1, \\ f(t) = P^+f(t) & \text{if } t > 1; \end{cases}$$

and, if $t \in [-1,1]$, then $P^+f(t) = \beta(t)p_0(t)$, and from (3.2.7),

$$\begin{aligned} P^+(P^+f(t)) &= P^+(\beta(t)p_0(t)) \\ &= \beta(t)\left[\beta^2(t)p_0(t) + \widetilde{\beta}^2(t)p_0(-t)\right] \\ &= \beta(t)\left[\beta^2(t) + \widetilde{\beta}^2(t)\right]p_0(t) \\ &= \beta(t)p_0(t) = P^+f(t), \end{aligned}$$

where we have used the facts that p_0 is even and β satisfies (3.2.6).

We show that P^+ is self-adjoint by showing $\langle P^+f, g\rangle = \langle f, P^+g\rangle$. We get

$$\begin{aligned} \langle P^+f, g\rangle &= \int_{-\infty}^{\infty} \beta(t)\left[\beta(t)f(t) + \widetilde{\beta}(t)f(-t)\right]\overline{g(t)}\,dt \\ &= \int_{-1}^{1} \beta^2(t)f(t)\overline{g(t)}\,dt + \int_{-1}^{1} \beta(t)\widetilde{\beta}(t)f(-t)\overline{g(t)}\,dt \\ &\quad + \int_{1}^{\infty} f(t)\overline{g(t)}\,dt. \end{aligned} \tag{3.2.9}$$

Recalling that $\widetilde{\beta}(t) = \beta(-t)$, in the second integral we make the change of variables, $t' = -t$, to get

$$\begin{aligned} \int_{-1}^{1} \beta(t)\beta(-t)f(-t)\overline{g(t)}\,dt &= -\int_{1}^{-1} \beta(-t')\beta(t')f(t')\overline{g(-t')}\,dt' \\ &= \int_{-1}^{1} \beta(t')\beta(-t')f(t')\overline{g(-t')}\,dt'. \end{aligned}$$

This is enough to show that in all three integrals in (3.2.9), f and g can be interchanged; therefore $\langle P^+f, g\rangle = \langle f, P^+g\rangle$. The same results apply to P^- and therefore both P^+ and P^- satisfy Definition 3.2.1.

Now we show that P^+ and P^- are mutually orthogonal projections. This is true if and only if $\langle P^+ f, P^- f\rangle = 0$. Since $\operatorname{supp} P^+ f \cap \operatorname{supp} P^- f \subset [-1,1]$, we need only show that

$$\int_{-1}^{+1} \beta(t)\widetilde{\beta}(t)p(t)\overline{q(t)}\,dt = 0.$$

This is true because $\beta(t)\widetilde{\beta}(t) = \beta(t)\beta(-t)$ and $p(t)$ are even and $q(t)$ is odd.

Finally, we show that $P^+ + P^- = I$, meaning that $P^+ f + P^- f = f$. On the parts where the supports of $P^+ f$ and $P^- f$ do not overlap, if $t > 1$, $P^+ f = f$ and if $t < -1$, $P^- f = f$, so we need only consider $t \in [-1,1]$. In that case,

$$\begin{aligned}
P^+ f(t) + P^- f(t) &= \beta(t)\left[\beta(t)f(t) + \widetilde{\beta}(t)f(-t)\right] + \widetilde{\beta}(t)\left[\widetilde{\beta}(t)f(t) - \beta(t)f(-t)\right] \\
&= \beta^2(t)f(t) + \beta(t)\widetilde{\beta}(t)f(-t) + \widetilde{\beta}^2(t)f(t) - \widetilde{\beta}(t)\beta(t)f(-t) \\
&= \left[\beta^2(t) + \widetilde{\beta}^2(t)\right] f(t) \\
&= f(t),
\end{aligned}$$

because β satisfies (3.2.6).

□

We can generalize both Definition 3.2.3 and the projections P^+ and P^- to any interval $(a-\eta, a+\eta)$, $a \in \mathbb{R}$, $\eta > 0$ as follows.

Definition 3.2.6 *An increasing profile function β associated to the interval $(a - \eta, a+\eta)$, $a \in \mathbb{R}$, $\eta > 0$ will denote a smooth monotone increasing function β such that*

$$\begin{aligned}
\beta(t) &= 0, & t &< a-\eta, \\
\beta^2(t) + \beta^2(2a-t) &= 1, & a-\eta \le t &\le a+\eta, \\
\beta(t) &= 1, & t &> a+\eta.
\end{aligned} \tag{3.2.10}$$

The associated decreasing profile function, $\widetilde{\beta}$, such that $\widetilde{\beta}(t) = \beta(2a-t)$, is the reflection of β about the center of the interval, a.

Given an increasing profile function β associated to the interval $(a-\eta, a+\eta)$, the projections $P^+_{a,\eta}$ and $P^-_{a,\eta}$ from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$ are then defined by

$$\begin{aligned} P^+_{a,\eta}f(t) &= \beta(t)\left[\beta(t)f(t)+\widetilde{\beta}(t)f(2a-t)\right] \\ &:= \beta(t)p_a(t) \end{aligned} \tag{3.2.11}$$

and

$$\begin{aligned} P^-_{a,\eta}f(t) &= \widetilde{\beta}(t)\left[\widetilde{\beta}(t)f(t)-\beta(t)f(2a-t)\right] \\ &:= \widetilde{\beta}(t)q_a(t), \end{aligned} \tag{3.2.12}$$

where p_a and q_a are functions which are even and odd with respect to a given by the second terms in square brackets on the right hand sides of (3.2.11) and (3.2.12), respectively.

As was the case for P^+ and P^-, the projections $P^+_{a,\eta}$ and $P^-_{a,\eta}$ are mutually orthogonal and $P^+_{a,\eta}+P^-_{a,\eta}=I$.

Let $W^+_{a,\eta}=P^+_{a,\eta}(L^2(\mathbb{R}))$ (resp. $W^-_{a,\eta}=P^-_{a,\eta}(L^2(\mathbb{R}))$) denote the image of $P^+_{a,\eta}$ (resp. $P^-_{a,\eta}$) [1, p. 356]. For $f\in L^2(\mathbb{R})$, one easily checks that

$$P^+_{a,\eta}f(t)=\begin{cases} 0, & t<a-\eta, \\ \beta(t)p_a(t), & a-\eta\le t\le a+\eta, \\ f(t), & t>a+\eta \end{cases}$$

and

$$P^-_{a,\eta}f(t)=\begin{cases} f(t), & t<a-\eta, \\ \widetilde{\beta}(t)q_a(t), & a-\eta\le t\le a+\eta, \\ 0, & t>a+\eta, \end{cases}$$

where $p_a(t)=p_a(2a-t)$ and $q_a(t)=-q_a(2a-t)$ are functions that are even and odd with respect to a.

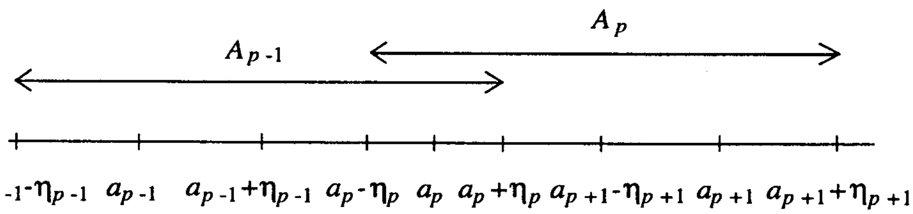

Figure 3: Two overlapping intervals.

3.3 Bell Functions

To introduce the notion of Malvar wavelets, let $\{a_p\}_{p\in\mathbb{Z}}$, be a strictly increasing sequence of real numbers such that

$$\lim_{p\to-\infty} a_p = -\infty \quad \text{and} \quad \lim_{p\to\infty} a_p = \infty.$$

For each $p \in \mathbb{Z}$, let I_p denote the interval $[a_p, a_{p+1})$ and $l_p = a_{p+1} - a_p$ its length. The sequence $\{I_p\}_{p\in\mathbb{Z}}$ forms a partition of $\mathbb{R}$. Let also $\{\eta_p\}_{p\in\mathbb{Z}}$ denote a sequence of strictly positive real numbers such that

$$\eta_p + \eta_{p+1} \leq l_p. \tag{3.3.13}$$

This condition ensures that $a_p + \eta_p \leq a_{p+1} - \eta_{p+1}$.

The sequence $\{A_p\}_{p\in\mathbb{Z}}$ of intervals $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$ has then the following properties:

$$A_p \cap A_q = \emptyset \quad \text{if} \quad |p-q| \geq 2 \quad \text{and} \quad A_{p-1} \cap A_p = [a_p - \eta_p, a_p + \eta_p),$$

meaning they form a set of nearly-disjoint intervals, as shown in Figure 3.

Definition 3.3.7 *Let $a_1 < b_1 < a_2 < b_2$ be four real numbers. A bell function g associated to (a_1, b_1, a_2, b_2) is defined by $g(t) = \min\{g_1(t), g_2(t)\}$, where g_1 (resp. g_2) is an increasing (resp. decreasing) profile function associated to (a_1, b_1) (resp. (a_2, b_2)).*

The three distinct portions of a bell function, $[a_1, b_1)$, $[b_1, a_2)$ and $[a_2, b_2)$, shown in Figure 4 will represent the attack, sustain and decay portions of a musical note. We say that the bell function associated to $(a_p - \eta_p, a_p + \eta_p, a_{p+1} - \eta_{p+1}, a_{p+1} + \eta_{p+1})$ is centered on the interval $I_p = [a_p, a_{p+1})$.

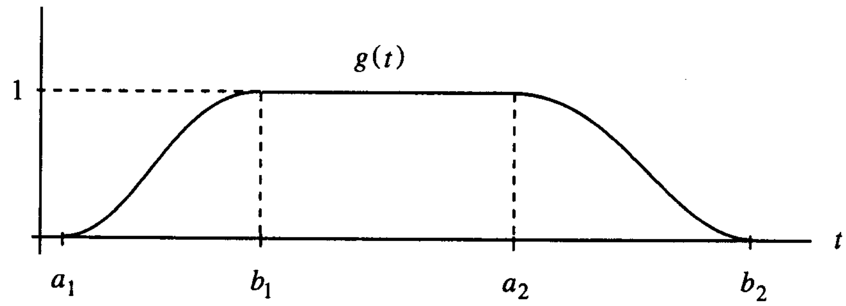

Figure 4: A bell function g.

Definition 3.3.8 *Let $\{A_p\}_{p\in\mathbb{Z}}$ be a sequence of nearly-disjoint intervals as above. For each $p \in \mathbb{Z}$, let g_p be a bell function associated to $(a_p - \eta_p, a_p + \eta_p, a_{p+1} - \eta_{p+1}, a_{p+1} + \eta_{p+1})$. A sequence $\{g_p\}_{p\in\mathbb{Z}}$ satisfying the condition*

$$g_{p-1}(a_p + \tau) = g_p(a_p - \tau), \quad \textit{for } |\tau| \leq \eta_p, \quad p \in \mathbb{Z}, \tag{3.3.14}$$

will be called a bell sequence associated to $\{A_p\}_{p\in\mathbb{Z}}$.

Let $\{\beta_p\}_{p\in\mathbb{Z}}$ and $\{\widetilde{\beta}_p(t) = \beta_p(2a_p - t)\}_{p\in\mathbb{Z}}$ denote sequences of monotone increasing and decreasing profile functions associated to the intervals $\{(a_p - \eta_p, a_p + \eta_p)\}_{p\in\mathbb{Z}}$. We remark that $\beta_p^2(t) + \widetilde{\beta}_p^2(t) = 1$, for all $p \in \mathbb{Z}$. An example of a bell sequence associated to the sequence of intervals $\{A_p\}_{p\in\mathbb{Z}}$ is then given by

$$g_p(t) = \begin{cases} \beta_p(t), & a_p - \eta_p \leq t \leq a_p + \eta_p, \\ 1, & a_p + \eta_p < t < a_{p+1} - \eta_{p+1}, \\ \widetilde{\beta}_{p+1}(t), & a_{p+1} - \eta_{p+1} \leq t \leq a_{p+1} + \eta_{p+1}, \\ 0, & t < a_p - \eta_p \text{ or } t > a_{p+1} + \eta_{p+1}. \end{cases} \tag{3.3.15}$$

3.4 Malvar Wavelets

Let $\{g_p\}_{p\in\mathbb{Z}}$ be a bell sequence as in Definition 3.3.8 and as shown in Figure 5 associated to a sequence of nearly disjoint intervals $\{A_p\}_{p\in\mathbb{Z}}$. The bell functions $\{g_p\}$ are used as the envelopes for Malvar wavelets. Recall that $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$ and $l_p = a_{p+1} - a_p$.

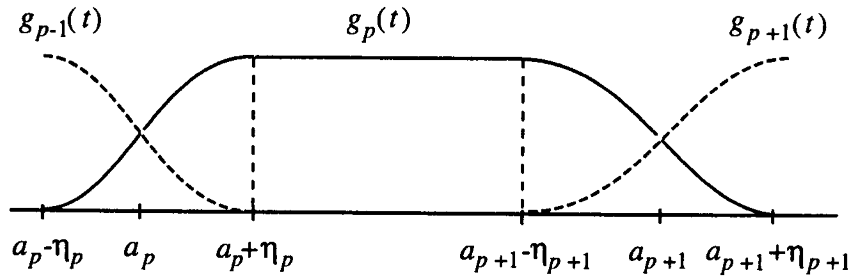

Figure 5: Bell sequence.

Definition 3.4.9 a) *For* $p \in \mathbb{Z}$ *and* $k \in \mathbb{Z}_+$ $(= \{0,1,2,\ldots\})$, *let* $u_{p,k} \in L^2(\mathbb{R})$ *be given by*

$$u_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]. \tag{3.4.16}$$

b) *For* $p \in 2\mathbb{Z}$, *let* $v_{p,k} \in L^2(\mathbb{R})$ *be given by*

$$v_{p,0}(t) = \sqrt{\frac{1}{l_p}}\, g_p(t) \quad \textit{and} \quad v_{p,k}(t) = \sqrt{\frac{2}{l_p}}\, g_p(t)\cos\frac{k\pi}{l_p}(t-a_p), \quad \textit{for } k \geq 1, \tag{3.4.17}$$

and, for $p \in 2\mathbb{Z}+1$,

$$v_{p,k}(t) = \sqrt{\frac{2}{l_p}}\, g_p(t)\sin\frac{k\pi}{l_p}(t-a_p), \quad \textit{for } k \geq 1. \tag{3.4.18}$$

Proposition 3.4.10 *The functions* $\{u_{p,k};\ p \in \mathbb{Z}, k \in \mathbb{Z}_+\}$ *form an orthonormal basis of* $L^2(\mathbb{R})$.

Proof. We first show that the functions are orthonormal. It suffices to show that

$$\langle u_{p-1,k}, u_{p,m}\rangle = 0, \quad \forall k, m \in \mathbb{Z}_+,$$

and

$$\langle u_{p,k}, u_{p,m}\rangle = \delta(k-m), \qquad \text{where } \delta \text{ denotes the Kronecker symbol,}$$

since $|i-j|>1 \implies \operatorname{supp}(u_{i,k}) \cap \operatorname{supp}(u_{j,m}) = \emptyset \implies \langle u_{i,k}, u_{j,m}\rangle = 0$. First, $\langle u_{p-1,k}, u_{p,m}\rangle =$

$$\int_{-\infty}^{\infty} \frac{2}{l_{p-1}l_p} g_{p-1}(t)g_p(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_{p-1}}{l_{p-1}}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt$$
$$=\frac{2}{l_{p-1}l_p}\int_{a_p-\eta_p}^{a_p+\eta_p} g_{p-1}(t)g_p(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_{p-1}}{l_{p-1}}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt.$$

Now, by Definition 3.3.8 of the bell sequence, using condition (3.3.14), we have $\langle u_{p-1,k}, u_{p,m}\rangle =$

$$\frac{2}{l_{p-1}l_p}\int_{a_p-\eta_p}^{a_p+\eta_p} g_p(2a-t)g_p(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_{p-1}}{l_{p-1}}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt$$
$$=0,$$

since $g_p(2a_p-t)g_p(t)$ is even with respect to a_p, $\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_{p-1}}{l_{p-1}}\right]$ is odd with respect to a_p and $\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]$ is even with respect to a_p. Now, $\langle u_{p,k}, u_{p,m}\rangle =$

$$\int_{-\infty}^{\infty} \frac{2}{l_p} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt$$
$$=\frac{2}{l_p}\int_{a_p-\eta_p}^{a_{p+1}+\eta_{p+1}} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt$$
$$=\frac{2}{l_p}\int_{a_p-\eta_p}^{a_p+\eta_p} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt$$
$$+\frac{2}{l_p}\int_{a_p+\eta_p}^{a_{p+1}-\eta_{p+1}} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt$$
$$+\frac{2}{l_p}\int_{a_{p+1}-\eta_{p+1}}^{a_{p+1}+\eta_{p+1}} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt$$
$$:=\mathrm{I}+\mathrm{II}+\mathrm{III}.$$

In I, both $\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]$ and $\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]$ are even with respect to a_p. So, for I we get

$$\begin{aligned}
&\frac{2}{l_p}\int_{a_p-\eta_p}^{a_p+\eta_p} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt\\
=&\frac{2}{l_p}\int_{a_p}^{a_p+\eta_p}\left[g_p^2(t)+g_p^2(2a_p-t)\right]\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt,\\
=&\frac{2}{l_p}\int_{a_p}^{a_p+\eta_p}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt,
\end{aligned}$$

because the bell functions g_p satisfy (3.2.10) by definition. Since $\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]$ and $\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]$ are odd with respect to a_{p+1}, their product is even. Similarly, for III we get

$$\begin{aligned}
&\int_{a_{p+1}-\eta_{p+1}}^{a_{p+1}+\eta_{p+1}} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt\\
=&\int_{a_{p+1}-\eta_{p+1}}^{a_{p+1}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt.
\end{aligned}$$

For II, we remark that $g_p(t)=1$ for $t\in[a_p+\eta_p, a_{p+1}-\eta_{p+1}]$. Therefore putting all three integrals together, we have $\langle u_{p,k},u_{p,m}\rangle =$

$$\begin{aligned}
&\frac{2}{l_p}\int_{a_p-\eta_p}^{a_{p+1}+\eta_{p+1}} g_p^2(t)\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt\\
=&\frac{2}{l_p}\int_{a_p}^{a_{p+1}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\cos\left[\pi\left(m+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]dt\\
=&\delta(k-m).
\end{aligned}$$

In order to show that the functions $\{u_{p,k};\ p\in\mathbb{Z}, k\in\mathbb{Z}_+\}$ generate the space $L^2(\mathbb{R})$, we first require the following result from [1].

Define the subspace $W^p\subset L^2(\mathbb{R})$ by

$$W^p = W^+_{a_p,\eta_p}\cap W^-_{a_{p+1},\eta_{p+1}}.$$

Then, $\bigoplus_{p\in\mathbb{Z}} W^p = L^2(\mathbb{R})$.

Thus it suffices to show that for $p\in\mathbb{Z}$, p fixed, the functions $\{u_{p,k};\ k\in\mathbb{Z}_+\}$ form a basis of W^p. The functions $\left\{\sqrt{2}\cos\left[\left(k+\frac{1}{2}\right)\pi t\right]\right\}_{k\in\mathbb{Z}_+}$ form an orthonormal basis of $L^2([0,1])$ and this implies that

$$\left\{\sqrt{\frac{2}{l_p}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\ ;\ k\in\mathbb{Z}_+\right\}$$

is an orthonormal basis of $L^2([a_p, a_{p+1}])$.

Let $f \in W^p$, then $f(t) = h(t)g_p(t)$, where h is a function that is even (resp. odd) with respect to a_p (resp. a_{p+1}) and $g_p(t)$ is, by definition, the minimum of two monotone increasing (decreasing) profile functions. The restriction of h to the interval $[a_p, a_{p+1}]$ can be decomposed into elements of this basis. Furthermore, this decomposition remains valid for all $t \in [a_p-\eta_p, a_{p+1}+\eta_{p+1}]$ since $h(t)$ and $\sqrt{\frac{2}{l_p}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]$ have the same symmetry with respect to a_p and a_{p+1} [1]. Therefore, for p fixed, $f \in W^p$ can be decomposed into elements of the family $\{u_{p,k};\ k \in \mathbb{Z}_+\}$.

$\square$

Remark.

i) This result is proved in [1, p. 361] where there is one profile function associated with each bell function, g_p, $p \in \mathbb{Z}$. This means that the attack and decay portions are all the same. However this result is independent of the function β_p chosen for each $p \in \mathbb{Z}$. This allows us to have different shapes for the attacks and decays of the bell functions as long as condition (3.3.14) is satisfied: that is on the portion where two adjacent bell functions overlap, their profiles be reflections of one another.

ii) The functions $\{v_{p,k}; p \in 2\mathbb{Z}, k \in \mathbb{Z}_+\} \cup \{v_{p,k}; p \in 2\mathbb{Z}+1, k \in \mathbb{N}\}$ also form an orthonormal basis of $L^2(\mathbb{R})$. This result is proved in [13, p. 106].

For the discrete case, we restrict our attention to $t \in \mathbb{Z}$. Letting h be the sampling rate, we need to substitute the discrete grid $h\mathbb{Z}$ for the real line, $\mathbb{R}$. The signal f is given as a sequence denoted $f(hk)$, $k \in \mathbb{Z}$ and we may omit h in all that follows [13, p. 116]. For the strictly increasing sequence of real numbers $\{a_p\}_{p\in\mathbb{Z}}$ and the sequence of positive real numbers $\{\eta_p\}_{p\in\mathbb{Z}}$, we require that a_1 be a half-integer and that $a_p \notin \mathbb{Z}$ to form a partition $\{I_p \cap \mathbb{Z}\}_{p\in\mathbb{Z}}$ of $\mathbb{Z}$, where $\{I_p\}_{p\in\mathbb{Z}}$ is a partition of $\mathbb{R}$ with $I_p = [a_p, a_{p+1})$. We recall again that $\eta_p + \eta_{p+1} \leq l_p$, where by l_p we denote the number of points in $[a_p, a_{p+1}] \cap \mathbb{Z}$ [13, p. 116]. Let $\{g_p\}_{p\in\mathbb{Z}}$ be a bell sequence, as in the continuous case, associated to the sequence of nearly disjoint intervals $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$. Now, the functions

$$\left\{ u_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}} \cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\right\}, \quad p \in \mathbb{Z},\ k = 0,1,2,\dots l_p - 1, \tag{3.4.19}$$

form an orthonormal basis of $l^2(\mathbb{Z})$ [1, p. 366] and [13, p. 117].

We can replace $l^2(\mathbb{Z})$ and consider $l^2\{1,2,\dots,N\}$ instead. Let $a_0 = \frac{1}{2}$, and $a_{j_0+1} = N + \frac{1}{2}$. The only conditions needed are that $g_0(t) = 1$ on $[\frac{1}{2}, a_1 - \eta_1]$ and that $g_{j_0}(t) = 1$ on $[a_{j_0+1} - \eta_{j_0+1}, a_{j_0+1}]$ with no additional constraints on those intervals [13].

3.5 Harmonic Amplitude Envelopes

In this section we describe the method for constructing N distinct harmonic amplitude envelopes. We begin with the following musical definition.

Definition 3.5.11 *A harmonic amplitude envelope for harmonic i in the synthesis of a musical instrument playing a note with fundamental frequency F is the envelope function that is used to define the amplitude over time of frequency iF (the i^{th} harmonic).*

We require that each harmonic have its own amplitude envelope because of what has been referred to in the literature as 'non-synchrony' of the harmonics [23, p. 223]: the higher numbered harmonics (≥ 2) appear later in the attack portion of a note and disappear sooner during the decay period than the fundamental. Similar to the exercise in Section 1.3 of recreating the amplitudes of harmonics using a combination of terms in a sine series to recreate the sound of an instrument, called additive synthesis, we need to recreate the amplitudes of harmonics, b_i, $i = 1, 2, \dots, N$, throughout the three portions of a musical note: the attack, sustain and decay. In this sense, we synthesize the sound of an instrument with a superposition of Malvar wavelets in a way that harmonic frequencies get added together to produce a sound with one fundamental frequency. Synthesized instruments sound more realistic when the harmonic amplitude changes are each accounted for separately [22]. This model is chosen in order to present Super Malvar wavelets defined in Section 3.6 that mimic the timbre of a musical instrument.

We introduce the following notation. For a function f, we denote by $f^{[0]}$ the identity and by $f^{[n]}$ the composition of f with itself n times.

Let $\{A_p\}_{p\in\mathbb{Z}}$, with $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$ be a sequence of nearly disjoint intervals as in Section 3.3. Let us introduce the following notation, which we will use in Definition 3.5.12. Let $\mathbf{m} = [m_1, m_2, \dots, m_N]$ and $\mathbf{n} = [n_1, n_2, \dots, n_N]$ denote vectors with entries in $\mathbb{N}$ and let $\mathbf{c} = [c_1, c_2, \dots, c_N]$ and $\mathbf{d} = [d_1, d_2, \dots, d_N]$ be vectors with non-negative real entries such that

$$\max_i\{c_i\} + \max_i\{d_i\} \leq \inf_p\{a_{p+1} - a_p - \eta_p - \eta_{p+1}\}. \tag{3.5.20}$$

Moreover, for $p \in \mathbb{Z}$, let $\eta_p^{(i)}$ be defined by

$$\eta_p^{(i)} = \eta_p + c_i, \quad \text{if } p \text{ odd}, \quad \text{and} \quad \eta_p^{(i)} = \eta_p + d_i, \quad \text{if } p \text{ even}. \tag{3.5.21}$$

Definition 3.5.12 *Let* $\mathbf{m}$, $\mathbf{n}$, $\mathbf{c}$ *and* $\mathbf{d}$ *be given as above. For each* $p \in \mathbb{Z}$, *let* $B_p^{(i)}$ *denote the function given by*

$$B_p^{(i)}(t) = \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[m_i]}\left(\frac{\pi}{2}\left(\frac{t-(a_p-\eta_p^{(i)})}{2\eta_p^{(i)}}\right)\right)\right), \quad \textit{if } p \textit{ odd}$$

and

$$B_p^{(i)}(t) = \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n_i]}\left(\frac{\pi}{2}\left(\frac{t-(a_p-\eta_p^{(i)})}{2\eta_p^{(i)}}\right)\right)\right), \quad \textit{if } p \textit{ even.}$$

Remark.

i) The choice of $\mathbf{m}$ and $\mathbf{n}$ will describe the suddenness of the attacks and decays of each harmonic, *i.e.* the greater the number of compositions of $\frac{\pi}{2}\sin^2$, the more sudden the attack or decay.

ii) The choice of $\mathbf{c}$ and $\mathbf{d}$ will determine the non-synchrony of the attacks and decays of the harmonics. This is shown for the attack portion in Figure 6.

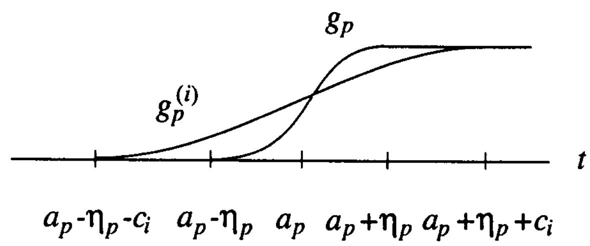

Figure 6: Attack portion of a harmonic amplitude envelope with non-asynchrony.

iii) Although a greater non-synchrony in (3.5.21), p odd, implies that the beginnings of the attacks come sooner which is seemingly the opposite of the desired effect, this is not considered a drawback since the endings of the attacks come later. The amplitudes of the higher numbered harmonics are often much less than the fundamental and this means that it would be difficult to hear the added-on portions resulting from a large non-synchrony for a higher harmonic. We have chosen this model for the desired symmetry of all the envelopes occurring at the exact same points, a_p.

We give an explicit definition for bell sequences that will be used in the definition of Super Malvar wavelets in Section 3.6. We denote by $\widetilde{B}_p^{(i)}(t) = B_p^{(i)}(2a_p - t)$, the reflection of $B_p^{(i)}(t)$ about a_p. For each $p \in \mathbb{Z}$, let $g_p^{(i)}$ be given by

$$g_p^{(i)}(t) = \begin{cases} B_p^{(i)}(t), & a_p - \eta_p^{(i)} \leq t \leq a_p + \eta_p^{(i)}, \\ 1, & a_p + \eta_p^{(i)} < t < a_{p+1} - \eta_{p+1}^{(i)}, \\ \widetilde{B}_{p+1}^{(i)}(t), & a_{p+1} - \eta_{p+1}^{(i)} \leq t \leq a_{p+1} + \eta_{p+1}^{(i)}, \\ 0, & t < a_p - \eta_p^{(i)} \text{ or } t > a_{p+1} + \eta_{p+1}^{(i)}. \end{cases} \tag{3.5.22}$$

Proposition 3.5.13 *The functions* $\{g_p^{(i)}\}_{p\in\mathbb{Z}}$ *form bell sequences associated to* $\{A_p^{(i)}\}_{p\in\mathbb{Z}}$, *with* $A_p^{(i)} = [a_p - \eta_p^{(i)}, a_{p+1} + \eta_{p+1}^{(i)})$, *for each* $i = 1, 2, \ldots, N$.

Proof.

It suffices to show that $[B_p^{(i)}(t)]^2 + [\widetilde{B}_p^{(i)}(t)]^2 = 1$ for $t \in [a_p - \eta_p^{(i)}, a_p + \eta_p^{(i)}]$ and that the functions $g_p^{(i)}$ are continuous. We consider any one of the sequences so we omit the superscript (i) and simply write B_p, $\widetilde{B}_p$ and g_p. To verify that B_p satisfies (3.2.10), we will use repeatedly the following trigonometric identity:

$$\sin\left(\frac{\pi}{2}\,(1-t)\right)=\cos\left(\frac{\pi}{2}\,t\right).$$

We evaluate

$$\begin{aligned}\widetilde{B}_p(t)=B_p(2a_p-t) &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\left(\frac{2a_p-t-(a_p-\eta_p)}{2\eta_p}\right)\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\left(\frac{a_p-t+\eta_p}{2\eta_p}\right)\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\left(\frac{a_p-t+2\eta_p-\eta_p}{2\eta_p}\right)\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\left(1-\frac{t-(a_p-\eta_p)}{2\eta_p}\right)\right)\right).\end{aligned}$$

With the change of variables $\dfrac{t-(a_p-\eta_p)}{2\eta_p}=\tilde{t}$, for $0\le\tilde{t}\le 1$, we have

$$a_p-\eta_p\le t\le a_p+\eta_p.$$

It follows that

$$B_p(t)=\sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\,\tilde{t}\right)\right)$$

and

$$\begin{aligned}\widetilde{B}_p(t) &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\,(1-\tilde{t})\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n-1]}\left(\frac{\pi}{2}\sin^2\left(\frac{\pi}{2}\,(1-\tilde{t})\right)\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n-1]}\left(\frac{\pi}{2}\,\cos^2\left(\frac{\pi}{2}\,\tilde{t}\right)\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n-1]}\left(\frac{\pi}{2}\left(1-\sin^2\left(\frac{\pi}{2}\,\tilde{t}\right)\right)\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n-2]}\left(\frac{\pi}{2}\,\sin^2\left(\frac{\pi}{2}\left(1-\sin^2\left(\frac{\pi}{2}\,\tilde{t}\right)\right)\right)\right)\right)\\ &= \sin\left(\left[\frac{\pi}{2}\sin^2\right]^{[n-2]}\left(\frac{\pi}{2}\,\cos^2\left(\frac{\pi}{2}\,\sin^2\left(\frac{\pi}{2}\,\tilde{t}\right)\right)\right)\right)\\ &\;\;\vdots\\ &= \cos\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\,\tilde{t}\right)\right).\end{aligned}$$

So,

$$B_p^2(t) + \widetilde{B}_p^2(t) = \sin^2\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\tilde{t}\right)\right) + \cos^2\left(\left[\frac{\pi}{2}\sin^2\right]^{[n]}\left(\frac{\pi}{2}\tilde{t}\right)\right) = 1,$$

for $a_p - \eta_p \leq t \leq a_p + \eta_p$. It is clear that when $t = a_p - \eta_p$,

$$B_p(t) = \sin 0 = 0$$

and, when $t = a_p + \eta_p$,

$$B_p(t) = \sin\frac{\pi}{2} = 1,$$

which shows that g_p is continuous.

□

In the following section, we create Super Malvar wavelets (SMW's) which consist of superpositions of Malvar wavelets. Rather than performing a Fourier analysis on the cut-up pieces of the signal corresponding to the intervals $[a_p, a_{p+1})$, we analyze the frequency in sets matching the harmonic overtone series from music theory localized to these intervals.

3.6 Super Malvar Wavelets

In this paper, we introduce new terminology, Super Malvar wavelets (SMW's), a superposition of N Malvar wavelets, for N less than 30, with frequencies following the harmonic overtone series discussed in Section 2.1. There is a physical meaning for this construction. The N Malvar wavelets represent N harmonics in a type of 'additive synthesis' and the coefficients in the superposition are harmonic amplitudes in this synthesis. Due to the importance of this harmonic structure throughout this thesis, we have given this a special name. The term 'Super' is added to 'Malvar wavelets' since it is a superposition of Malvar wavelets; also the word 'wavelets' is left in the definition because mathematically wavelets are small 'waves' in a very general sense and SMW's are still small waves. The sonic interpretation of these wavelets becomes the concept of timbre (pitch with a colour) whereas, for Malvar wavelets, it was simply the concept of a specific frequency (pitch).

For our construction of Super Malvar wavelets (SMW's), we will require that the individual Malvar wavelets in the superposition have frequencies that match with the harmonic overtone series. Given bell sequences $g_p^{(i)}$ as in (3.5.22), we define the function $\Omega_{p,k} \in L^2(\mathbb{R})$ for $p \in \mathbb{Z}$ and $k \in \mathbb{N}$ by

$$\Omega_{p,k}(t) = \frac{1}{\sqrt{\|\mathbf{b}\|^2}} \sum_{i=1}^{N} b_i v_{p,ik}^{(i)}(t), \tag{3.6.23}$$

where $\mathbf{b} = [b_1, b_2, \ldots, b_N]$ is a vector with non-negative real entries with $b_1 \neq 0$ and where

$$v_{p,k}^{(i)}(t) = \begin{cases} \sqrt{\frac{2}{l_p}}\, g_p^{(i)}(t) \cos \frac{k\pi}{l_p}(t - a_p), & p \text{ even}, \\ \sqrt{\frac{2}{l_p}}\, g_p^{(i)}(t) \sin \frac{k\pi}{l_p}(t - a_p), & p \text{ odd}. \end{cases}$$

In the case where p is odd,

$$\begin{aligned} \Omega_{p,k}(t) = & \sqrt{\frac{2}{l_p\|b\|^2}} \left[b_1 g_p^{(1)}(t) \sin \frac{k\pi}{l_p}(t - a_p) + b_2 g_p^{(2)}(t) \sin \frac{2k\pi}{l_p}(t - a_p) \right. \\ & \left. + \cdots + b_N g_p^{(N)}(t) \sin \frac{Nk\pi}{l_p}(t - a_p) \right], \quad k = 1, 2, \ldots. \end{aligned} \tag{3.6.24}$$

We recognize the form of (1.3.2) where the amplitudes of the harmonics, $b_1, b_2, \ldots, b_N$, change over time by way of the harmonic amplitude envelopes, $g_p^{(1)}(t), g_p^{(2)}(t), \ldots, g_p^{(N)}(t)$. Here, we can consider these harmonics as all belonging to the same emission of sound as each window is centered about the same interval, $[a_p, a_{p+1})$. In (3.6.24), $\frac{k\pi}{2\pi l_p} = \frac{k}{2l_p}$ is the fundamental frequency and

$$\frac{2k}{2l_p}, \frac{3k}{2l_p}, \ldots, \frac{Nk}{2l_p}$$

are the frequencies of the harmonics. The family of functions $\{\Omega_{p,k}(t)\}_{p\in\mathbb{Z},k\in\mathbb{N}}$ are Super Malvar wavelets (SMW's). A Super Malvar wavelet possesses a distinct timbre (to the human ear) relative to the instrument it models. In this thesis, we construct a different Super Malvar wavelet family for each different instrument we wish to consider in our instrument identification experiment explained in Chapter 4 in order to form a sample set (library of SMW's).

For Malvar wavelets, we had an orthonormal basis of $L^2(\mathbb{R})$. This is due to the fact that the individual Malvar wavelets in the linear combination each have distinct frequencies. In the case of Super Malvar wavelets, we do not have orthogonality in general. The Super Malvar wavelets, $\Omega_{p,k}$, $k = 0, 1, 2, \ldots$, do not each possess distinct frequencies. For the method proposed in this thesis for instrument identification, we have not used the fact that some pairs of Super Malvar wavelets have distinct frequencies and can be shown to have a near-zero inner product because of this. However, in Section 5.3, we show that two different Malvar wavelets with differing amplitude envelopes are nearly orthogonal as they are bases we will call highly coherent bases.

3.7 Backwards Characterization

This section describes a method which considers attack and decay features of the instruments considered in Chapter 4. We call this method 'backwards characterization'. We first remark that any signal can be reversed and, in the case where the signal consists of the musical recording of one note, then the new signal consists of a note with the same fundamental frequency but with an attack which used to be the decay and a decay which used to be the attack. We then wish to test whether the backwards signal corresponds to the same SMW instrument from the library but with the attack and decay characteristics ($\mathbf{n}$, $\mathbf{m}$, $\mathbf{c}$ and $\mathbf{d}$) switched. When this is true, then the attack and decay characteristics of that SMW are well suited or similar to the attack and decay characteristics of the note in the signal.

To see how this remark can be used to help identify the instrument recorded in a test signal, F, in which each note is separated by some silence, we first require some terminology. We suppose the signal contains some number of notes, say k. We characterize the signal by two sets of numbers, $\{a_p\}$ and $\{\eta_p\}$, $p = 1, 2, \ldots, 2k$, where $a_p - \eta_p$ and $a_{p+1} + \eta_{p+1}$, k odd, denote the beginning and end of note $\frac{p+1}{2}$, respectively, and where $2\eta_p$ and $2\eta_{p+1}$ are the durations of the attacks and decays of that note. This method is flexible for characterizing a recorded signal of many types of instruments.

For the backwards characterization method, we reverse the numbers η_p and the signal as well as inverting the partition by the following algorithms:

$$F[i+1] \longleftarrow F[L-i], \qquad i = 0, 1, \ldots, L-1,$$

$$\{a_{p+1}\} \longleftarrow L - \{a_{2k-p}\},$$

and

$$\{\eta_{p+1}\} \longleftarrow \{\eta_{2k-p}\}, \qquad p = 0, 1, \ldots, 2k-1,$$

where L is the length of the signal, and k is the number of notes in the signal.

What is of interest here is that, in the backwards characterization, when the chosen Super Malvar wavelet corresponds to the instrument in the recording, switching the attack and decay characteristics should produce a closer match. This backward characterization is shown in Figures 7 and 8 where we called the envelopes with reversed characteristics the 'ghost' envelopes, shown as dashed lines. In Figure 7, the 'ghost' envelopes replace in Figure 8 the regular envelopes, accomplished by shifting all windows by one p value, conforming to the notes in the signal. For the backward characterization method, we look at four different analyses of the signal (performed by taking inner products of the signal and the SMW): forwards/regular, forwards/ghost, backwards/ghost and backwards/regular (F/R, F/G, B/G and B/R). When the inner products coming from the F/R and B/G cases are higher in absolute value than those of the F/G and B/R cases, respectively, then we can conclude that the attack and decay characteristics are the correct characteristics for the notes in the signal. This, in itself, is not sufficient to identify the recorded instrument, but will be of help in the process when used in conjunction with the method described in Section 4.5.

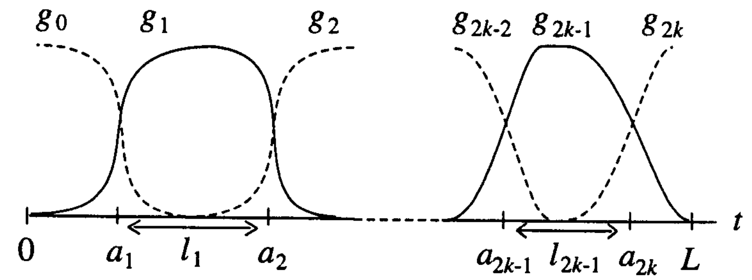

Figure 7: Bell functions for the characterization of a signal.

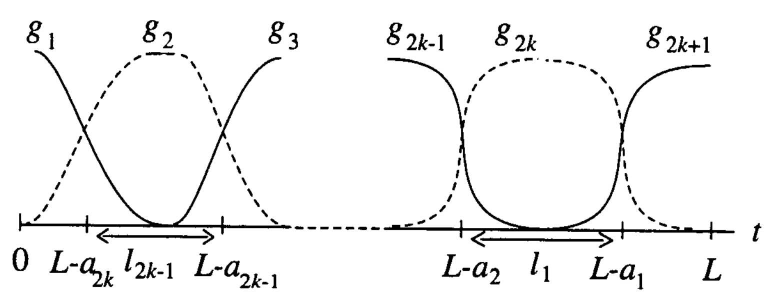

Figure 8: Bell functions for the characterization of the reversed signal.

Chapter 4

Instrument Identification Using SMW's

We would like to develop methods for instrument identification by computer which take into account attack and decay periods of notes. We believe that for developing a more reliable identification method, the computer needs to 'think like a human'. Part of the motivation is the use of wavelets in itself. Wavelet properties are adapted to musical signals.

One fact to be considered is that such a method should be capable of working with signals with several notes played by the same instrument. When a human hears some instrument, it is easier to identify the instrument if there are several notes played. We would like the method to reflect this.

4.1 Tests of Instrument Recordings

The experiment consisted of five separate tests. For each test, there was a recording of one of four instruments: the clarinet, the oboe, the trumpet and the violin. Figures 15–19 in Appendix C, which analyze the fundamental frequency content of the five recordings, were created using a MATLAB program, 'GUI–STRAIGHT', written by Hideki Kawahara [26].

The first recording is of four synthesizer instruments (clarinet, oboe, trumpet and violin) each playing the three notes, C3, G3 and E4, consecutively. The next two recordings for Tests 2 and 3 are of a real oboe playing the two notes, B♭3 and B3, mezzo-forte and then the twelve notes that are separated by semi-tones in between and including C4 and B4, fortissimo, respectively. The last two recordings for Tests 4 and 5 are of a real violin playing the three notes A4, B♭4 and B4, bowed on the A string mezzo-forte and then fortissimo, respectively. The figures for each of the five tests appear just after the test result data tables for that test in Appendix C.

4.2 Hypothesis

It is hypothesized that with the four different Super Malvar wavelets modelling the four musical instruments, from recordings where one instrument plays several notes with some silence between the notes, the computer can identify which of the four instruments is playing by analyzing the inner products of the super Malvar wavelets with the music signals in a way that takes into account various attack and decay features. We hypothesize that the first of two methods that searches only for the highest inner-product will produce above 25% accuracy (better than chance) and the second method that uses 'backwards characterization' will produce a significantly higher accuracy than when compared to the first method.

4.3 Purpose

The internet is a resourceful search engine for many different fields of interest, in particular that of music. For example, there is a musical tune library that allows a user to input a tune and the engine will search for similar tunes and provide their names, called the Greenstone Digital Library[1] of music. Sometimes but not always, a tune can be associated to a specific instrument that would usually play it. It would be possible to narrow a search by specifying the instrument in the input, but such classifications could be done automatically by a computer identification system. The

[1]New Zealand Digital Library Project at the University of Waikato: http://www.sadl.uleth.ca/nz/cgi-bin/library

experiment is designed to be helpful toward the development of such a system.

We believe that Super Malvar wavelets will have many further applications. In the field of music analysis, particularly jazz, we would like to analyze the intricate mix of different instruments and, for the aid of this analysis, computers may be employed for many different purposes.

For one, the quality of older jazz recordings can be very poor and computer algorithms can be used to increase the listenability. A second application is music transcription: the time it takes for a human to write out the various parts in a music performance is enormous and very tedious but may be done by computer.

Another concern is the jumbling effect that is inherent but not unique to improvisational music: when one part crosses another part or when too much is happening at one time. In an effort to sort out these types of situations, a computer algorithm could be used to help keep track of the lines of each of the instruments. In this regard, a jazz music analyst would be able to keep concentrating on the things that a computer would very likely not be able to do which is provide a musical analysis of the improvisations.

Also, in the field of speech recognition, Super Malvar wavelets could be of use.

4.4 Harmonic Grouping for SMW's

We design here a method for our SMW constructions to group the harmonics numbered 1 to N depending on their positions in the Western tonal system. Namely, when two harmonics are the same note modulo one octave, they belong to the same group. The natural numbers are placed in an array shown in Table 5. The array has the definition $A = \{a_{n,k}; n \geq 1, k \geq 0\}$, where the n^{th} row is given by the following: if l denotes the non-negative integer such that $2^l \leq 2n-1 < 2^{l+1}$, then

$$a_{n,j} = \begin{cases} 0 & \text{if } 0 \leq j \leq l-1, \\ (2n-1)2^{j-l} & \text{if } j \geq l. \end{cases}$$

In general, for any n,

$$l = \left\lfloor \log_2 \frac{2n-1}{2} \right\rfloor.$$

Table 5: Harmonic groupings array by musical note, modulo one octave.

1	2	4	8	16	$\cdots$
	3	6	12	24	$\cdots$
		5	10	20	$\cdots$
		7	14	28	$\cdots$
			9	18	$\cdots$
			11	22	$\cdots$
			13	26	$\cdots$
			15	30	$\cdots$
				17	$\cdots$
				$\vdots$	$\ddots$

The harmonic groupings array groups harmonics by their musical notes. Harmonics are grouped by row: each row corresponds to one musical note, modulo one octave. Given the row n, the number l indicates which octave harmonic $2n-1$, the first non-zero entry in the row, is in with respect to the fundamental (zeros in the array are left empty). For example, for row five, $l = 3$, meaning that the ninth harmonic is in the third octave above the fundamental. As a result of this, harmonics are also organized into columns: the harmonics with numbers in column k are in the k^{th} octave above the fundamental. For example, $a_{1,3}$, which corresponds to the eighth harmonic, is in the third octave above the fundamental.

Lemma 4.4.1 *For $1 \leq n \leq \lfloor \frac{N+1}{2} \rfloor$, let $q_n = \log_2 \frac{N}{2n-1}$ and $I_n = \{(2n-1)2^m; 0 \leq m \leq q_n\}$. Then $\{I_n; 1 \leq n \leq \lfloor \frac{N+1}{2} \rfloor\}$ forms a partition of $\{1, \ldots, N\}$.*

Proof.

We show that the lemma holds with the following:

1. For all $p \in I_n$, $1 \leq n \leq \lfloor \frac{N+1}{2} \rfloor$, $p \leq N$.
2. For any natural number $k \leq N$, we can write $k = (2n-1)2^m$ where n and m are such that $m \leq q_n = \log_2 \frac{N}{2n-1}$.

From the definition of q_n, for $p \in I_n$,

$$\begin{aligned} p =& (2n-1)2^m \\ \leq& (2n-1)2^{\log_2\left(\frac{N}{2n-1}\right)} \\ =& (2n-1)\left(\frac{N}{2n-1}\right) = N. \end{aligned}$$

We show the second by contrapositive. Let $k \leq N$ be such that $k = (2n-1)2^m$ where n and m do not satisfy the condition. Then

$$\begin{aligned} m &> \log_2 \frac{N}{2n-1} \\ 2^m &> \frac{N}{2n-1} \\ (2n-1)2^m &> N, \end{aligned}$$

but this says that $k > N$, which produces a contradiction.

$\square$

Commonly, a note played by any given instrument will have six or seven harmonics audible to the human ear [12, p.2]. For the construction of the SMW's, we first require definitions of the harmonic amplitude envelopes depending on the instrument we wish to model. In our models we consider twelve harmonics and after harmonic grouping by the harmonic groupings array we get six sets containing the numbers less than 13 in the first six rows of Table 5. They are

$$\{1,2,4,8\}, \{3,6,12\}, \{5,10\}, \{7\}, \{9\}, \{11\}.$$

For each instrument, we define six harmonic amplitude envelopes $g^{(1)}, g^{(2)}, \ldots, g^{(6)}$ as in (3.5.22) in Section 3.5 by way of $\mathbf{m} = [m_1, \ldots, m_6]$, $\mathbf{n} = [n_1, \ldots, n_6]$, $\mathbf{c} = [c_1, \ldots, c_6]$ and $\mathbf{d} = [d_1, \ldots, d_6]$ subject to conditions described in Section 3.5. All the Malvar wavelets in the SMW superposition $\Omega_{p,k} = \sum_{i=1}^{12} b_i u_{p,ik}^{(i)}$ with harmonic i falling within group n will share the harmonic amplitude envelope $g^{(n)}$, $n = 1, 2, \ldots, 6$. Simplifying the number of harmonic amplitude envelope calculations this way will cause less ill-effects on the synthesis of the instrument's timbre than other simplifications might:

we could measure the non-synchronies only for odd numbered harmonics and because of the choice for the groupings, we still arrive at an approximation of the instrument's timbre since notes separated by octaves sound alike.

Remark.

i) Since the harmonic amplitude envelope $g^{(n)}$ is used for more than one harmonic, namely all the harmonics in group n, in particular it may not portray the non-synchrony and attack/decay profile characteristics of all the harmonics in the group. For this reason, we model envelope $g^{(n)}$'s characteristics after the harmonic in the group with the largest amplitude. For example, if $\mathbf{b} = [50\ 40\ 15\ 34\ 18\ 20\ 14\ 12\ 3\ 6\ 6\ 10]$ is the vector with non-negative real entries for the harmonic amplitudes of a given instrument's tone, then $b_3 = 15$, $b_6 = 20$ and $b_{12} = 10$ and then $g^{(2)}$ would be modelled after the characteristics portrayed by the sixth harmonic rather than the third.

ii) We modelled twelve harmonics since in [20], [21] and [22], twelve is the number of harmonics analyzed in the case of the trumpet, whereas 20 for the violin and 21 for the clarinet and oboe.

4.5 Method

The method was, for each instrument, using the diagrams and tables from [20], [21] and [22] to estimate the values of the following: $b = [b_1, b_2, \ldots, b_{12}]$, such that $b_i \geq 0$, $i = 1, 2, \ldots 12$, and $\|b\| \neq 0$, the amplitudes of the first twelve harmonics on an arbitrary scale during the sustain periods, $\mathbf{c} = [c_1, c_2, \ldots, c_6]$ and $\mathbf{d} = [d_1, d_2, \ldots, d_6]$, measuring the attack and decay non-synchrony in seconds as well as $\mathbf{m} = [m_1, m_2, \ldots, m_6]$ and $\mathbf{n} = [n_1, n_2, \ldots, n_6]$ which characterize the suddenness of the harmonic amplitude envelopes' attacks and decays, respectively. In Table 6, we give these values for the four instruments in the experiment.

Here, we found that by introducing the $\log_2$ factor for all the harmonic amplitudes of the trumpet and violin produced more audible resemblance to the trumpet and violin to the ear.

Table 6: Parameter measurements by instrument.

Clarinet

$$\begin{aligned}
\mathbf{b} &= [700\ 25\ 220\ 28\ 460\ 28\ 105\ 50\ 150\ 33\ 55\ 45] \\
\mathbf{c} &= [0\ 0\ 0\ 0\ 0\ 0] \\
\mathbf{d} &= [0\ 0\ 0\ 0.01\ 0.02\ 0.03] \\
\mathbf{m} &= [1\ 1\ 2\ 1\ 1\ 1] \\
\mathbf{n} &= [1\ 2\ 2\ 2\ 2\ 2]
\end{aligned}$$

Oboe

$$\begin{aligned}
\mathbf{b} &= [430\ 310\ 480\ 460\ 575\ 110\ 35\ 14\ 65\ 295\ 170\ 110] \\
\mathbf{c} &= [0\ 0\ 0\ 0\ 0\ 0] \\
\mathbf{d} &= [0\ 0.03\ 0.05\ 0\ 0.06\ 0.05] \\
\mathbf{m} &= [1\ 1\ 1\ 2\ 1\ 1] \\
\mathbf{n} &= [1\ 2\ 2\ 2\ 2\ 2]
\end{aligned}$$

Trumpet

$$\begin{aligned}
\mathbf{b} &= \log_2([420\ 440\ 450\ 300\ 160\ 120\ 100\ 90\ 80\ 35\ 35\ 12]) \\
\mathbf{c} &= [0\ 0.005\ 0.005\ 0.02\ 0.015\ 0.04] \\
\mathbf{d} &= [0\ 0.01\ 0\ 0\ 0\ 0] \\
\mathbf{m} &= [1\ 1\ 2\ 2\ 2\ 1] \\
\mathbf{n} &= [1\ 1\ 1\ 1\ 1\ 1]
\end{aligned}$$

Violin

$$\begin{aligned}
\mathbf{b} &= \log_2([80\ 180\ 70\ 350\ 250\ 50\ 110\ 140\ 30\ 25\ 60\ 40]) \\
\mathbf{c} &= [0\ 0\ 0.015\ 0\ 0\ 0] \\
\mathbf{d} &= [0\ 0.09\ 0.09\ 0.155\ 0.13\ 0.16] \\
\mathbf{m} &= [1\ 1\ 1\ 1\ 1\ 1] \\
\mathbf{n} &= [1\ 1\ 1\ 1\ 1\ 1]
\end{aligned}$$

4.6 Procedure

All the signals considered in this thesis will have one instrument playing one note at a time. Given one signal to be analyzed, we need to first decide on a partition of a segment of the positive time axis into intervals $[a_p, a_{p+1})$ and the values $\{\eta_p\}$ describing the half-lengths of the attacks and decays of the notes in the signal. Generally, we will use a new partition for each new signal to be analyzed. In Chapter 3, we considered an infinite bell sequence and infinite bell sequences for Malvar wavelets and Super Malvar wavelets, beginning with the sequence of nearly disjoint intervals $\{A_p\}_{p\in\mathbb{Z}}$, recall $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$, for $\{a_p\}$ and $\{\eta_p\}$ appropriately given as in Section 3.3. However, because all signals for processing will be finite, we need neither consider the negative $p = 0, -1, \ldots$ part of the time axis intervals, nor the part for $p \geq J$, where J is some positive integer to be determined based on the partition and the signal being analyzed.

In this chapter, we know more facts about the test signals to be analyzed than the fact that they have one note played at a time: we know when each note was played and for how long and the letter name each note. From Figures 15–19 in Appendix C, the durations and time of occurrences of all attacks, sustains and decays for every note was observed. For each test signal, the sets $\{a_p\}_{p=1,\ldots,J}$ and $\{\eta\}_{p=1,\ldots,J}$, J even with $J/2$ being the number of notes in the signal, were created such that $t = a_p$ was the time directly in the center of the attack portion of note $(p+1)/2$, p odd, or the time directly in the center of the decay portion of note $p/2$, p even, and η_p would equal half of the attack or decay durations, p odd or even, respectively. Furthermore, information about each of the note's fundamental frequency was noted and each note was assigned a frequency range which contained five adjacent frequencies in Hz in integers. All of these measurements were taken using a software tool and have been included in Tables 15–16 in Appendix C. Remark that for Test 1, the sample rate was $h = \frac{1}{11025}$ second, while for each of the other tests, $h = \frac{1}{44100}$ second.

We have four different Super Malvar wavelets, one for each instrument in the sample set, to perform analyses on the signals by calculating inner products. In Tables 17–21 in Appendix C, each of the entries in the arrays show the absolute values of the inner products of the corresponding Super Malvar wavelets, $\Omega_{p,k}$, with the signal in question. The columns determine p and k for the specific notes in the signal and frequencies within the frequency range for that note. The rows determine several factors: first, which Super Malvar wavelet (clarinet, oboe, trumpet or violin) was used for the analysis; and second, whether the analysis was performed with the forwards or backwards signal and whether the regular or ghost windows were used in the analysis: forwards/regular, forwards/ghost, backwards/ghost and backwards/regular (F/R, F/G, B/G and B/R, respectively). The results for Tests 1 and 3 both required a multiple number of pages for the data.

4.7 Results

We used two methods to pick one instrument from the sample set for the likely candidate as playing a note in the tests. Each decision is made on a note to note basis. For note $\frac{p+1}{2}$, method one makes its choice by choosing the instrument with the largest inner product, $\langle \Omega_{p,k}, S \rangle$, in any of the four F/R, F/G, B/G and B/R analyses with k within the five Hz frequency range given for that particular note, out of the four sample set instruments. The significance of picking this instrument is that, because the SMW's are normalized by the weights of the harmonic amplitudes, the relative amplitudes of the SMW in question were a closer match for that note than were those of the other three SMW models. This is the main component described by Helmholtz as being responsible for the particular timbre of an instrument.

For our second method, a choice is contingent on backwards characterization. For a particular note, method two first ranks the four instruments by largest inner products in absolute value. Then, starting at the top of the list and moving downwards, it chooses the first instrument that satisfies the condition that, for the specific column corresponding to the largest inner product, the inner products corresponding to the F/R and the B/G cases are both greater than the F/G and the B/R cases, respectively. In case of a tie, the condition is not satisfied. This method does not make a selection in every case. If none of the instruments satisfy this condition, the note is left undecided.

This method considers specific attack and decay characteristics of the notes. Selecting the instrument at the top of the list, as for method one, does not necessarily consider these characteristics.

We have in Table 22, instrument identifications by two methods described in Section 4.6. In the *Ranking* columns, we have listed the four instruments by highest inner product in descending order for each note. Thus, the first instrument to appear in the ranking was identified by method 1 as having played the note in question. A × mark has been placed in case of an incorrect choice whereas a ✓ denotes a correct choice. In the *Winner* columns, there was an identification by the second method whenever there appears, in this order:

1. the four values of the inner products which correspond to those listed in Tables 17–21 for the successful SMW analyses by F/R, F/G, B/G and B/R (top to bottom),

2. the chosen instrument's name and

3. a ✓ or ×.

Nothing appears in the column when the note is undecided by method 2.

For method one, a correct identification occurred in nine of the 32 notes which is 28.125% of the time. However, for method two, a correct identification occurred in six of the 14 identified notes which is 42.857% of the time. We also have that if we considered only the F/R and F/G analyses, then method one produced eleven correct of 32 identifications or 34.375% notes correctly identified whereas the second method, which had to be modified slightly by making a choice of an instrument if the F/R case was higher than the F/G case, would have produced eight correct out of 25 identifications or 32% of the time. A final method, which considers only the F/R case and selects the instrument producing the highest inner product, gave eight correct identification in 32 or 25%.

Method two gains an advantage over method one where it decides, i n a way, that method one was an unreliable identification which lowers the number of errors. In fact, the method abandoned twelve misidentifications but only six correct choices. It changed five wrong choices to right choices, retained wrong answers only three times and switched from one wrong answer to another three times. Method two only changed a correct identification to an incorrect one twice. We deem these results supportive of the hypothesis.

4.8 Conclusions

Although seemingly poor reliability (constantly below 50%), when these results are compared to those in Table 1, where graduate music students were made to identify instruments from altered tones (attacks and decays removed) from a list of nine instruments, we see similar percentages of correct identifications.

It is interesting to note that in Test 1, where notes were played by a synthesizer, the results for method 1 gave only two out of twelve correct identifications whereas, from the experiments with synthesizers in Chapter 1, we concluded that it would be easier for a computer to tell synthesizer instruments apart. Perhaps the reason for the lower than expected accuracy for the **RolandXV5050** synth-module test (Test 1) is due to the fact that this time the SMWs (which made up the sample set) were modelling instruments from analyses of real instrument recordings in [20], [21] and [22].

One fact we noticed is that it seems that the violin and trumpet tend to have smaller inner products whereas the clarinet and oboe have the larger ones, especially the clarinet which accounted for 18 of the 32 choices in method 1. What could account for this flaw is the various harmonic amplitudes in the SMW's: the SMW's are normalized, meaning that we have divided each by their norm in their definition; although this was thought to give uniformity between the different instruments, but because each instrument has different harmonic amplitudes, this could influence the values of the inner products. In fact, the clarinet has the strongest fundamental with respect to the other harmonics of the four instruments. Twelve clarinet identifications by method 1 were abandoned by method 2, of which only two were in fact correct.

We can also note that both the clarinet and the oboe have strong attack profiles,

$\mathbf{n} = [122222]$. This leads us to conclude that the highest inner products may not provide an accurate way to tell which instrument is playing. However, for method 2, it seems that the trumpet is identified most often followed by the oboe, the violin and finally the clarinet. This indicates that the aforementioned bias seems to be counteracted, however, another bias may have been introduced. Perhaps the trumpet SMW attack and decay profiles are most like the majority of notes in the five tests.

One way this problem could be averted is, instead of just looking for the highest coefficient in the first method, we would use the concept of entropy, defined in Section 5.2, which gives us a way of measuring the efficiency of a representation. For future tests, we also hope to create a larger database of instruments and sample tones, as we had done in the **Technics KN6500 Experiment**.

Chapter 5

Future Directions

In this thesis, we attempted to see if a computer could tell the difference between four instruments by using SMW's. For the completion of this, many tasks were done manually. However, many ideas rely heavily on computer programming and, in this chapter, we present some of the ideas that are both original, specifically when they pertain to Super Malvar wavelets, and non-original, such as some pertaining simply to Malvar wavelets. We present these as possible future directions. They are meant to be a starting point for the development of an automated instrument recognition computer program which will use SMW's.

5.1 Merge Techniques

We begin by recalling some of the terms we will be working with in the following sections. We introduced bell sequences in Section 3.3, followed by Malvar wavelets in Section 3.4, both of which required first a sequence of nearly disjoint intervals $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$.

For the signal S being considered there are three known facts.

1. The length of the signal is L. Therefore, $\operatorname{supp} S \subset [0, L]$ is the segment of the time axis of concern.

2. There exists a real positive number $T > 0$ such that each note in the signal is

longer in duration than T.

3. All notes in the signal are separated by some silence.

Let l be the smallest integer with $[0,3^l) \supset [0,L]$. Knowing T, we pick a natural number q large enough to have $\frac{1}{3^q} < T$ and define the initial partition of $[0,3^l)$ into intervals of equal length $I_p = [a_p, a_{p+1}) = [p3^{-q}, (p+1)3^{-q})$, $p = 0, 1, \ldots, 3^{q+l} - 1$. In this way, $\operatorname{supp} S \subset \bigcup_p I_p = [0, 3^l)$. For each $p \geq 0$, let $\eta_p = \eta = \frac{1}{3} 3^{-q}$. With this, we can associate to the partition $\{I_p\}$ a bell sequence $\{g_p\}$ defined in Section 3.3 which would correspond with the nearly disjoint intervals $[a_p - \eta, a_{p+1} + \eta)$.

We allow a modification to be made to both the partition and the bell sequence by merging three adjacent intervals I_k, I_{k+1} and I_{k+2} into one new interval $\widetilde{I}_k = I_k \cup I_{k+1} \cup I_{k+2}$ and the corresponding merging of the bell functions obtained by replacing the three corresponding bell functions with one new bell function,

$$\widetilde{g}_k = \sqrt{g_k^2 + g_{k+1}^2 + g_{k+2}^2}. \tag{5.1.25}$$

In doing this, we obtain a new pair of sequences $\{\widehat{I}_p\}$ and $\{\widehat{g}_p\}$ such that

$$\widehat{I}_p \Longleftarrow \begin{cases} I_p, & \text{if } p < k, \\ \widetilde{I}_p, & \text{if } p = k, \\ I_{p+2}, & \text{if } p > k \end{cases} \tag{5.1.26}$$

and similarly for $\widehat{g}_p$.

In [13], there is a merge procedure which joins **two** adjacent bell functions. We explain the reason for this modification as it relates to our particular problem. Recall that for any signal we are considering, there is a rest between any two notes. This allows us to place the necessary 'ghost' windows which has the reversed attack and decay characteristics between the notes in the signal. The problem with merging **two** bell functions is that we have lost the desired attack and decay profiles. To see this, in Figure 9, A, D, $\overline{A}$ and $\overline{D}$ indicate the profile portions in terms of attack or decay or the reflections of either. The process of merging **three** bell functions g_p defined in (3.3.15) produces a new window that has retained the attack and decay profiles.

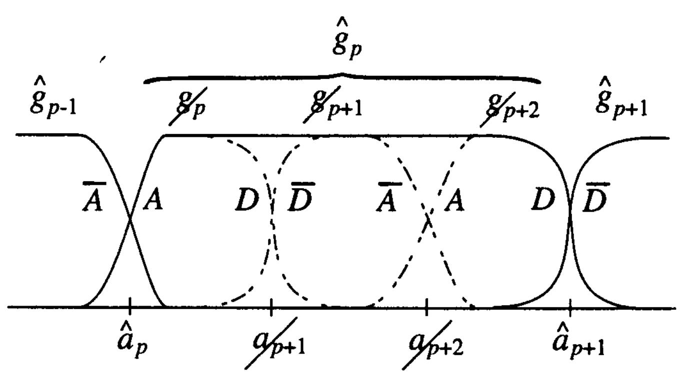

Figure 9: Merge procedure.

5.2 Entropy-Based Algorithm

In this section, we describe the method for determining when it is considered advantageous to perform a merging procedure and introduce the entropy-based algorithm for best basis selection [3]. We begin with the following definition of Shannon entropy which is the basis for the entropy-based algorithm.

Definition 5.2.1 *Let H be a hilbert space having representation by an orthogonal direct sum, $H = \oplus \sum H_i$. Letting $v \in H$, $\|v\| = 1$, for the decomposition $v = \oplus \sum_i v_i$ into its H_i components, we define Shannon entropy [3] given by*

$$\epsilon^2(v; \{H_i\}) = -\sum_i \|v_i\|^2 \log \|v_i\|^2,$$

a measure of distance between v and the orthogonal decomposition.

We consider sequences in $l^2(\mathbb{Z})$. We require the following definition.

Definition 5.2.2 *For $v \in H$, $\|v\| = 1$, $\epsilon^2(v; \{H_i\}) < +\infty$, the theoretical dimension of the sequence $\{\|v_i\|\}$ is defined by*

$$d = \exp\left(\epsilon^2(v; \{H_i\}\right).$$

We can take $l^2(\mathbb{Z})$ as H. Hereafter we put $H = l^2(\mathbb{Z})$.

In Section 5.1, we created an initial partition of $[0, 3^l)$ into intervals of the form $I_p = [p3^{-q}, (p+1)3^{-q})$, $p = 0, 1, \ldots 3^{q+l} - 1$ for some $q \in \mathbb{N}$. The goal is to modify the initial partition by a sequence of merging procedures to arrive at a new partition that is best suited to the signal—as we had manually set up for our experiments in Chapter 4. By a simple change of scale, we can arrive at the case where $q = 0$ and then the initial partition is the simple partition into intervals $I_p = [a_p, a_{p+1}) = [p, p+1)$, $p = 0, 1, \ldots 3^l - 1$. We require the following definition.

Definition 5.2.3 *An interval I is said to be triadic if it can be written in the form*

$$I_{n,m} = [3^m n, 3^m(n+1)), \quad \textit{for } m,\ n \in \mathbb{Z}_+,$$

and $I_{n,m}$, m, $n \geq 0$, denotes the triadic interval $[3^m n, 3^m(n+1))$.

From this definition, we give conditions for three intervals to be considered *mergeable.* Three adjacent intervals are said to be *mergeable* if and only if they are of the form $I_{3n,m}$, $I_{3n+1,m}$ and $I_{3n+2,m}$, n, $m \geq 0$. It can be checked that the resulting interval obtained by the merging of three mergeable intervals is of the form $I_{n,m+1}$.

Let $\mathcal{I}$ denote the set of all triadic intervals. Let u_k^I denote the Malvar wavelets $u_{p,k}$ in (3.4.16) where p indicates the interval I on which the bell function g_p is centered. The functions $\{u_k^I\}_{k \in \mathbb{Z}_+}$, $I \in \mathcal{I}$, forms an orthonormal basis of W_I [13], which we define as the subspace of $L^2(\mathbb{R})$–functions in the closure of the span of the family of Malvar wavelets, $\{u_k^I\}_{k \in \mathbb{Z}_+}$.

Remark.

i) For m fixed, the set of intervals $\{I_{n,m}\}$, $n = 0, 1, \ldots, 3^{l-m} - 1$, forms a partition of $[0, 3^l)$.

ii) The initial partition corresponds to the case where $m = 0$.

iii) For any collection of intervals $\{I_{n,m}\}$ that form a partition of $[0, 3^l)$, we have a corresponding orthonormal Malvar wavelet basis of $W_{[0,3^l)}$.

Definition 5.2.4 *A library of orthonormal bases is a (ternary) tree if it satisfies:*

1. *The basis vectors are centered on intervals of $\mathbb{N}$ of the form*

$$I_{n,m} = [3^m n, 3^m(n+1)), \quad \text{for } m, n \in \mathbb{Z}_+.$$

x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	S $\swarrow$						$\searrow$ **D**	
s_0	s_1	s_2	g_0	g_1	g_2	d_0	d_1	d_2
S $\swarrow$		$\searrow$ **D**	**S** $\swarrow$		$\searrow$ **D**	**S** $\swarrow$		$\searrow$ **D**
ss	gs	ds	sg	gg	dg	sd	gd	dd

Figure 10: Wavelet packets organized as a ternary tree.

2. *Each basis in the library corresponds to a disjoint cover of* $\mathbb{N}$ *by intervals* $I_{n,m}$.

3. *If* $W_{I_{n,m}}$ *is the subspace identified with the triadic interval* $I_{n,m}$, *then* $W_{I_{n,m+1}} = W_{I_{3n,m}} \oplus W_{I_{3n+1,m}} \oplus W_{I_{3n+2,m}}$.

In Figure 10, we show the ternary tree structure. The tree has the following structure: each level of the tree corresponds to a value of $m \geq 0$, $m = 0$ denoting the lowest level of the tree; referring to l from Section 5.1, the value $m = l$ represents the root (top) of the tree. The tree has the following interpretation: every leaf of the tree corresponds to a triadic interval $I \subset [0, 3^l)$; the parent nodes are the union of its three children nodes; every connected subtree corresponds to a partition of $[0, 3^l)$.

Given a signal S, for $I \in \mathcal{I}$, we define

$$c_k^I = \int_0^\infty S(t) u_k^I(t)\, dt,$$

and

$$\epsilon(I) = -\sum_{k=0}^{\infty} |c_k^I|^2 \log |c_k^I|^2, \tag{5.2.27}$$

and finally

$$\epsilon^*(I) = \inf \sum_p \epsilon(J_p), \tag{5.2.28}$$

where the infimum is taken over all partitions $\{J_p\}$ of the interval I into triadic intervals J_p belonging to $\mathcal{I}$ [13]. We describe the algorithm for best-basis selection.

The entropy-based algorithm consists of induction on the level of the (ternary) tree, $m \geq 0$, for finding the optimal partition of $[0, 3^l)$ starting with the most refined partition and merging intervals as described above. Starting with $m = 0$, we have only one possible basis, namely the basis corresponding to the initial partition. In this case, $\epsilon^*(I_{n,0}) = \epsilon(I_{n,0})$, $n = 0, 1, \ldots 3^l$, because the most refined partition available for these intervals are the intervals themselves. The induction hypothesis at level m is that we know:

a) $\epsilon^*(I_{n,m})$ for each $n = 0, 1, \ldots, 3^{l-m}$;

b) The partition $\{J_p\}$, $J_p \in \mathcal{I}$, for each interval $I_{n,m}$ at level m that gives the value of $\epsilon^*(I_{n,m})$.

To proceed to the next level, $m+1$, if

$$\epsilon(I_{n,m+1}) < \epsilon^*(I_{3n,m}) + \epsilon^*(I_{3n+1,m}) + \epsilon^*(I_{3n+2,m}), \tag{5.2.29}$$

then we retain $I_{n,m+1}$, and drop all prior information regarding $I_{3n,m}, I_{3n+1,m}$ and $I_{3n+2,m}$. The partition retained of interval $I_{n,m+1}$ is the coarsest, trivial, partition (composed of $I_{n,m+1}$ by itself) [13]. If

$$\epsilon(I_{n,m+1}) \geq \epsilon^*(I_{3n,m}) + \epsilon^*(I_{3n+1,m}) + \epsilon^*(I_{3n+2,m}), \tag{5.2.30}$$

then we let

$$\epsilon^*(I_{n,m+1}) = \epsilon^*(I_{3n,m}) + \epsilon^*(I_{3n+1,m}) + \epsilon^*(I_{3n+2,m}),$$

and the partition of $I_{n,m+1}$ is chosen as the respective partitions of $I_{3n,m}, I_{3n+1,m}$ and $I_{3n+2,m}$ that are known from the calculations of $\epsilon^*(I_{3n,m}), \epsilon^*(I_{3n+1,m})$ and $\epsilon^*(I_{3n+2,m})$ at the previous level. The algorithm terminates when a partition is found of the interval $L = [0, 3^l)$ [13].

5.3 Highly Coherent Bases

We develop new terminology which arises from terminology in [4], where bases are said to be 'highly incoherent'. This new terminology is used to describe some facts about Malvar wavelets with differing bell functions, $g_p^{(i)}$. We use this terminology in order to describe dependence from the elements of one basis on the elements of another basis for the same space.

We begin with the following theorem.

Theorem 5.3.5 *Let β_1 and β_2 be two strictly monotone increasing profile functions associated to the two intervals $[a-\eta_1, a+\eta_1]$ and $[a-\eta_2, a+\eta_2]$, respectively, as in Definition 3.2.6. Then*

$$\frac{1}{\sqrt{2}} \leq \beta_1(t)\beta_2(t) + \beta_1(2a-t)\beta_2(2a-t) \leq 1. \tag{5.3.31}$$

Proof. Without loss of generality, we may assume that $a = 0$ and $0 < \eta_1 \leq \eta_2$. It is enough to consider $t \geq a = 0$ since the middle terms in (5.3.31) are symmetric with respect to a (*i.e.* for $t = a+\tau$, $\beta_i(t) = \beta_i(a+\tau)$ and $\beta_i(2a-t) = \beta_i(a-\tau)$, and when $t = a-\tau$, $\beta_i(2a-t) = \beta_i(a+\tau)$ and $\beta_i(t) = \beta_i(a-\tau)$, $i = 1, 2$). Therefore we wish to show that

$$\frac{1}{\sqrt{2}} \leq \beta_1(t)\beta_2(t) + \beta_1(-t)\beta_2(-t) \leq 1.$$

Define,

$$\mathbf{u}(t) = \begin{bmatrix} \beta_1(t) \\ \beta_1(-t) \end{bmatrix} \qquad \text{and} \qquad \mathbf{v}(t) = \begin{bmatrix} \beta_2(t) \\ \beta_2(-t) \end{bmatrix}.$$

From property (3.2.10) of monotone increasing profile functions, we have that $\|\mathbf{u}\| = 1$ and $\|\mathbf{v}\| = 1$. Also from this property, we have that

$$\mathbf{u}(0) = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \mathbf{v}(0).$$

Let $\theta_1(t)$ and $\theta_2(t)$ denote the angles between $\mathbf{u}(t)$ and $\mathbf{v}(t)$ and the positive x-axis at time t, respectively. For $t \in [0, \eta_2]$, $\theta_1(t)$ and $\theta_2(t)$ are decreasing, such that

$$\frac{\pi}{4} \geq \theta_1(t) \geq 0 \qquad \text{and} \qquad \frac{\pi}{4} \geq \theta_2(t) \geq 0 \tag{5.3.32}$$

since β_1 and β_2 are smooth monotone increasing functions. Let $\theta(t)$ be the angle between $u(t)$ and $v(t)$ at time t. Then we have for $0 \leq t \leq \eta_2$,

$$\cos(\theta(t)) = \mathbf{u}(t) \cdot \mathbf{v}(t) = \beta_1(t)\beta_2(t) + \beta_1(-t)\beta_2(-t).$$

It suffices to that

$$\frac{1}{\sqrt{2}} \leq \cos(\theta(t)) \leq 1,$$

but due to the properties (5.3.32), we have that $-\frac{\pi}{4} \leq \theta(t) \leq \frac{\pi}{4}$ and so this completes the proof.

□

We introduce the following definition.

Definition 5.3.6 *Let $\{e_\alpha^{(1)}\}_{\alpha\in J}$ and $\{e_\beta^{(2)}\}_{\beta\in J}$ be two orthonormal bases of a Hilbert space. We call the two bases highly coherent when the following condition holds:*

There exists a bijection ρ of J, such that $\left|\left\langle e_\alpha^{(1)}, e_{\rho(\alpha)}^{(2)}\right\rangle\right| \geq \frac{1}{\sqrt{2}}$.

Remark. We say that the bases are 'nearly orthogonal' in the sense that, as a consequence resulting from the condition in Definition 5.3.6, if $\{e_\alpha^{(1)}\}_{\alpha\in J}$ and $\{e_\beta^{(2)}\}_{\beta\in J}$ are two highly coherent bases, then by Parseval's identity for any α, we have

$$\sum_{\beta\neq\rho(\alpha)} \left|\left\langle e_\alpha^{(1)}, e_\beta^{(2)}\right\rangle\right|^2 \leq \frac{1}{2}.$$

Let us recall the following well-known result needed for the proof of Theorem 5.3.8.

Proposition 5.3.7 *Let $g(x)$ be a bounded non-negative function on $[a,b]$ and $f(x)$ a function defined on $[a,b]$, such that*

$$m \leq f(x) \leq M, \quad x \in [a,b],$$

then

$$m\int_a^b g(x)\,dx \leq \int_a^b f(x)g(x)\,dx \leq M\int_a^b g(x)\,dx.$$

Then we have,

Theorem 5.3.8 *Let $a_1, \eta_1^{(1)}, a_2, \eta_2^{(1)}, \eta_1^{(2)}$ and $\eta_2^{(2)}$ be six real numbers such that $a_1 < a_2$, with $\eta_1^{(1)} + \eta_2^{(1)} \leq a_2 - a_1 = l$ and $\eta_1^{(2)} + \eta_2^{(2)} \leq l$ and let g_1 and g_2 be two bell functions associated to $(a_1-\eta_1^{(1)}, a_1+\eta_1^{(1)}, a_2-\eta_2^{(1)}, a_2+\eta_2^{(1)})$ and $(a_1-\eta_1^{(2)}, a_1+\eta_1^{(2)}, a_2-\eta_2^{(2)}, a_2+\eta_2^{(2)})$, respectively, as in Definition 3.3.7 (see Figure 11). As in (3.4.18), let $\{v_k\}$ and $\{w_k\}$ be two orthonormal Malvar bases of $W_{([a_1,a_2])}$ given, for $k \geq 1$, by*

$$v_k = \sqrt{\frac{2}{l}}\, g_1(t) \sin\frac{k\pi}{l}(t-a_1), \quad \text{and} \quad w_k = \sqrt{\frac{2}{l}}\, g_2(t) \sin\frac{k\pi}{l}(t-a_1).$$

Then these two bases, $\{v_k\}$ and $\{w_k\}$, are highly coherent.

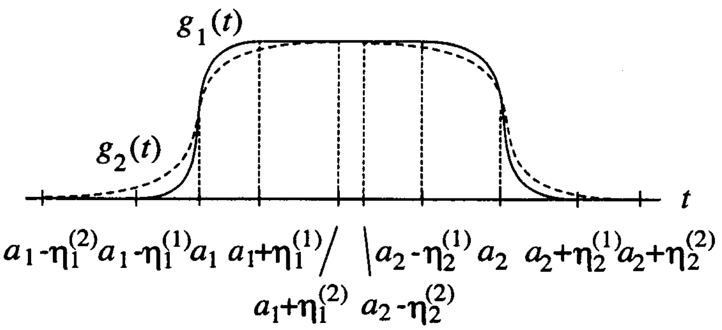

Figure 11: Two bell functions, g_1 and g_2.

Proof. To simplify notation, let $s_k(t)$ denote $\sin \frac{k\pi}{l}(t-a_1)$, $\eta_1 = \max\{\eta_1^{(1)}, \eta_1^{(2)}\}$ and $\eta_2 = \max\{\eta_2^{(1)}, \eta_2^{(2)}\}$. Now,

$$\begin{aligned}
\langle v_k, w_k \rangle &= \frac{2}{l}\int_{-\infty}^{\infty} g_1(t)g_2(t)s_k(t)s_k(t)\,dt \\
&= \frac{2}{l}\int_{a_1-\eta_1}^{a_2+\eta_2} g_1(t)g_2(t)s_k^2(t)\,dt \qquad (5.3.33)
\end{aligned}$$

Letting

$$h(t) = \begin{cases} [g_1(t)g_2(t) + g_1(2a_1-t)g_2(2a_1-t)] & \text{if } t \in [a_1, a_1+\eta_1], \\ 1 & \text{if } t \in [a_1+\eta_1, a_2-\eta_2], \\ [g_1(t)g_2(t) + g_1(2a_2-t)g_2(2a_2-t)] & \text{if } t \in [a_2-\eta_2, a_2], \end{cases}$$

we can rewrite (5.3.33) as in the proof of Proposition 3.4.10. We get

$$\langle v_k, w_k \rangle = \frac{2}{l}\int_{a_1}^{a_2} h(t)s_k^2(t)\,dt.$$

The function h has, by Theorem 5.3.5, $\frac{1}{\sqrt{2}} \leq h(t) \leq 1$, for $t \in [a_1, a_2]$, since g_1 and g_2 are defined as monotone increasing (decreasing) profile functions. Then by Proposition 5.3.7,

$$\frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}}\underbrace{\left(\frac{2}{l}\int_{a_1}^{a_2} s_k^2(t)\,dt\right)}_{=1} \leq \frac{2}{l}\int_{a_1}^{a_2} h(t)s_k^2(t)\,dt \leq \frac{2}{l}\int_{a_1}^{a_2} s_k^2(t)\,dt = 1,$$

since $s_k^2(t)$ is bounded and positive and $\frac{1}{\sqrt{2}} \leq h(t) \leq 1$ on $[a_1, a_2]$. This proves that

$$\langle v_k, w_k \rangle \geq \frac{1}{\sqrt{2}}, \quad \forall k \geq 1.$$

□

5.4 Ideal Atomic Decomposition

We begin with a collection of bases of the finite dimensional space $\mathbb{R}^N$, Φ_1, Φ_2, …, Φ_D, that are expressed as $N \times N$ matrices with the basis elements as column vectors.

Let a dictionary of the bases,

$$\Phi = \Phi_1 \bigcup \Phi_2 \bigcup \cdots \bigcup \Phi_D,$$

be the union of the bases Φ_d, $d = 1, \ldots, D$. This is expressed as an $N \times DN$ matrix. Given a signal S, we seek a representation in terms of atoms in the dictionary, by

$$S = \sum_{\gamma} \alpha_\gamma \phi_\gamma, \tag{5.4.34}$$

where $\gamma = (d, i)$ is an index into the dictionary, naming both the basis and the specific basis element [4].

Definition 5.4.9 *An overcomplete system in the space $\mathbb{R}^N$ is a set of elements of $\mathbb{R}^N$ containing at least one proper subset of vectors that forms a basis.*

Remark. In particular, any element can be expressed in more than one way using linear combinations of the elements in an overcomplete system.

Definition 5.4.10 *We say that a vector $\mathbf{x} \in \mathbb{R}^N$ is highly sparse if all but a few of its components are zero.*

Remark. This term is a necessary component for two optimization problems defined below. The number of components which are non-zero depends on the particular instances in each of these problems: the type of bases used and the particularities of the signal in question may quantitatively affect this notion.

> If Φ is an overcomplete system, any representation, $S = \sum_\gamma \alpha_\gamma \phi_\gamma$, is an *atomic decomposition* using *atoms from the dictionary.* If S in fact can be generated by a highly sparse sum, with the term "highly sparse" given an appropriate definition, and there is in fact only one such highly sparse way of doing so, and if an optimization principle finds that decomposition, we say that the principle leads to *ideal atomic decomposition* under the stated sparsity hypothesis [4].

The representation (5.4.34) is not unique. We want to find an optimal representation, which is highly sparse. We define two constrained optimization problems:

(P_0): Minimize $\|\alpha\|_0$, subject to $S = \sum_\gamma \alpha_\gamma \phi_\gamma$,
where $\|\alpha\|_0 = \#\{\gamma : \alpha_\gamma \neq 0\}$, giving the number of non-zero components, which is called the l^0 quasi-norm [4] and

(P_1): Minimize $\|\alpha\|_1$, subject to $S = \sum_\gamma \alpha_\gamma \phi_\gamma$,
where $\|\alpha\|_1 = \sum_\gamma |\alpha_\gamma|$ is the l^1-norm.

The first has the solution that is the sparsest of the set of feasible vectors (that satisfy the constraint) but which is difficult to solve with nonlinear optimization methods. The second is a convex optimization problem which has methods available for solving it but that finds the most sparse of the feasible vectors indirectly provided some facts can be verified. In general the solution to the second optimization problem is not highly sparse except in some situations for which we require the notion of incoherence of two orthogonal bases.

Definition 5.4.11 *Let Φ_1 and Φ_2 be two orthogonal bases of $\mathbb{R}^N$. We define the mutual incoherence* [4] *of Φ_1 and Φ_2 by*

$$M(\Phi_1, \Phi_2) = \sup\{|\langle \phi_1, \phi_2 \rangle| : \phi_1 \in \Phi_1, \phi_2 \in \Phi_2\}.$$

We have the following [4, Theorem VII.1]

Theorem 5.4.12 *Let Φ_1 and Φ_2 be orthonormal bases for $\mathbb{R}^N$ and let $M(\Phi_1, \Phi_2)$ be their mutual incoherence as in Definition 5.4.11. Let $\Phi = \Phi_1 \cup \Phi_2$ be the dictionary obtained by their union. Let $S = \Phi\alpha$, where α obeys*

$$\|\alpha\|_0 < \frac{1}{2}(1 + M^{-1})$$

then α is the unique solution to (P_1) and also the unique solution to (P_0).

The situation can be generalized to the case of an overcomplete system $\Phi = \Phi_1 \cup \Phi_2$ that consists of two non-orthogonal bases. Then we can proceed with the following [4]:

Definition 5.4.13 *Let Φ_1 and Φ_2 be two bases of $\mathbb{R}^N$. We define the mutual incoherence of Φ_1 and Φ_2 by*

$$\widetilde{M}(\Phi_1, \Phi_2) = \max \left\{ \max_{i,j} |\Phi_1^{-1}\Phi_2|_{i,j}, \max_{i,j} |\Phi_2^{-1}\Phi_1|_{i,j} \right\},$$

where Φ_1^{-1} is the inverse matrix of Φ_1, similarly for Φ_2^{-1}.

Remark. This definition is congruous to Definition 5.4.11 in the case that Φ_1 and Φ_2 are orthogonal bases [4]. We have the following [4, Theorem XII-1],

Theorem 5.4.14 *Let Φ_1 and Φ_2 be bases for $\mathbb{R}^L$ and let $\Phi = \Phi_1 \cup \Phi_2$ be the dictionary obtained by merging the two bases. Suppose that S can be represented as a superposition of N_1 atoms from Φ_1 and N_2 atoms from Φ_2. If*

$$N_1 + N_2 \leq \frac{1}{2}\widetilde{M}(\Phi_1, \Phi_2)^{-1},$$

then the solution to (P_1) is unique, the solution to (P_0) is unique and they are the same.

We can pose the problems for arbitrary collections of $L_i < N$ independent vectors Φ_i in $\mathbb{R}^N$ instead of bases. Here, Φ_i is an $N \times L_i$ matrix with full rank L_i. However, there are problems of existence of solutions and Theorem 5.4.14, which states when the solution to (P_1) is the solution we seek, may no longer apply. The question of existence can be addressed by incorporating at least one orthonormal basis Φ_d into the dictionary along with the collections of vectors in $\mathbb{R}^N$. We can associate the matrices Φ_i^+, pseudo-inverses of the matrices Φ_i. Then, we can calculate the mutual incoherence of the sets Φ_1 and Φ_2 of vectors in $\mathbb{R}^N$, by a modified Definition 5.4.13 simply using pseudo-inverses rather than inverse matrices. Recall that for a matrix $A \in \mathbb{R}^{m,n}$, $m > n$ with full rank n, the pseudo-inverse matrix $A^+ \in \mathbb{R}^{n,m}$ is given by $A^+ = (A^T A)^{-1} A^T$.

We propose to create a dictionary using Super Malvar wavelets to find the ideal atomic decomposition for a signal S. Due to their properties concerning timbre, we believe the resulting decomposition of the signal into time-fundamental frequency atoms will possess some desirable characteristics which are explained in the next section.

5.5 A Program for Part Separation

We introduce the concept for this section with the question: what can be done to quickly obtain a written score of a piece of music where, for some reason, only the recording of it is available? By a written score of the piece we require all distinct parts written out separately on their own line. The task is easy for recordings of one instrument playing one note at a time, either by automated means or by human ear and becomes difficult as the number of instruments as well as the number of notes each is allowed to play at one time increase, to the extreme cases of full symphony orchestra or big band. The time it might take to decipher the individual parts by ear could be very long for trained aural music transcribers.

Before we discuss this further, we review the main topics of this thesis. In Section 3.3, with a strictly increasing sequence $\{a_p\}_{p\in\mathbb{Z}}$ with $\lim_{p\to\pm\infty} a_p = \pm\infty$ and a sequence of small enough, strictly positive numbers $\{\eta_p\}$, we created a sequence of nearly disjoint intervals $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$. A sequence of bell functions $\{g_p\}$ was used to define Malvar wavelets $u_{p,k}$ and $v_{p,k}$, $p \in \mathbb{Z}$, $k \in \mathbb{Z}_+$, that both form orthonormal bases of $L^2(\mathbb{R})$.

Given a signal of finite length L, with music notes having durations at least $T > 3^{-q}$ for some q chosen large enough, that are separated by some silence, we considered in Section 5.2 a library of orthonormal Malvar wavelet bases of the space

$L^2([0,3^l))$, where l is such that $[0,L] \subset [0,3^l)$. By some merging procedures using an entropy-based algorithm for best basis selection, we were able to choose a partition of $[0,3^l)$ into intervals $[a_p, a_{p+1})$ that have integer lengths. With $\{\eta_p\}$, strictly positive real numbers small enough, we associated a sequence of nearly disjoint intervals $A_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1})$ to the partition.

In Section 3.5 we created N bell functions $g_p^{(1)}, g_p^{(2)}, \ldots, g_p^{(N)}$, each centered on the same interval $I_p = [a_p, a_{p+1})$, that model a particular musical instrument's harmonic amplitude envelopes. Then, N Malvar wavelets $v_{p,k}^{(i)}$, $i = 1, 2, \ldots, N$, were defined using these bell functions. Finally, in Section 3.6, Super Malvar wavelets were made as a weighted sum of the N Malvar wavelets such that their frequencies match the harmonic overtone series starting from some fundamental frequency described in Section 2.1, and have the sound of the particular instrument in question. We have called these wavelets time-fundamental frequency atoms.

Now, we wish to describe how to choose a dictionary $\Phi = \Phi_1 \cup \Phi_2 \cup \cdots \cup \Phi_D$ of Super Malvar wavelets and Malvar wavelet bases. We outline a program that uses ideal atomic decomposition for instrument part separation. For a given signal input into the program, a dictionary is formed by Super Malvar wavelets, $\Phi_d = \{\Omega_{p,k}^{(d)}\}$, that each represent one instrument in the signal. To do this, the user must already have knowledge of what the parts of the signal are. For example, the user should be able to tell that the signal is a duet between the violin and oboe.

Remark. If, instead of considering new dictionaries with every new signal analyzed, we wish to have one, fixed, dictionary, for a particular class of signals, it should be possible to create one. Let $\mathcal{S}$ be the class of all music signals for which the lengths are no longer than L and in which each note is played for at least a duration $T > 0$. It is possible to build a finite, static dictionary capable of analyzing these types of signals with the purpose of finding ideal atomic decompositions because:

1. There are finitely many types of instruments;

2. Coming from the initial partition in the merging procedure, there is a finite number of possible resulting partitions, albeit very large.

By finding the ideal atomic decomposition of the signal in terms of atoms from the dictionary,

$$S = \sum_{\gamma} \alpha_\gamma \phi_\gamma,$$

$\gamma = (d, i)$ being an index into the dictionary, we have found the distinct parts of the signal by their instruments: all coefficients $\alpha_\gamma \neq 0$ denote the presence of instument d in the signal. If all coefficients pertaining to one particular instrument are zero then that instrument is not present in the signal. Separating the parts is equivalent to sorting out the indices $\gamma = (d, i)$ by their values of d.

5.6 Conclusions

In this thesis, we considered two main methods for identifying instruments by their characteristic timbre. For this thesis we considered four musical instruments: the clarinet, oboe, violin and trumpet.

The first method considered the steady-state harmonic spectrum of a note, something that is known to affect the timbre or musical colour of an instrument. Using this method, in Chapter 1 and in Appendices A and B we conducted two experiments using synthesizers. Results were very successful for the **Yamaha** synthesizer instruments test but remained mainly unsuccessful for the **Technics KN6000** test.

A new approach was adopted based on observations made in 1975 in a research experiment performed on 57 graduate Music students [19]: it was found that the attacks and decays of the musical instruments played a large role in the identification of nine musical instruments.

To address this problem, we needed a Mathematical model for the musical instruments that considered both steady-state spectrum and the attack and decay portions of the note.

We defined Malvar wavelets and developed Super Malvar wavelets. Where Malvar wavelets had the physical meaning of a sound with a specific frequency, Super Malvar wavelets had a specific timbre as well.

Using specially designed SMW's which modelled the attack and decay characteristics of the four instruments as well as their steady-state spectrum, we formulated a method for 'instrument de-synthesis' a term referring to the inverse procedure of additive synthesis of a musical instrument's sound.

Wavelets are a useful, multipurpose tool in signal analysis. Now that we have created wavelets that physically represent the sound of musical instruments, they will be best suited to music signals created by specific musical instruments.

Different wavelets have differing properties, such as orthogonal or orthonormal. In the case of SMW's, we have modified Malvar Wavelets and in doing so we modified their orthogonal properties. This modification would render the wavelets useless as a tool for compression. However, not so for our proposed purpose of musical analysis: we do not attempt to recover the signal from its transform. It is for the purpose of recognizing the timbre of an instrument's sound. In this thesis, SMW's were shown to be versatile with possibilities for further study.

Appendix A

Yamaha Synthesizer Experiment

			E2	E2	E2	E2	E2	E2	E2	E2	E2	E2
			C3	C3	C3	C3	C3	C3	C3	C3	C3	C3
			Bas	Vio	Cel	Tpt	Trb	FHn	Obo	Cla	Flu	Pan
test	C2	Bas	0.980	0.454	0.386	0.533	0.891	0.256	0.729	0.735	0.086	0.205
			0.999	0.457	0.411	0.497	0.856	0.224	0.732	0.726	0.080	0.162
	D2	Vio	0.443	0.809	0.738	0.379	0.759	0.831	0.457	0.433	0.301	0.277
			0.472	0.861	0.805	0.376	0.774	0.776	0.478	0.437	0.284	0.264
	E2	Cel	0.382	0.902	0.994	0.943	0.690	0.894	0.367	0.629	0.696	0.696
			0.417	0.902	0.996	0.934	0.713	0.889	0.383	0.615	0.686	0.689
	F2	Tpt	0.528	0.974	0.897	0.998	0.811	0.935	0.525	0.783	0.753	0.767
			0.550	0.966	0.927	0.996	0.836	0.931	0.540	0.776	0.743	0.754
	G2	Trb	0.797	0.835	0.754	0.878	0.977	0.759	0.788	0.817	0.440	0.497
			0.796	0.848	0.798	0.852	0.988	0.723	0.801	0.824	0.428	0.464
	A3	FHn	0.217	0.965	0.841	0.916	0.596	0.999	0.337	0.593	0.799	0.750
			0.245	0.961	0.879	0.926	0.646	0.999	0.354	0.578	0.788	0.753
	B3	Obo	0.280	0.787	0.719	0.744	0.641	0.844	0.333	0.317	0.308	0.255
			0.315	0.843	0.786	0.718	0.663	0.791	0.354	0.317	0.291	0.250
	C3	Cla	0.744	0.712	0.550	0.751	0.793	0.594	0.821	0.997	0.691	0.781
			0.710	0.661	0.562	0.761	0.826	0.611	0.823	0.994	0.689	0.748
	D3	Flu	0.077	0.768	0.638	0.738	0.323	0.768	0.233	0.663	1.000	0.979
			0.080	0.704	0.632	0.784	0.384	0.824	0.239	0.635	1.000	0.987
	E3	Pan	0.143	0.742	0.618	0.732	0.350	0.728	0.297	0.717	0.995	0.994
			0.135	0.672	0.608	0.777	0.409	0.786	0.302	0.688	0.996	0.998

Appendix B

Technics KN6000 Experiment

Strings

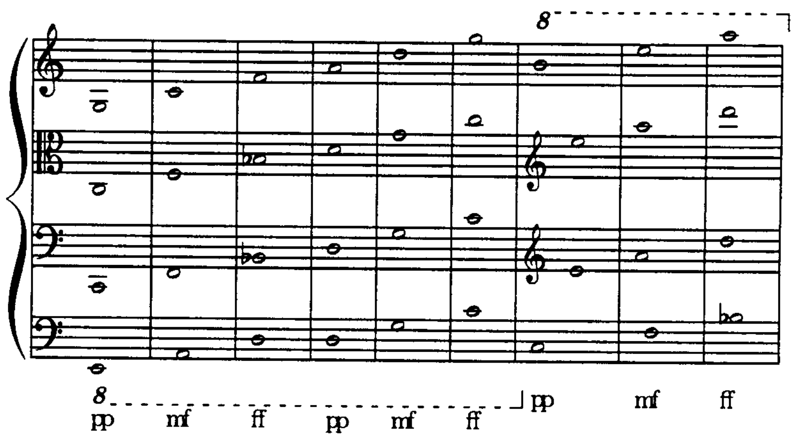

Figure 12: Database of stringed instruments (Vio, Vla, Vc and DB) in 9 different registers/dynamics.

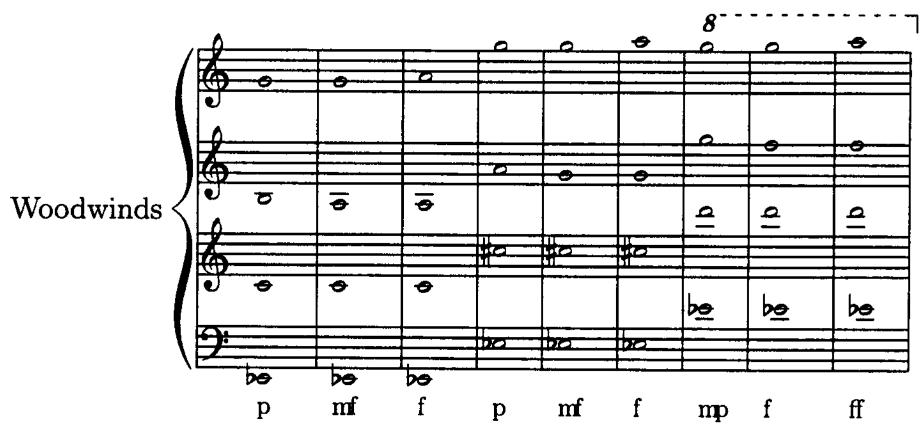

Figure 13: Database of woodwind instruments (Flu, Cla, Obo and Bsn) in 9 different registers/dynamics.

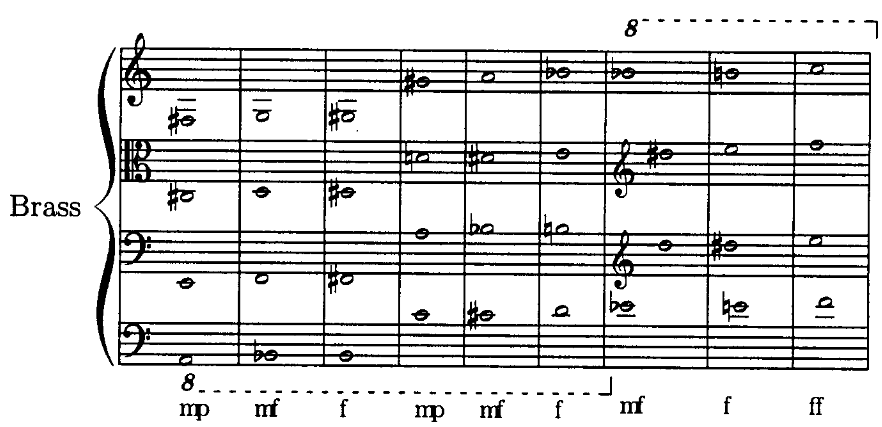

Figure 14: Database of brass instruments (Tpt, FHn, Trb and Tba) in 9 different registers/dynamics.

Table 7: Violin test tone.

reg/dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.747	0.734	0.873	0.815	0.930	0.720	0.653	0.716	0.640	0.810	0.484	0.712
low/mf	0.748	0.751	0.897	0.816	0.909	0.726	0.696	0.710	0.654	0.800	0.439	0.752
low/ff	0.716	0.756	0.829	0.816	0.948	0.702	0.683	0.709	0.674	0.805	0.458	0.697
med/pp	0.852	0.841	0.849	0.803	0.724	0.729	0.914	0.739	0.956	0.776	0.494	0.713
med/mf	0.775	0.778	0.859	0.811	0.712	0.716	0.912	0.737	0.0948	0.789	0.517	0.723
med/ff	0.625	0.673	0.803	0.815	0.723	0.727	0.920	0.728	0.949	0.813	0.505	0.692
high/pp	0.911	0.935	0.795	0.826	0.706	0.723	0.856	0.897	0.954	0.872	0.854	0.803
high/mf	0.930	0.945	0.932	0.715	0.653	0.728	0.847	0.910	0.909	0.894	0.896	0.728
high/ff	0.925	0.632	0.917	0.708	0.534	0.738	0.875	0.881	0.954	0.855	0.801	0.770
mean	0.803	0.783	0.862	0.792	0.760	0.723	0.817	0.781	0.849	0.824	0.606	0.732

Table 8: Oboe test tone.

reg/dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.951	0.950	0.742	0.531	0.835	0.649	0.464	0.561	0.400	0.964	0.272	0.500
low/mf	0.961	0.952	0.764	0.550	0.813	0.696	0.502	0.553	0.430	0.969	0.250	0.563
low/ff	0.963	0.963	0.932	0.548	0.868	0.703	0.501	0.548	0.429	0.966	0.250	0.495
med/pp	0.881	0.808	0.929	0.532	0.989	0.680	0.865	0.601	0.770	0.960	0.298	0.514
med/mf	0.945	0.942	0.914	0.553	0.986	0.664	0.862	0.603	0.773	0.968	0.318	0.520
med/ff	0.911	0.927	0.893	0.547	0.988	0.710	0.884	0.571	0.774	0.958	0.330	0.513
high/pp	0.539	0.578	0.903	0.548	0.979	0.922	0.988	0.541	0.924	0.980	0.581	0.653
high/mf	0.571	0.595	0.926	0.810	0.960	0.852	0.991	0.555	0.950	0.948	0.611	0.592
high/ff	0.547	0.943	0.873	0.815	0.897	0.838	0.982	0.562	0.915	0.994	0.535	0.606
mean	0.808	0.851	0.875	0.604	0.924	0.746	0.782	0.566	0.707	0.967	0.383	0.551

Table 9: Bassoon test tone 1.

reg/dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.522	0.502	0.841	0.902	0.906	0.558	0.649	0.669	0.644	0.595	0.506	0.705
low/mf	0.521	0.528	0.853	0.902	0.904	0.538	0.707	0.665	0.663	0.584	0.457	0.722
low/ff	0.473	0.533	0.689	0.907	0.904	0.515	0.689	0.664	0.690	0.593	0.475	0.697
med/pp	0.811	0.848	0.715	0.905	0.503	0.549	0.999	0.671	0.954	0.552	0.505	0.700
med/mf	0.548	0.546	0.756	0.897	0.487	0.546	0.888	0.670	0.953	0.573	0.513	0.699
med/ff	0.367	0.440	0.638	0.894	0.500	0.535	0.881	0.660	0.955	0.607	0.501	0.661
high/pp	0.935	0.974	0.620	0.916	0.469	0.460	0.691	0.933	0.831	0.740	0.838	0.750
high/mf	0.953	0.985	0.851	0.539	0.398	0.469	0.676	0.945	0.748	0.800	0.888	0.680
high/ff	0.977	0.364	0.874	0.524	0.247	0.483	0.729	0.911	0.851	0.697	0.765	0.728
mean	0.678	0.636	0.760	0.821	0.591	0.517	0.755	0.754	0.810	0.638	0.605	0.705

Table 10: Bassoon test tone 2.

reg dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.410	0.392	0.711	0.733	0.685	0.681	0.620	0.654	0.680	0.493	0.540	0.698
low/mf	0.400	0.413	0.737	0.710	0.654	0.655	0.643	0.652	0.663	0.470	0.495	0.712
low/ff	0.355	0.409	0.501	0.709	0.695	0.607	0.620	0.653	0.698	0.480	0.524	0.675
med/pp	0.502	0.525	0.532	0.699	0.311	0.673	0.618	0.669	0.784	0.447	0.541	0.691
med/mf	0.465	0.476	0.538	0.699	0.297	0.664	0.616	0.661	0.757	0.452	0.570	0.708
med/ff	0.276	0.322	0.533	0.724	0.312	0.644	0.619	0.683	0.759	0.499	0.538	0.671
high/pp	0.954	0.934	0.516	0.728	0.306	0.428	0.493	0.918	0.709	0.502	0.856	0.719
high/mf	0.949	0.928	0.623	0.478	0.251	0.518	0.479	0.921	0.633	0.541	0.882	0.657
high/ff	0.937	0.246	0.620	0.465	0.146	0.548	0.515	0.870	0.702	0.479	0.837	0.706
mean	0.583	0.516	0.590	0.660	0.406	0.602	0.580	0.742	0.710	0.485	0.643	0.693

Table 11: Flute test tone.

reg/dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.965	0.966	0.679	0.415	0.752	0.715	0.405	0.513	0.364	0.973	0.224	0.451
low/mf	0.785	0.694	0.883	0.405	0.987	0.751	0.776	0.565	0.690	0.971	0.256	0.469
low/ff	0.509	0.526	0.901	0.428	0.992	0.978	0.595	0.484	0.898	0.932	0.550	0.618
med/pp	0.975	0.963	0.709	0.429	0.717	0.769	0.435	0.506	0.379	0.975	0.206	0.520
med/mf	0.961	0.962	0.849	0.430	0.987	0.731	0.772	0.566	0.685	0.973	0.286	0.480
med/ff	0.533	0.542	0.860	0.808	0.984	0.922	0.966	0.497	0.940	0.881	0.572	0.558
high/pp	0.984	0.974	0.887	0.425	0.796	0.769	0.431	0.501	0.385	0.971	0.206	0.442
high/mf	0.946	0.942	0.888	0.433	0.988	0.780	0.801	0.541	0.685	0.962	0.294	0.476
high/ff	0.497	0.979	0.787	0.816	0.947	0.911	0.943	0.496	0.874	0.961	0.526	0.570
mean	0.795	0.839	0.827	0.510	0.905	0.814	0.721	0.519	0.655	0.956	0.346	0.509

Table 12: Clarinet test tone.

reg dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.956	0.955	0.715	0.385	0.697	0.879	0.498	0.606	0.521	0.970	0.371	0.575
low/mf	0.959	0.955	0.748	0.383	0.640	0.916	0.525	0.599	0.512	0.964	0.366	0.632
low/ff	0.959	0.957	0.852	0.375	0.743	0.917	0.506	0.593	0.532	0.965	0.357	0.565
med/pp	0.650	0.555	0.847	0.357	0.900	0.904	0.663	0.668	0.658	0.970	0.417	0.597
med/mf	0.970	0.973	0.793	0.379	0.900	0.891	0.657	0.664	0.635	0.962	0.457	0.607
med/ff	0.960	0.953	0.930	0.410	0.903	0.925	0.699	0.650	0.635	0.965	0.456	0.612
high/pp	0.578	0.553	0.940	0.384	0.918	0.981	0.884	0.540	0.884	0.829	0.665	0.725
high/mf	0.583	0.542	0.763	0.872	0.920	0.990	0.887	0.546	0.933	0.765	0.661	0.681
high/ff	0.501	0.932	0.668	0.877	0.910	0.987	0.854	0.544	0.848	0.870	0.675	0.687
mean	0.791	0.819	0.806	0.491	0.837	0.932	0.686	0.601	0.684	0.918	0.492	0.631

Table 13: Cello test tone.

reg/dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.673	0.658	0.984	0.922	0.965	0.661	0.838	0.873	0.825	0766	0.675	0.870
low/mf	0.668	0.679	0.982	0.914	0.953	0.646	0.895	0.869	0.852	0.754	0.652	0.887
low/ff	0.622	0.678	0.860	0.919	0.954	0.643	0.887	0.869	0.851	0.767	0.652	0.872
med/pp	0.817	0.826	0.872	0.914	0.622	0.653	0.867	0.876	0.983	0.736	0.712	0.871
med/mf	0.730	0.724	0.896	0.909	0.606	0.656	0.865	0.877	0.973	0.753	0.715	0.870
med/ff	0.579	0.655	0.809	0.937	0.617	0.650	0.875	0.859	0.971	0.781	0.696	0.860
high/pp	0.900	0.923	0.797	0.912	0.586	0.580	0.803	0.916	0.916	0.798	0.940	0.929
high/mf	0.912	0.882	0.876	0.805	0.535	0.628	0.786	0.923	0.967	0.834	0.960	0.897
high/ff	0.853	0.509	0.868	0.796	0.417	0.629	0.816	0.917	0.937	0.775	0.894	0.921
mean	0.750	0.726	0.883	0.892	0.695	0.638	0.848	0.887	0.908	0.774	0.766	0.886

Table 14: Bass test tone.

reg/dyn	Vio	Vla	Cel	Bas	Flu	Cla	Obo	Bsn	Tpt	FHn	Trb	Tba
low/pp	0.790	0.781	0.913	0.737	0.788	0.824	0.885	0937	0.915	0.808	0.854	0.954
low/mf	0.772	0.796	0.923	0.725	0.751	0.804	0.898	0.934	0.920	0.808	0.847	0.969
low/ff	0.721	0.778	0.846	0.711	0.795	0.829	0.872	0.928	0.929	0.816	0.838	0.951
med/pp	0.675	0.648	0.848	0.706	0.623	0.812	0.691	0.962	0.805	0.808	0.880	0.963
med/mf	0.829	0.822	0.827	0.719	0.615	0.828	0.688	0.960	0.781	0.803	0.894	0.962
med/ff	0.738	0.797	0.911	0.758	0.623	0.817	0.719	0.949	0.780	0.841	0.897	0.959
high/pp	0.733	0.716	0.903	0.721	0.612	0.659	0.722	0.825	0.842	0.703	0.919	0.978
high/mf	0.729	0.655	0.732	0.910	0.591	0.736	0.711	0.817	0.831	0.689	0.898	0.956
high/ff	0.620	0.588	0.688	0.894	0.539	0.734	0.709	0.855	0.837	0.707	0.910	0.960
mean	0.734	0.731	0.844	0.764	0.660	0.783	0.766	0.907	0.849	0.776	0.882	0.961

Appendix C

Experiment Data and Results

Table 15: Values of $\{a_p\}$ and $\{\eta_p\}$ in seconds for Tests 1–5.

Test 1	$\{a_p\} =$	{0	2.9195	4.0268	6.8792	8.0201	10.839
		12.013	14.832	15.973	18.792	19.933	22.819
		23.96	26.812	27.953	30.839	31.913	34.832
		35.939	38.892	39.933	42.718	43.926	46.678}
	$\{\eta_p\} =$	{0.01072	0.77181	0.13421	0.77181	0.16782	0.77165
		0.134	0.73804	0.134	0.67158	0.134	0.73858
		0.13453	0.70481	0.16776	0.63728	0.10023	0.73858
		0.16776	0.50329	0.20153	0.57082	0.16776	0.57082}
Test 2	$\{a_p\} =$	{0.11638	1.6983	4.0733	5.7974}		
	$\{\eta_p\} =$	{0.11638	0.31036	0.21982	0.26292}		
Test 3	$\{a_p\} =$	{0.1906	1.6715	1.9327	3.4017	3.6482	5.1665
		5.393	6.9646	7.1965	8.6215	8.8679	10.205
		10.466	11.847	12.108	13.489	13.736	15.234
		15.466	16.789	17.05	18.607	18.824	20.366}
	$\{\eta_p\} =$	{0.043998	0.058642	0.03206	0.058642	0.0174	0.034426
		0.0028094	0.043794	0.017563	0.058548	0.017329	0.088056
		0.0025752	0.058548	0.032083	0.029508	0.046603	0.044028
		0.017329	0.043794	0.046837	0.014754	0.031849	0.16136}
Test4	$\{a_p\} =$	{2.5257	8.5189	12.393	18.023	22.625	28.211}
	$\{\eta_p\} =$	{0.17121	0.5135	0.19282	0.25675	0.14974	0.29948}
Test 5	$\{a_p\} =$	{1.1939	6.657	9.2438	14.888	18.234	23.715}
	$\{\eta_p\} =$	{0.072348	0.43397	0.12684	0.19907	0.10864	0.19878}

Table 16: Frequency ranges in Hz for individual notes in Tests 1–5.

Test	Notes	Freq range
Test 1	1, 4, 7, 10	261–265 Hz.
	2, 5, 8, 11	391–395
	3, 6, 9, 12	657–661
Test 2	1	233–237
	2	244–248
Test 3	1	258–262
	2	275–279
	3	291–295
	4	309–313
	5	327–331
	6	348–352
	7	369–373
	8	390–394
	9	413–417
	10	437–441
	11	462–466
	12	493–497
Test 4	1	437–441
	2	458–462
	3	487–491
Test 5	1	436–440
	2	465–469
	3	500–504

Table 17: Results Test 1

Analyzing Instrument — Clarinet notes

Clarinet	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.013	0.028	0.001	0.146	0.066	0.046	0.407	0.490	0.046	0.084	0.561	0.310	0.577	1.047	0.012
F/G	0.589	0.405	.376	0.114	0.022	0.369	4.163	0.357	0.221	0.122	0.234	0.219	0.677	0.290	0.321
B/G	0.198	0.138	0.043	0.041	0.015	0.285	1.713	0.463	0.042	0.072	0.009	0.251	0.503	0.749	0.215
B/R	0.379	0.351	0.369	0.182	0.042	0.436	3.973	0.273	0.205	0.079	0.558	0.357	0.671	0.990	0.050

Oboe	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.044	0.010	0.015	0.074	0.035	0.217	0.287	0.230	0.025	0.046	0.291	0.152	0.345	0.560	0.026
F/G	0.365	0.602	0.282	0.139	0.039	0.251	0.824	0.136	0.064	0.021	0.225	0.200	0.266	0.550	0.070
B/G	0.334	0.159	0.065	0.095	0.026	0.250	1.790	0.061	0.080	0.003	0.234	0.114	0.345	0.314	0.181
B/R	0.131	0.039	0.187	0.095	0.026	0.220	1.806	0.173	0.100	0.037	0.292	0.182	0.400	0.525	0.011

Trumpet	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.072	0.022	0.010	0.102	0.036	0.207	0.200	0.200	0.022	0.042	0.265	0.138	0.324	0.499	0.023
F/G	0.199	0.086	0.105	0.148	0.049	0.181	2.014	0.066	0.076	0.036	0.025	0.131	0.321	0.278	0.095
B/G	0.334	0.582	0.252	0.149	0.041	0.186	0.837	0.145	0.042	0.001	0.048	0.116	0.327	0.064	0.144
B/R	0.131	0.107	0.171	0.087	0.022	0.202	1.703	0.141	0.091	0.034	0.263	0.169	0.368	0.469	0.015

Violin	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.009	0.160	0.039	0.013	0.013	0.157	0.211	0.167	0.017	0.033	0.208	0.110	0.256	0.387	0.021
F/G	0.183	0.264	0.005	0.086	0.005	0.125	1.031	0.249	0.023	0.007	0.090	0.155	0.039	0.172	0.067
B/G	0.058	0.064	0.163	0.098	0.007	0.130	0.753	0.105	0.050	0.025	0.018	0.005	0.234	0.018	0.095
B/R	0.113	0.075	0.135	0.069	0.018	0.153	1.300	0.117	0.071	0.027	0.190	0.069	0.164	0.410	0.007

Results Test 1 (Cont.)

Analyzing Instrument — Oboe Notes

Clarinet

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.067	0.037	0.022	0.016	0.010	0.015	0.162	0.530	0.014	0.062	0.059	0.024	0.047	0.067	0.551
F/G	0.005	0.008	0.002	0.017	0.004	0.047	0.067	0.431	0.032	0.030	0.030	0.039	0.029	0.116	0.160
B/G	0.020	0.001	0.058	0.016	0.005	0.015	0.085	0.455	0.020	0.060	0.022	0.029	0.016	0.007	0.408
B/R	0.115	0.003	0.017	0.016	0.005	0.037	0.074	0.446	0.011	0.017	0.060	0.031	0.045	0.074	0.528

Oboe

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.042	0.031	0.058	0.021	0.005	0.016	0.110	0.646	0.016	0.017	0.009	0.005	0.060	0.006	0.327
F/G	0.001	0.001	0.069	0.015	0.002	0.002	0.113	0.123	0.084	0.033	0.029	0.003	0.014	0.021	0.281
B/G	0.005	0.036	0.014	0.000	0.013	0.019	0.032	0.532	0.050	0.016	0.016	0.028	0.150	0.059	0.020
B/R	0.091	0.007	0.039	0.012	0.011	0.011	0.004	0.001	0.123	0.034	0.054	0.007	0.019	0.000	0.327

Trumpet

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.051	0.032	0.044	0.030	0.006	0.015	0.123	0.654	0.006	0.005	0.001	0.005	0.063	0.014	0.288
F/G	0.108	0.003	0.007	0.009	0.002	0.022	0.063	0.508	0.048	0.018	0.028	0.012	0.031	0.032	0.049
B/G	0.023	0.034	0.026	0.011	0.000	0.023	0.040	0.025	0.067	0.017	0.051	0.032	0.093	0.040	0.000
B/R	0.063	0.010	0.031	0.018	0.008	0.014	0.009	0.096	0.105	0.045	0.058	0.004	0.019	0.007	0.298

Violin

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.086	0.041	0.034	0.026	0.006	0.005	0.131	0.445	0.064	0.012	0.011	0.038	0.070	0.003	0.226
F/G	0.026	0.054	0.051	0.011	0.007	0.029	0.079	0.360	0.013	0.065	0.041	0.018	0.046	0.001	0.087
B/G	0.001	0.014	0.017	0.037	0.013	0.009	0.114	0.126	0.105	0.084	0.045	0.015	0.168	0.002	0.055
B/R	0.053	0.008	0.025	0.019	0.009	0.007	0.001	0.108	0.107	0.045	0.042	0.006	0.119	0.013	0.202

Results Test 1 (Cont.)
Trumpet Notes

Analyzing Instrument

Clarinet	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.079	0.062	0.113	0.056	0.016	0.046	0.480	0.071	0.046	0.011	0.106	0.070	0.010	0.176	0.067
F/G	0.060	0.150	0.085	0.032	0.007	0.046	0.928	0.037	0.034	0.015	0.012	0.131	0.483	0.076	0.009
B/G	0.079	0.111	0.075	0.016	0.002	0.039	0.046	0.076	0.033	0.016	0.052	0.123	0.444	0.010	0.069
B/R	0.009	0.022	0.003	0.033	0.018	0.045	0.845	0.002	0.021	0.025	0.054	0.051	0.094	0.185	0.067

Oboe	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.025	0.028	0.042	0.052	0.012	0.047	0.790	0.016	0.023	0.005	0.053	0.040	0.067	0.079	0.016
F/G	0.034	0.148	0.051	0.002	0.014	0.017	0.304	0.003	0.013	0.007	0.019	0.010	0.102	0.075	0.036
B/G	0.026	0.001	0.067	0.027	0.019	0.044	0.193	0.031	0.005	0.006	0.099	0.018	0.342	0.056	0.011
B/R	0.039	0.126	0.030	0.003	0.006	0.009	0.434	0.009	0.037	0.020	0.041	0.011	0.033	0.139	0.046

Trumpet	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.041	0.008	0.028	0.047	0.016	0.053	0.858	0.004	0.017	0.005	0.062	0.040	0.023	0.059	0.009
F/G	0.014	0.029	0.001	0.052	0.025	0.063	0.423	0.115	0.006	0.017	0.013	0.028	0.221	0.009	0.033
B/G	0.144	0.024	0.013	0.021	0.029	0.004	0.346	0.031	0.009	0.009	0.064	0.057	0.208	0.001	0.033
B/R	0.040	0.142	0.029	0.005	0.007	0.015	0.455	0.012	0.031	0.024	0.029	0.001	0.029	0.138	0.050

Violin	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.035	0.023	0.040	0.030	0.007	0.021	1.327	0.006	0.009	0.004	0.046	0.036	0.021	0.040	0.004
F/G	0.081	0.026	0.031	0.015	0.007	0.010	0.223	0.014	0.010	0.003	0.022	0.009	0.080	0.053	0.013
B/G	0.090	0.118	0.032	0.006	0.009	0.069	0.299	0.077	0.010	0.005	0.012	0.055	0.099	0.053	0.015
B/R	0.036	0.125	0.031	0.002	0.006	0.017	0.452	0.069	0.027	0.021	0.042	0.001	0.029	0.114	0.041

Results Test 1 (Cont.)

Analyzing Instrument — Violin Notes

Clarinet	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.070	0.283	0.081	0.067	0.047	0.044	1.000	0.038	0.004	0.007	0.010	0.026	0.488	0.131	0.010
F/G	0.065	0.230	0.147	0.073	0.050	0.047	0.628	0.023	0.016	0.012	0.044	0.072	0.021	0.051	0.048
B/G	0.099	0.276	0.169	0.042	0.015	0.059	0.855	0.040	0.004	0.012	0.038	0.066	0.166	0.012	0.099
B/R	0.013	0.608	0.129	0.018	0.008	0.018	0.945	0.020	0.015	0.013	0.003	0.010	0.444	0.125	0.009

Oboe	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.030	0.128	0.022	0.043	0.015	0.030	0.509	0.014	0.005	0.002	0.001	0.009	0.278	0.075	0.032
F/G	0.011	0.284	0.050	0.007	0.029	0.010	0.355	0.034	0.015	0.003	0.025	0.010	0.134	0.071	0.112
B/G	0.012	0.105	0.082	0.011	0.027	0.016	0.556	0.018	0.003	0.000	0.013	0.004	0.229	0.083	0.045
B/R	0.012	0.300	0.084	0.011	0.012	0.029	0.536	0.004	0.009	0.008	0.003	0.001	0.186	0.053	0.051

Trumpet	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.038	0.119	0.018	0.043	0.015	0.028	0.431	0.009	0.006	0.003	0.003	0.013	0.265	0.064	0.025
F/G	0.001	0.285	0.001	0.040	0.017	0.001	0.328	0.016	0.006	0.001	0.013	0.030	0.172	0.013	0.057
B/G	0.001	0.272	0.013	0.006	0.024	0.030	0.429	0.018	0.009	0.006	0.010	0.024	0.097	0.014	0.057
B/R	0.027	0.275	0.072	0.010	0.011	0.019	0.527	0.008	0.015	0.008	0.002	0.004	0.161	0.049	0.056

Violin	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.030	0.094	0.007	0.035	0.012	0.033	0.386	0.026	0.007	0.008	0.004	0.008	0.218	0.049	0.013
F/G	0.006	0.069	0.010	0.023	0.010	0.008	0.276	0.002	0.020	0.002	0.009	0.034	0.123	0.021	0.046
B/G	0.065	0.182	0.004	0.027	0.018	0.012	0.402	0.015	0.014	0.004	0.003	0.005	0.208	0.011	0.047
B/R	0.025	0.213	0.057	0.008	0.009	0.022	0.424	0.013	0.017	0.007	0.002	0.002	0.130	0.037	0.040

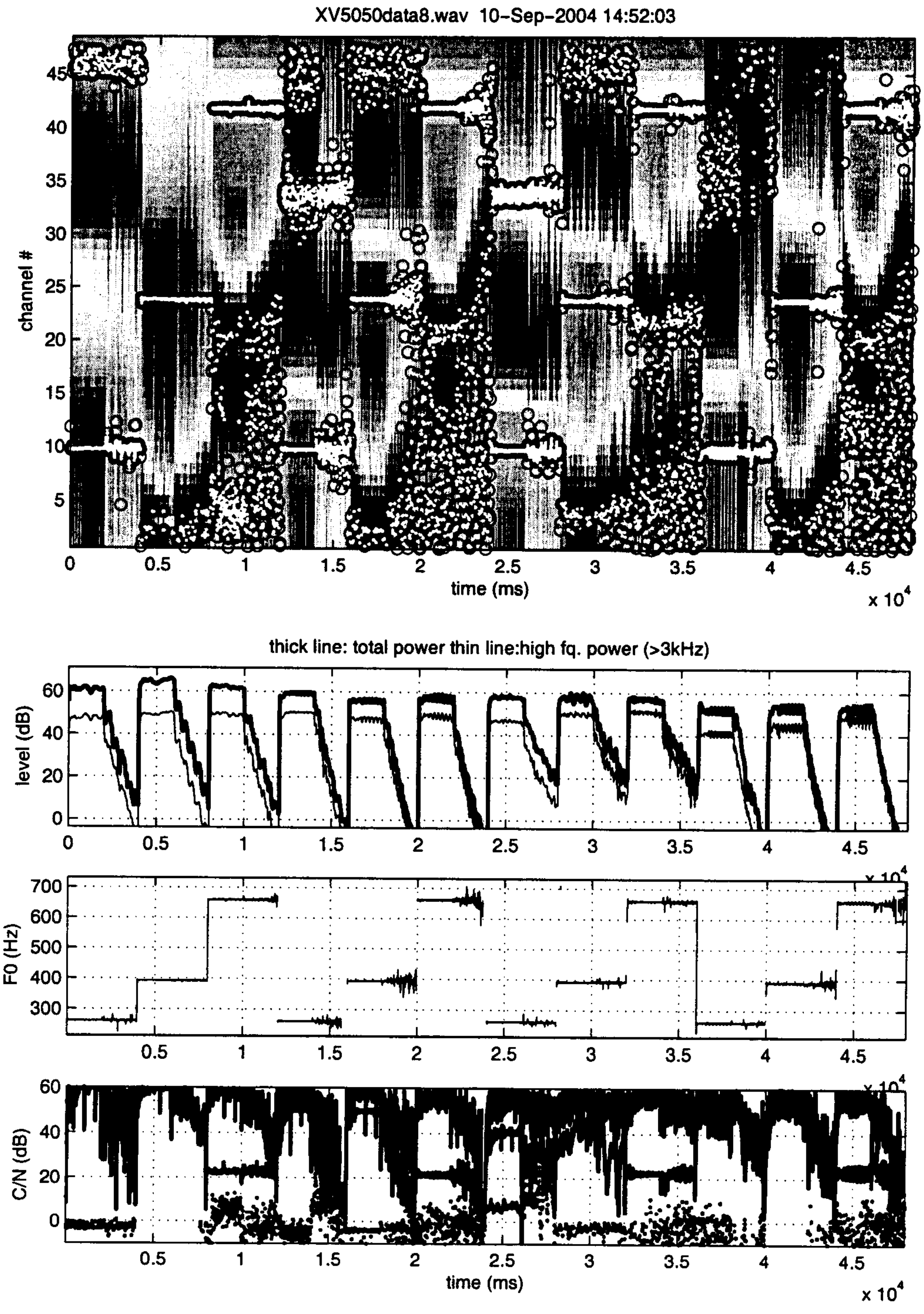

Figure 15: Recording for Test 1.

Table 18: Results Test 2

Analyzing Instrument — oboe.mf.B♭3B3

Clarinet	Frequency range Note 1					Frequency range Note 2				
F/R	0.106	0.242	0.927	0.201	0.017	0.028	0.151	1.284	0.111	0.056
F/G	0.065	0.258	1.407	0.186	0.030	0.024	0.556	0.582	0.043	0.069
B/G	0.193	0.259	0.810	0.151	0.001	0.052	0.596	0.293	0.066	0.051
B/R	0.181	0.274	1.237	0.228	0.001	0.067	0.178	0.591	0.067	0.060

Oboe	Frequency range Note 1					Freq. range Note 2				
F/R	0.103	0.230	0.518	0.065	0.028	0.036	0.292	1.745	0.056	0.029
F/G	0.003	0.301	1.852	0.042	0.020	0.152	0.183	0.222	0.005	0.041
B/G	0.113	0.230	0.158	0.084	0.006	0.027	1.038	0.351	0.012	0.012
B/R	0.106	0.309	0.555	0.133	0.019	0.073	0.418	0.066	0.047	0.017

Trumpet	Freq. range Note 1					Freq. range Note 2				
F/R	0.007	0.118	0.088	0.138	0.038	0.020	0.147	1.488	0.004	0.021
F/G	0.011	0.116	0.907	0.114	0.014	0.046	0.135	0.367	0.004	0.002
B/G	0.048	0.301	0.198	0.102	0.023	0.020	0.007	1.780	0.005	0.017
B/R	0.052	0.276	0.205	0.110	0.013	0.024	0.171	0.307	0.007	0.023

Violin	Freq. range Note 1					Freq. range Note 2				
F/R	0.028	0.047	0.240	0.068	0.005	0.024	0.199	1.395	0.015	0.020
F/G	0.023	0.055	0.275	0.008	0.026	0.012	0.050	0.160	0.026	0.010
B/G	0.043	0.010	0.232	0.058	0.017	0.078	0.147	1.192	0.014	0.033
B/R	0.039	0.237	0.132	0.091	0.010	0.010	1.020	0.069	0.016	0.001

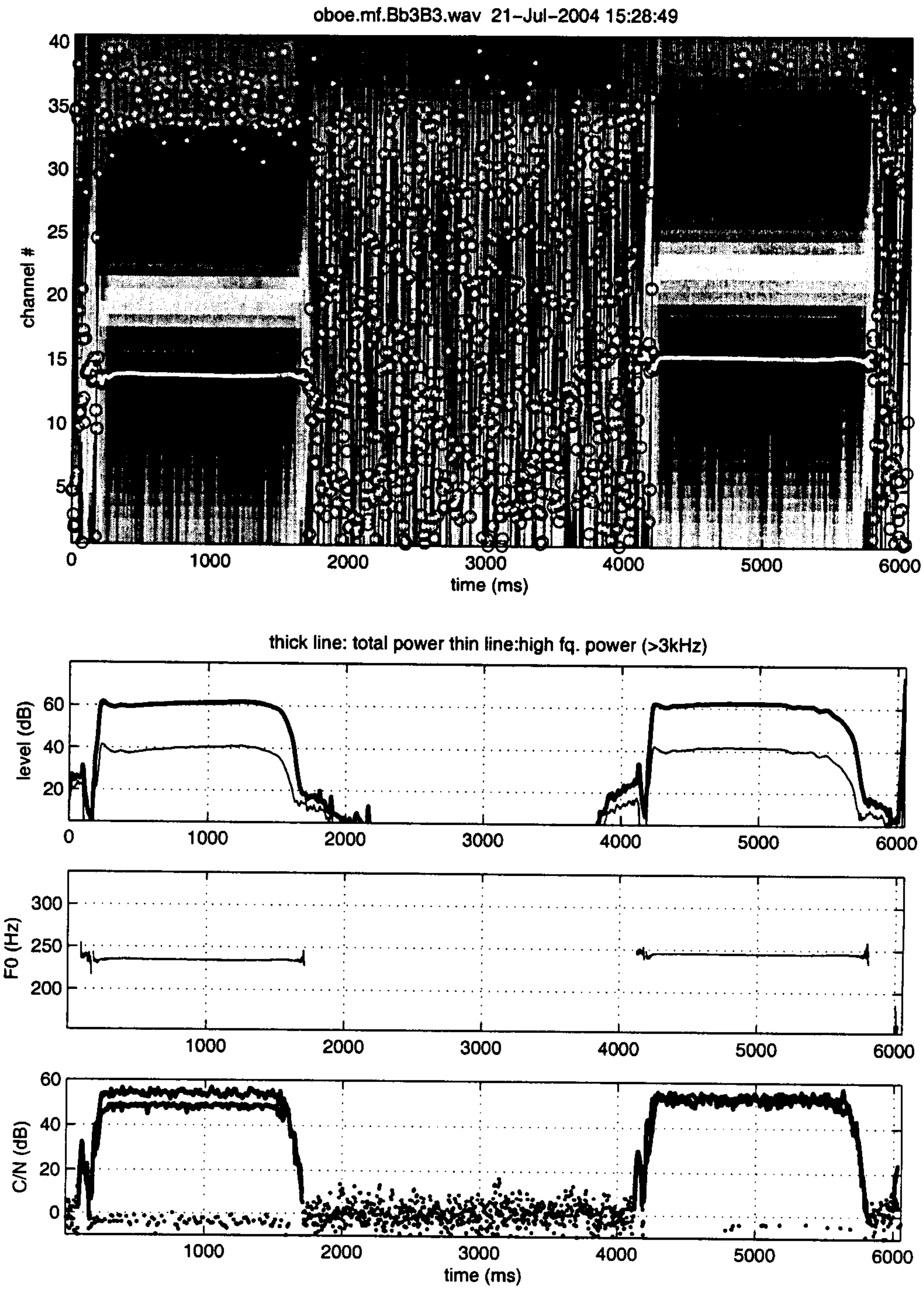

Figure 16: Recording for Test 2.

Table 19: Results Test 3

oboe.ff.C4B4

Analyzing Instrument: Clarinet

	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.003	0.130	0.496	0.349	0.072	0.086	0.208	0.419	0.622	0.148	0.109	0.022	0.564	0.539	0.438
F/G	0.112	0.094	0.321	2.682	0.162	0.135	0.297	0.358	1.545	0.329	0.055	0.071	0.195	2.753	0.156
B/G	0.162	0.395	1.887	3.791	0.100	0.101	0.144	0.839	0.502	0.312	0.112	0.277	0.318	0.056	0.497
B/R	0.108	0.091	1.729	1.727	0.490	0.046	0.047	2.560	0.846	0.236	0.081	0.239	0.452	1.499	0.136

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.214	0.788	1.497	0.493	0.025	0.138	0.051	0.623	0.877	0.083	0.565	0.295	0.111	0.074	0.081
F/G	0.109	0.721	1.587	0.198	0.022	0.118	0.082	0.389	0.666	0.583	0.455	1.115	0.213	0.081	0.074
B/G	0.410	1.515	0.262	0.557	0.010	0.149	0.087	0.510	0.363	0.612	0.040	0.515	0.474	0.111	0.099
B/R	0.140	0.862	0.215	0.173	0.001	0.029	0.089	0.251	2.316	0.643	0.038	0.506	0.445	0.141	0.091

	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.032	0.227	1.806	0.552	0.284	0.078	0.220	0.410	0.091	0.053	0.175	0.193	0.181	0.055	0.049
F/G	0.104	0.367	1.257	0.355	0.207	0.186	0.066	0.382	0.055	0.040	0.077	0.318	0.341	0.150	0.026
B/G	0.008	0.367	0.321	0.533	0.241	0.274	0.092	0.584	0.129	0.001	0.069	0.371	0.748	0.263	0.033
B/R	0.079	0.145	0.896	0.560	0.065	0.132	0.620	0.245	0.035	0.024	0.116	0.193	0.294	0.532	0.063

	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.420	0.094	0.690	1.449	0.758	0.964	1.582	1.568	1.622	2.247	0.218	0.560	1.351	0.430	0.096
F/G	0.548	0.613	0.622	0.595	0.389	0.800	0.108	0.125	0.929	1.116	0.061	0.026	0.394	0.308	0.087
B/G	0.334	0.850	0.677	0.359	0.006	1.014	1.647	1.029	0.427	1.659	0.247	0.125	1.754	0.171	0.006
B/R	0.036	0.664	0.327	2.836	0.075	0.714	0.283	0.896	0.992	1.755	0.162	0.650	0.010	0.418	0.002

Results Test 3 (Cont.)

Analyzing Instrument: Oboe

	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.039	0.002	0.413	0.482	0.081	0.058	0.211	0.067	0.471	0.023	0.028	0.010	0.557	0.103	0.248
F/G	0.050	0.269	0.474	3.095	0.723	0.058	0.011	0.487	0.120	0.231	0.008	0.075	0.223	0.185	0.044
B/G	0.035	0.052	0.237	1.912	0.237	0.088	0.056	0.122	1.245	0.179	0.165	0.192	0.743	0.251	0.234
B/R	0.051	0.099	0.969	1.547	0.020	0.054	0.182	0.617	1.431	0.037	0.023	0.030	0.142	1.182	0.011

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.112	0.604	2.215	0.296	0.033	0.113	0.089	0.674	1.219	0.160	0.547	0.245	1.255	0.040	0.096
F/G	0.005	0.180	0.291	0.022	0.025	0.040	0.096	0.017	0.475	0.451	0.083	3.373	1.611	0.065	0.077
B/G	0.027	0.220	0.031	0.018	0.098	0.040	0.046	0.223	0.466	0.149	0.008	0.352	0.127	0.100	0.028
B/R	0.109	0.418	0.791	0.019	0.035	0.014	0.148	0.544	0.097	0.617	0.110	0.958	0.118	0.081	0.021

	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.003	0.038	1.490	1.367	0.137	0.073	0.074	0.311	0.074	0.031	0.093	0.265	1.060	0.446	0.131
F/G	0.081	0.026	2.486	0.900	0.156	0.065	0.176	0.178	0.112	0.020	0.008	0.108	0.343	0.374	0.047
B/G	0.023	0.273	2.232	0.470	0.092	0.040	0.252	0.131	0.061	0.034	0.248	0.040	0.793	1.384	0.175
B/R	0.099	0.119	0.949	0.480	0.078	0.112	0.918	0.065	0.020	0.021	0.059	0.099	0.753	0.180	0.149

	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.331	0.201	0.613	3.638	0.805	1.355	2.460	2.662	2.423	3.045	0.286	0.514	0.652	0.529	0.679
F/G	0.720	0.396	0.055	1.450	0.478	0.764	1.593	2.376	1.002	0.866	0.414	0.373	1.681	0.113	0.934
B/G	1.072	0.190	0.770	0.263	0.285	1.211	1.260	2.240	2.809	2.912	0.106	0.485	0.937	0.478	0.065
B/R	0.189	1.470	0.364	3.055	0.339	1.164	0.871	1.555	1.592	3.808	0.305	0.735	1.179	0.221	0.090

Results Test 3 (Cont.)

Analyzing Instrument: Trumpet

	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.001	0.042	0.099	0.062	0.146	0.041	0.125	0.460	0.159	0.029	0.065	0.042	0.435	0.147	0.294
F/G	0.015	0.074	0.433	0.983	0.109	0.041	0.057	1.442	1.063	0.082	0.023	0.259	0.480	2.245	0.115
B/G	0.004	0.007	0.097	1.230	0.487	0.029	0.129	0.334	0.615	0.047	0.003	0.074	0.056	0.118	0.487
B/R	0.009	0.010	0.500	1.049	0.031	0.013	0.110	0.272	0.743	0.023	0.008	0.013	0.157	1.348	0.034

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.101	0.635	1.682	0.271	0.013	0.110	0.079	0.606	0.930	0.149	0.517	0.129	0.938	0.056	0.087
F/G	0.018	0.155	0.557	0.068	0.004	0.024	0.048	0.015	0.536	0.033	0.024	0.225	0.487	0.109	0.043
B/G	0.047	0.114	0.135	0.014	0.020	0.024	0.056	0.073	1.011	0.161	0.128	0.006	0.140	0.014	0.024
B/R	0.064	0.418	0.281	0.045	0.037	0.001	0.114	0.421	0.302	0.494	0.032	0.784	0.109	0.076	0.025

	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.009	0.079	1.447	1.089	0.142	0.038	0.136	0.265	0.054	0.023	0.071	0.213	0.892	0.216	0.129
F/G	0.091	0.040	0.585	0.365	0.071	0.159	0.076	0.009	0.004	0.029	0.247	0.094	0.470	0.500	0.076
B/G	0.094	0.041	1.314	1.508	0.024	0.046	0.197	0.164	0.064	0.044	0.023	0.046	0.193	0.552	0.048
B/R	0.112	0.136	0.919	0.303	0.100	0.117	0.771	0.133	0.023	0.017	0.048	0.084	0.356	0.247	0.138

	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.405	0.202	0.551	3.154	0.512	1.046	2.062	2.330	1.906	2.811	0.300	0.482	0.292	0.494	0.597
F/G	0.397	0.058	0.530	2.194	0.437	0.755	0.883	0.463	1.397	3.895	0.253	0.091	0.520	0.638	0.237
B/G	0.262	0.363	0.384	0.463	0.639	0.567	0.615	1.608	2.078	2.277	0.096	0.223	1.160	0.313	0.931
B/R	0.119	1.374	0.376	2.096	0.327	1.002	0.535	1.240	1.265	2.936	0.256	0.687	1.070	0.138	0.140

Results Test 3 (Cont.)

Analyzing Instrument: Violin

	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.027	0.041	0.180	0.323	0.025	0.050	0.130	0.256	0.250	0.025	0.057	0.059	0.516	0.201	0.182
F/G	0.031	0.093	1.241	0.626	0.120	0.103	0.235	1.989	1.737	0.289	0.004	0.015	0.115	0.395	0.121
B/G	0.029	0.059	0.590	1.335	0.036	0.023	0.057	2.114	2.002	0.079	0.012	0.077	0.622	0.657	0.352
B/R	0.010	0.001	0.556	0.706	0.081	0.049	0.159	0.167	1.073	0.003	0.008	0.026	0.073	0.992	0.017

	Freq. range Note 4					Freq. range Note 5					Freq. range Note 6				
F/R	0.104	0.465	1.634	0.223	0.030	0.093	0.108	0.572	1.112	0.218	0.475	0.125	0.953	0.041	0.082
F/G	0.108	1.367	0.319	0.207	0.108	0.043	0.041	0.148	1.112	0.196	0.182	1.757	2.071	0.234	0.159
B/G	0.291	0.592	0.815	0.048	0.168	0.058	0.027	0.014	1.249	0.136	0.068	1.007	0.264	0.099	0.029
B/R	0.070	0.360	0.868	0.049	0.034	0.014	0.109	0.438	0.613	0.491	0.052	0.901	0.066	0.088	0.020

	Freq. range Note 7					Freq. range Note 8					Freq. range Note 9				
F/R	0.012	0.045	1.151	1.055	0.114	0.029	0.229	0.313	0.051	0.010	0.017	0.239	1.105	0.238	0.100
F/G	0.054	0.241	1.062	0.463	0.026	0.025	0.338	0.078	0.043	0.025	0.074	0.083	0.641	1.154	0.022
B/G	0.160	0.022	1.500	0.087	0.125	0.101	0.125	0.077	0.034	0.007	0.145	0.101	0.007	0.166	0.063
B/R	0.093	0.170	1.180	0.109	0.088	0.094	0.699	0.094	0.010	0.010	0.012	0.074	0.154	0.465	0.131

	Freq. range Note 10					Freq. range Note 11					Freq. range Note 12				
F/R	0.244	0.072	0.447	2.119	0.327	0.766	1.539	1.797	1.385	2.200	0.329	0.348	0.196	0.307	0.501
F/G	0.317	0.102	1.440	1.917	0.040	0.466	0.744	0.635	0.137	3.315	0.062	0.252	0.637	0.352	0.209
B/G	0.058	0.065	0.007	0.032	0.317	0.677	0.251	0.098	0.702	0.660	0.163	0.313	0.714	0.375	0.701
B/R	0.218	1.220	0.153	1.406	0.178	0.689	0.379	0.937	0.933	2.066	0.201	0.541	1.000	0.011	0.261

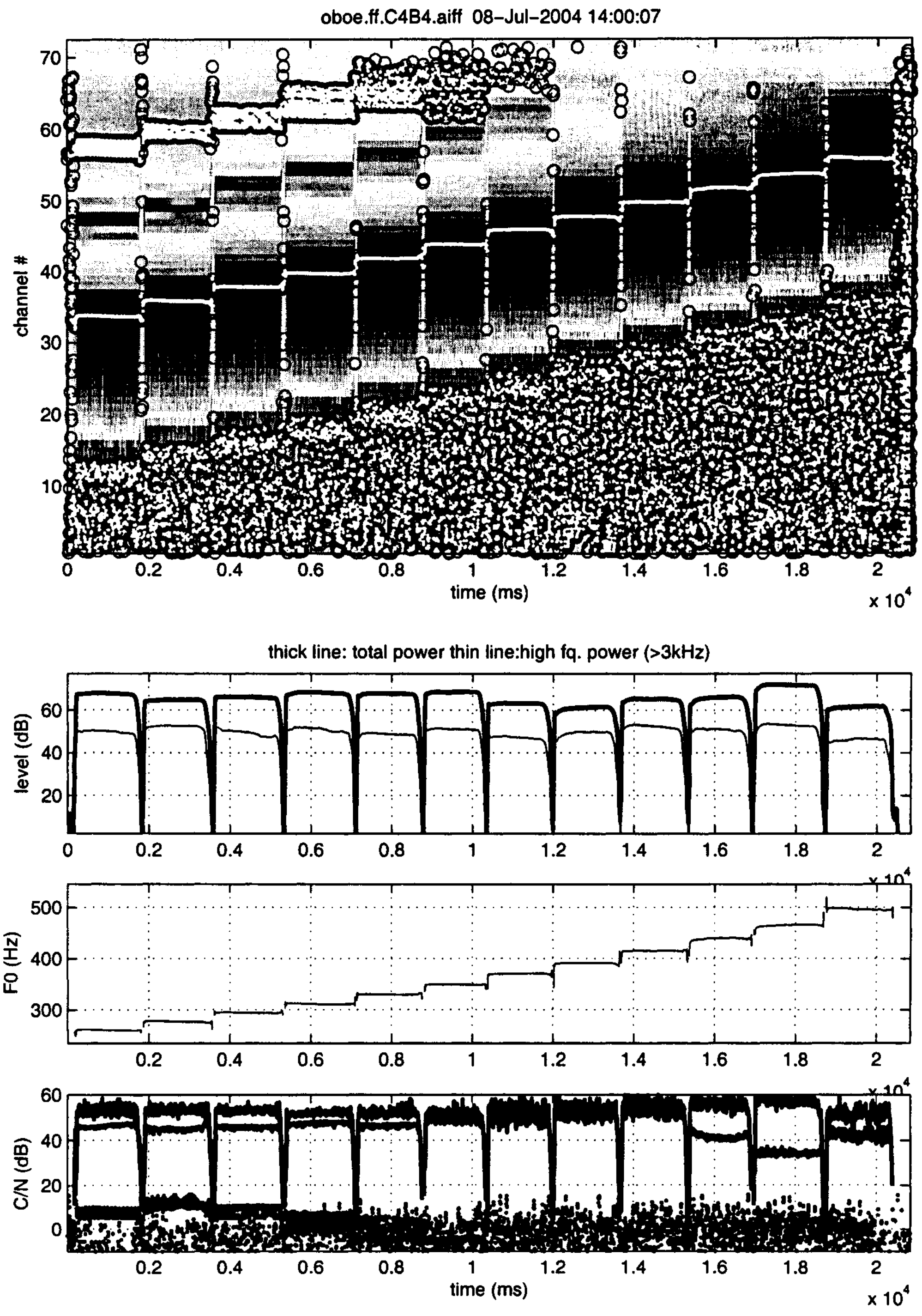

Figure 17: Recording for Test 3.

Table 20: Results Test 4

Violin.arco.mf.sulA.A4B4

Clarinet	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.936	0.014	0.893	0.000	0.082	0.032	0.261	0.827	0.819	0.545	0.060	0.164	0.488	0.085	0.005
F/G	0.778	2.781	0.516	0.059	0.066	0.045	0.340	0.760	2.177	0.585	0.179	0.009	0.191	0.074	0.020
B/G	0.177	3.568	0.106	0.062	0.088	0.047	0.181	0.175	0.750	0.883	0.074	0.006	0.205	0.029	0.022
B/R	0.688	1.905	0.709	0.045	0.040	0.081	0.367	1.171	1.587	0.107	0.034	0.190	0.221	0.060	0.001

Oboe	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.505	0.245	0.325	0.003	0.029	0.004	0.108	0.409	0.300	0.497	0.080	0.612	0.254	0.048	0.003
F/G	0.129	1.912	0.148	0.037	0.027	0.008	0.114	0.277	0.923	0.556	0.115	0.524	0.453	0.017	0.008
B/G	0.188	1.218	0.533	0.008	0.020	0.032	0.215	0.563	1.297	0.574	0.229	0.149	0.072	0.040	0.002
B/R	0.378	0.811	0.413	0.025	0.011	0.047	0.196	0.724	0.983	0.004	0.007	0.285	0.038	0.036	0.006

Trumpet	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.473	0.158	0.302	0.002	0.024	0.006	0.101	0.456	0.169	0.376	0.091	0.504	0.204	0.044	0.002
F/G	0.223	1.668	0.183	0.017	0.012	0.019	0.135	0.119	0.103	0.306	0.049	0.513	0.173	0.002	0.020
B/G	0.382	0.453	0.209	0.018	0.047	0.013	0.124	0.277	1.291	0.458	0.021	0.241	0.186	0.046	0.006
B/R	0.355	0.731	0.418	0.024	0.013	0.045	0.184	0.580	0.805	0.080	0.015	0.341	0.130	0.026	0.006

Violin	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.352	0.100	0.199	0.001	0.017	0.002	0.071	0.355	0.119	0.374	0.087	0.576	0.166	0.038	0.002
F/G	0.210	0.073	0.064	0.006	0.009	0.027	0.055	0.291	0.427	0.359	0.055	0.035	0.279	0.030	0.007
B/G	0.431	1.092	0.055	0.012	0.026	0.004	0.174	0.205	0.585	0.075	0.091	0.663	0.028	0.017	0.017
B/R	0.285	0.515	0.336	0.020	0.010	0.044	0.156	0.480	0.687	0.071	0.025	0.197	0.095	0.026	0.005

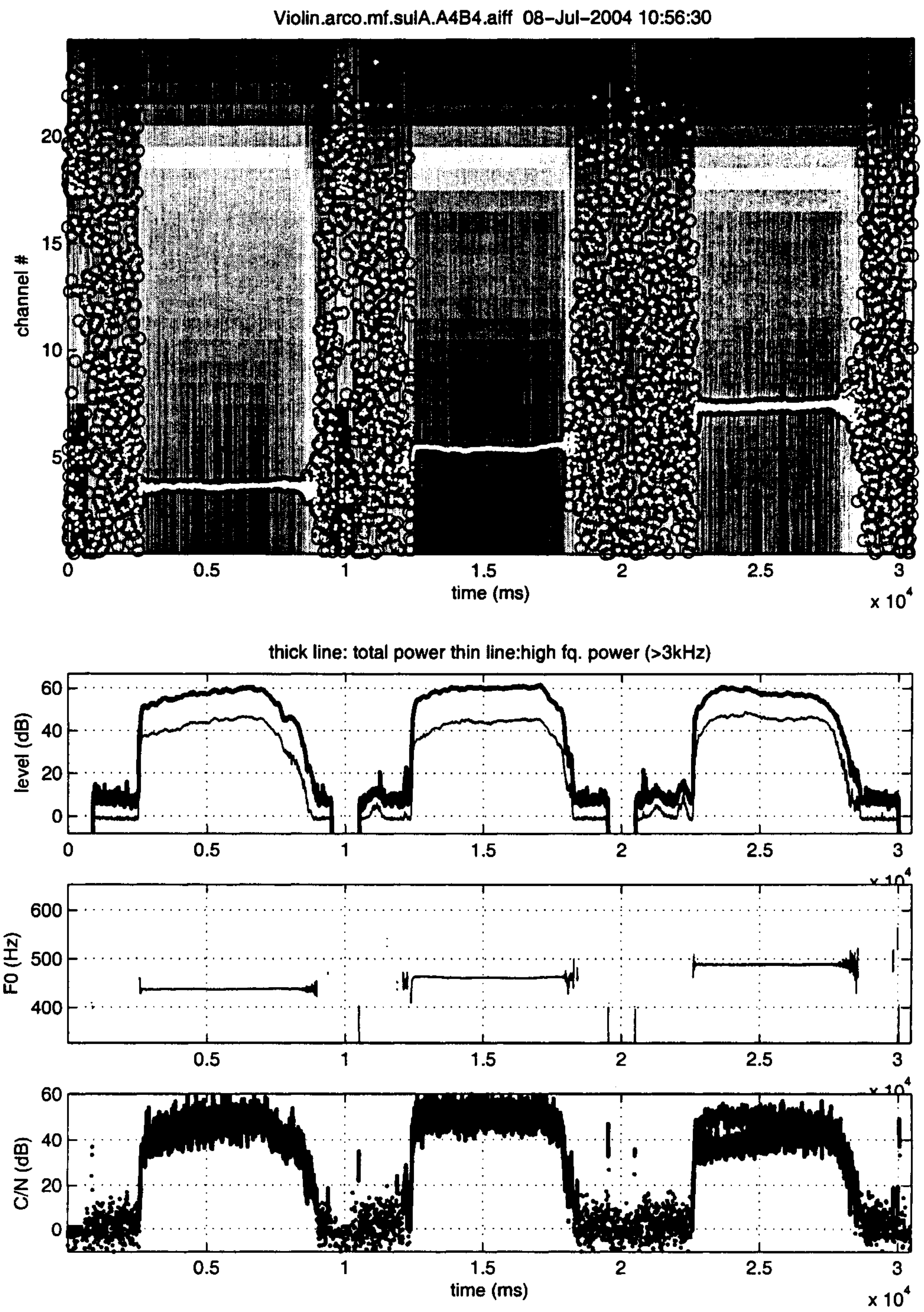

Figure 18: Recording for Test 4.

Table 21: Results Test 5

Violin.arco.ff.sulA.A4B4

Clarinet	Freq. range Note 1					Freq. range Note 2					Freq. range Note 3				
F/R	0.185	0.795	1.398	1.392	0.627	2.214	0.494	4.723	0.358	0.335	0.226	0.208	0.544	0.038	0.070
F/G	0.967	0.738	2.773	5.499	0.782	1.771	0.636	4.452	0.926	0.376	0.224	0.376	0.384	1.306	0.045
B/G	1.734	0.352	2.792	4.883	0.113	0.973	2.649	8.231	1.825	0.222	0.091	0.261	1.027	0.314	0.099
B/R	1.881	0.371	1.525	4.711	1.225	0.138	2.850	8.562	1.429	0.070	0.090	0.428	0.622	1.858	0.111
Oboe	**Freq. range Note 1**					**Freq. range Note 2**					**Freq. range Note 3**				
F/R	0.103	0.132	0.586	0.380	0.458	1.397	0.476	1.748	0.169	0.238	0.073	0.496	0.565	2.248	0.042
F/G	0.960	0.168	0.892	2.712	0.815	0.554	0.249	1.567	0.614	0.131	0.070	0.817	0.510	0.619	0.013
B/G	0.937	0.293	0.818	0.579	0.664	0.770	0.874	3.744	0.789	0.123	0.148	0.354	0.046	0.598	0.086
B/R	0.787	0.272	0.595	2.249	0.653	0.066	0.989	3.824	0.488	0.031	0.147	0.556	0.737	1.679	0.129
Trumpet	**Freq. range Note 1**					**Freq. range Note 2**					**Freq. range Note 3**				
F/R	0.025	0.186	0.581	0.334	0.282	1.347	0.460	1.572	0.255	0.229	0.016	0.589	0.414	2.524	0.057
F/G	0.268	0.507	0.236	1.425	0.540	1.305	0.467	1.683	0.528	0.124	0.017	0.598	0.404	0.372	0.162
B/G	0.489	0.383	0.214	0.087	0.210	0.274	0.947	3.972	0.469	0.024	0.266	0.373	0.511	2.538	0.026
B/R	0.634	0.238	0.507	1.994	0.398	0.173	0.689	3.898	0.442	0.043	0.267	0.562	0.841	1.291	0.149
Violin	**Freq. range Note 1**					**Freq. range Note 2**					**Freq. range Note 3**				
F/R	0.037	0.084	0.455	0.168	0.244	1.175	0.523	0.974	0.179	0.195	0.035	0.502	0.415	2.434	0.072
F/G	0.688	0.123	0.439	1.707	0.647	0.392	0.585	4.625	0.337	0.194	0.120	0.081	0.485	1.246	0.090
B/G	0.296	0.128	0.745	1.864	0.182	0.511	1.144	0.588	0.449	0.029	0.167	0.575	0.188	2.526	0.120
B/R	0.468	0.195	0.384	1.489	0.298	0.036	0.368	3.112	0.268	0.028	0.201	0.525	0.743	1.177	0.127

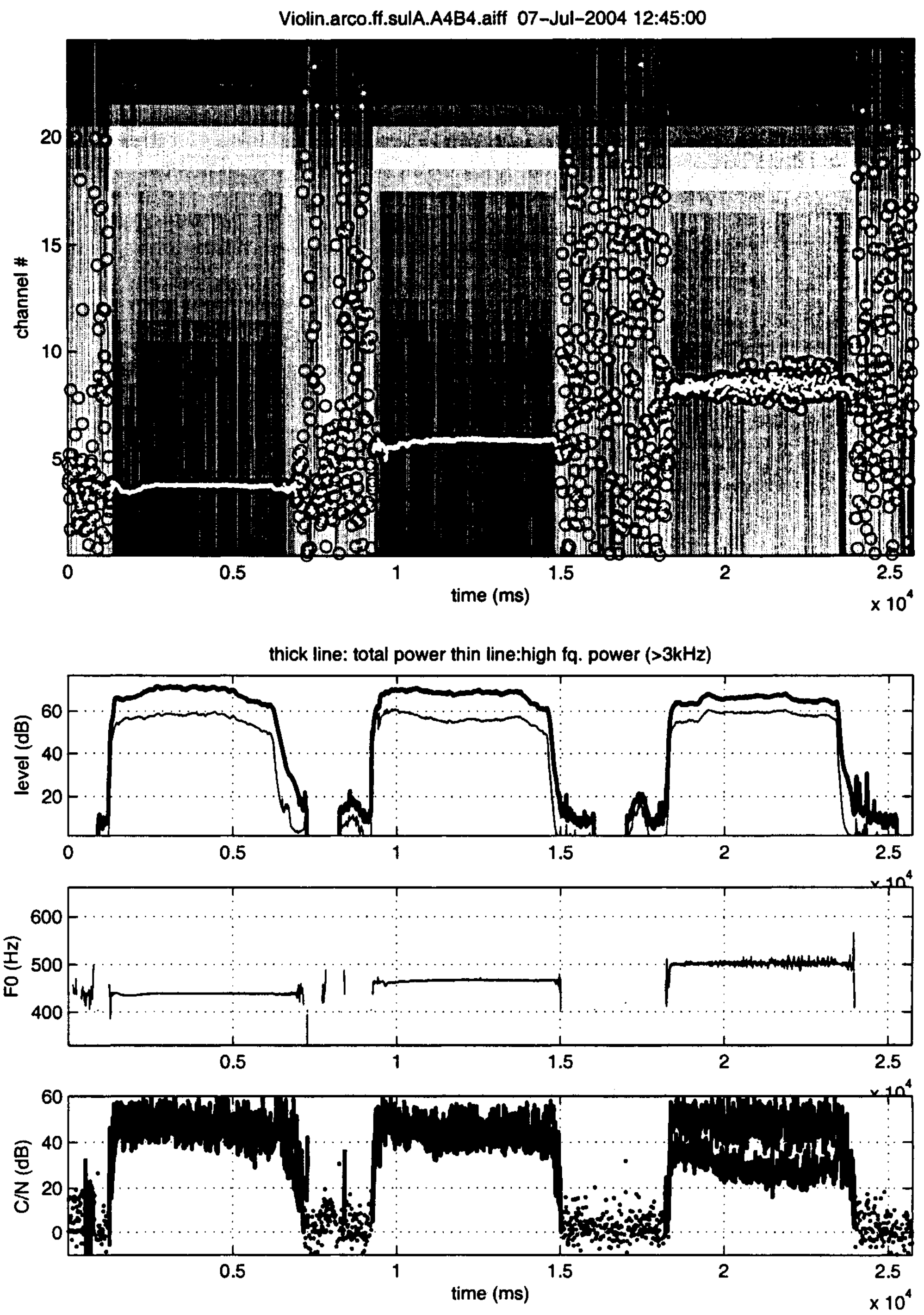

Figure 19: Recording for Test 5.

Table 22: Instrument identification results for two methods

		Method 1		Method 2
		Ranking by highest inner-products		Winner by backwards char.
Test 1	Note 1	0.602 Oboe ×		
		0.589 Clarinet		
		0.582 Trumpet		
		0.264 Violin		
	Note 2	4.163 Clarinet ✓		
		2.014 Trumpet		
		1.806 Oboe		
		1.300 Violin		
	Note 3	1.047 Clarinet ✓		
		0.560 Oboe		
		0.499 Trumpet		
		0.410 Violin		
	Note 4	0.115 Clarinet ×		
		0.108 Trumpet		
		0.091 Oboe		
		0.086 Violin		
	Note 5	0.654 Trumpet ×	F/R	0.646 Oboe ✓
		0.646 Oboe	F/G	0.123
		0.530 Clarinet	B/G	0.532
		0.445 Violin	B/R	0.001
	Note 6	0.551 Clarinet ×		
		0.327 Oboe		
		0.288 Trumpet		
		0.226 Violin		

Instrument identification results for two methods (cont.)

Test 1 (cont.)	Note 7	0.150 Clarinet ×	F/R	0.041
		0.148 Oboe	F/G	0.014
		0.144 Trumpet	B/G	0.144 Trumpet ✓
		0.125 Violin	B/R	0.040
	Note 8	1.327 Violin ×		
		0.928 Clarinet		
		0.858 Trumpet		
		0.790 Oboe		
	Note 9	0.483 Clarinet ×		
		0.342 Oboe		
		0.221 Trumpet		
		0.099 Violin		
	Note 10	0.608 Clarinet ×		
		0.285 Trumpet		
		0.284 Oboe		
		0.213 Violin		
	Note 11	1.000 Clarinet ×	F/R	0.509
		0.556 Oboe	F/G	0.355
		0.527 Trumpet	B/G	0.556 Oboe ×
		0.424 Violin	B/R	0.536
	Note 12	0.488 Clarinet ×	F/R	0.278 Oboe ×
		0.278 Oboe	F/G	0.134
		0.265 Trumpet	B/G	0.229
		0.218 Violin	B/R	0.186

Instrument identification results for two methods (cont.)

Test 2	Note 1	1.852 Oboe ✓		
		1.407 Clarinet		
		0.907 Trumpet		
		0.275 Violin		
	Note 2	1.780 Trumpet ×	F/R	1.488
		1.745 Oboe	F/G	0.367
		1.395 Violin	B/G	1.780 Trumpet ×
		1.284 Clarinet	B/R	0.307
Test 3	Note 1	3.791 Clarinet ×		
		3.095 Oboe		
		1.335 Violin		
		1.230 Trumpet		
	Note 2	2.560 Clarinet ×		
		2.114 Violin		
		1.442 Trumpet		
		1.431 Oboe		
	Note 3	2.753 Clarinet ×		
		2.245 Trumpet		
		1.182 Oboe		
		0.992 Violin		
	Note 4	2.215 Oboe ✓		
		1.682 Trumpet		
		1.634 Violin		
		1.587 Clarinet		

Instrument identification results for two methods (cont.)

Test 3 (cont.)	Note 5	2.316 Clarinet ×	F/R	1.219 Oboe ✓
		1.249 Violin	F/G	0.475
		1.219 Oboe	B/G	0.466
		1.011 Trumpet	B/R	0.097
	Note 6	3.373 Oboe ✓	F/R	0.938 Trumpet ×
		2.071 Violin	F/G	0.487
		1.115 Clarinet	B/G	0.140
		0.938 Trumpet	B/R	0.109
	Note 7	2.486 Oboe ✓	F/R	1.089
		1.806 Clarinet	F/G	0.365
		1.508 Trumpet	B/G	1.508 Trumpet ×
		1.500 Violin	B/R	0.303
	Note 8	0.918 Oboe ✓		
		0.771 Trumpet		
		0.699 Violin		
		0.620 Clarinet		
	Note 9	1.105 Violin ×	F/R	1.060 Oboe ✓
		1.060 Oboe	F/G	0.343
		0.892 Trumpet	B/G	0.793
		0.748 Clarinet	B/R	0.753
	Note 10	3.638 Oboe ✓		
		3.154 Trumpet		
		2.836 Clarinet		
		2.119 Violin		

Instrument identification results for two methods (cont.)

Test 3 (cont.)	Note 11	3.895 Trumpet ×		
		3.808 Oboe		
		3.315 Violin		
		2.247 Clarinet		
	Note 12	1.754 Clarinet ×	F/R	1.351
		1.681 Oboe	F/G	0.394
		1.160 Trumpet	B/G	1.754 Clarinet ×
		1.000 Violin	B/R	0.010
Test 4	Note 1	3.568 Clarinet ×	F/R	0.100
		1.912 Oboe	F/G	0.073
		1.668 Trumpet	B/G	1.092 Violin ✓
		1.092 Violin	B/R	0.515
	Note 2	2.177 Clarinet ×	F/R	0.169
		1.297 Oboe	F/G	0.103
		1.291 Trumpet	B/G	1.291 Trumpet ×
		0.687 Violin	B/R	0.805
	Note 3	0.663 Violin ✓	F/R	0.576
		0.612 Oboe	F/G	0.035
		0.513 Trumpet	B/G	0.663 Violin ✓
		0.488 Clarinet	B/R	0.197

Instrument identification results for two methods (cont.)

Test 5	Note 1	5.499 Clarinet ×		
		2.712 Oboe		
		1.994 Trumpet		
		1.864 Violin		
	Note 2	8.562 Clarinet ×		
		4.625 Violin		
		3.972 Trumpet		
		3.824 Oboe		
	Note 3	2.538 Trumpet ×	F/R	2.524
		2.526 Violin	F/G	0.372
		2.248 Oboe	B/G	2.538 Trumpet ×
		1.858 Clarinet	B/R	1.291

Appendix D

Matlab Programs

```
function c = coeff_sep04(a,eta,flow,fup,signal,inst,smplrt,bckwds,ghost)

%function c = coeff_sep04(a,eta,flow,fup,signal,inst,smplrt,bckwds,ghost)
%
%Returns a matrix of coefficients where each row corresponds to an interval
%in the partition, "a".
%For each interval, coefficients are calculated by taking the integral
%of the signal with the wavelets, U_{j,k}, where k is a function of the
%frequency, ranging from flow to fup and the size of the interval.
%Input item "inst" must be a three letter abbreviation for the wavelet
%desired: "vio" "obo" "cla" or "tpt". Setting "bckwds" = 1 will reverse
%the signal before computation. Input "smplrt" allows us to perform
%the transformation for music signals recorded at different sample rates,
%eg. 44100 or 22050 samples per second. "ghost" will reverse the attacks
%for the decays in each window, causing the 'ghost' windows to become used
%for producing the inner products.

if inst == 'tpt'
    har = log2([420 440 450 300 160 120 100 90  80  35  35  12]);
    asyn = [0 0.005 0.005 0.02 0.015 0.04
                0 0.01 0 0 0 0];
```

```
    s = [1 1 1 1 1 1
                1 1 2 2 2 1];
elseif inst == 'vio'
    har = log2([80  180 70  350 250 50  110 140 30  25  60  40]);
    asyn =[0 0 0.015 0 0 0
                0   0.09   0.09   0.155   0.13   0.16];
    s = [1  1   1   1   1   1
                1   1   1   1   1   1];
elseif inst == 'obo'
    har = [430  310 480 460 575 110 35  14  65  295 170 110];
    asyn = [0 0 0 0 0 0
                0   0.03   0.05   0   0.06   0.05];
    s = [1  2   2   2   2   2
                1   1   1   2   1   1];
elseif inst == 'cla'
    har = ([700 25  220 28  460 28  105 50  150 33  55  45]);
    asyn = [0   0   0   0   0   0
                0   0   0   0.01   0.02   0.03];
    s = [1  2   2   2   2   2
                1   1   2   1   1   1];
end %Variables to create the Super Malvar Wavelets for "inst"

len_a = length(a);

if bckwds == 1 %Backwards case
    sig = zeros(1,length(signal));
    for m = 0:len_a-1 %Reverse vectors: eta, flow and fup, plus invert a
        abckwds(m+1) = floor(1000*length(signal)/smplrt)/1000 - a(len_a-m);
        etabckwds(m+1) = eta(len_a-m);
    end
    for m = 1:len_a-1
        flowbckwds(m) = flow(len_a-m);
        fupbckwds(m) = fup(len_a-m);
```

```
    end
    for n = 1:length(signal) %Reverse the signal
        sig(n) = signal(length(signal)-n+1);
    end
    a = abckwds;
    eta = etabckwds;
    flow = flowbckwds;
    fup = fupbckwds;
    signal = sig;
end

masyn = [max(asyn(1,:)) max(asyn(2,:))]';
    %Vector containing the largest asynchronicities

for j = 1:len_a - 1
    I(j) = a(j+1)-a(j); %Vector containing the sizes of the intervals
end %Calculations to make matrices of coefficients in the transformation
    c = zeros(len_a-1,max(fup));
        %Prepare the output matrix

for j = 1:len_a-1    %All intervals created by the partition
    sus = a(j+1) - a(j) - eta(j+1) - eta(j); %Sustain length
    dur = floor(smplrt*(sus + 2*eta(j+1) + 2*eta(j) + masyn(1) + masyn(2)));
        %Duration of wavelet in number of samples
    wv=superMalvar_sep04(har,flow(j),fup(j),eta(j),eta(j+1),asyn(1,:),
                                 asyn(2,:),s(1,:),s(2,:),sus,smplrt,ghost);
        %Compute wavelet: SMW for interval j with k ranging from flow(j) to
        %fup(j)
    comm = max([floor(smplrt*(a(j) - eta(j) - masyn(1))) 0]);
        %Beginning of support of wavelet for interval j
    rightend = min([(comm+1+dur) length(signal)]);
        %End of support of wavelet
    for f = flow(j):fup(j)    %All frequencies between flow(j) and fup(j)
```

```
        integral = signal(comm+1:rightend)*wv(f-flow(j)+1,1:rightend-comm)';
        %Compute inner product between signal and wavelet with frequency f
        if f ~= 0
            if bckwds == 0 %Forwards
                c(j,f) = integral;
            else           %Backwards
                c(len_a-j,f) = integral;
            end %Forward/backward differentiation
        end %Avoid the frequency f = 0 case
    end %All frequencies
end %All intervals
```

```
function s = superMalvar_sep04(har, flow, fup, eta_a, eta_d, c_n,
                           d_n, phi_n_a, phi_n_d, sus, smplrt, ghost)

%s = superMalvar_sep04(har, flow, fup, eta_a, eta_d, c_n, d_n,
                                 phi_n_a, phi_n_d, sus, smplrt, ghost)
%
%Create super Malvar wavelets, U_{j,k}, with F_0 ranging from flow to
%fup. Generates a matrix with row f corresponding to the signal with
%samplerate samples per second which is given by the super Malvar wavelet
%U_{j,k} with F_0 = f. The other inputs determine the characteristics of
%the instrument being modelled, such as harmonic amplitudes; alph_a,
%alph_d, c, d and sus should be given in seconds and not the number of
%samples.

if ghost == 1 %use ghost windows
    asyn_temp = c_n; %Switch non-synchronicities
    c_n = d_n;
    d_n = asyn_temp;
    phi_temp = phi_n_a; %Switch phi(n), suddenness characteristics
    phi_n_a = phi_n_d;
    phi_n_d = phi_temp;
end

h = length(har);%Number of harmonics modelled
p = floor((h + 1)/2);%Number of harmonic amplitude windows required
c_max = max(c_n);%Highest attack asynchronicity
d_max = max(d_n);%Highest decay asynchronicity
attack = 2*eta_a;%Duration of the attack period
decay = 2*eta_d;%Duration of the decay period
dur = floor(smplrt*(sus + attack + decay + c_max + d_max));
%Duration of the interval of support for the wavelets
s = zeros(fup-flow+1, dur + 1);%Prepare output
if c_max + d_max > sus
```

```
    return
end %Ensure that condition on the asynchronicities is satisfied

t = 0 : dur;%Time variable as number of samples
w = zeros(p, dur + 1);%Matrix to represent window functions
I = floor(smplrt*(sus + eta_a + eta_d));%|I_j| (in samples)
f = zeros(h, dur + 1);
    %Matrix with a row for each Malvar wavelet. SMW is the sum of the rows.
a_a = floor(smplrt*(c_max + eta_a));%a_j
a_d = dur - floor(smplrt*(d_max + eta_d));%a_{j+1}

for n = 1 : p
    comm = a_a - floor(smplrt*(c_n(n) + eta_a));%Start of window
    dur_a = floor(smplrt*(attack + 2*c_n(n)));%Duration of attack
    arg = zeros(1, dur_a + 1);%Prepare the attack window vector
    arg(1, 1:dur_a + 1) = pi/2.*((sin(pi/2.*((t(1:dur_a + 1))/dur_a))).^2);
    if phi_n_a(n) > 1 %Number of iterations of (pi/2)sin^2()
        for j = 2:phi_n_a(n)
            arg = pi/2.*((sin(arg)).^2);
        end %Iterate that number of (pi/2)sin^2()
    end
    w_a = sin(arg); %Attack portion of window
    w(n,comm+1:comm+1+ dur_a) = w_a(1, 1: dur_a + 1); %Attack

    term = a_d + floor(smplrt*(eta_d + d_n(n))); %End of window
    dur_d = floor(smplrt*(decay + 2*d_n(n))); %Duration of decay
    arg = zeros(1, dur_d + 1); %Prepare the decay window vector
    arg(1, 1:dur_d + 1) = pi/2.*((sin(pi/2.*(1 - (t(1:dur_d + 1))
                                                                /dur_d))).^2);
    if phi_n_d(n) > 1 %Number of iterations of (pi/2)sin^2()
        for j = 2:phi_n_d(n)
            arg = pi/2.*((sin(arg)).^2);
        end %Iterate that number of (pi/2)sin^2()
```

```
    end
    w_d = sin(arg); %Decay portion of window
    w(n, term - dur_d + 1:term + 1) = w_d(1: dur_d + 1); %Decay
    w(n, comm + dur_a + 2 : term - dur_d) = 1; %Sustain portion
end %Create fundamental and harmonic windows

for o = flow:fup
    k = round(o*(2*I/smplrt)); %Value of k to give correct F_0 for wavelet
    for n = 1 : p %Each window
        q = floor(log2(h/(2*n - 1)));
            %Number of harmonics to share a window
        for m = 0 : q
            f((2*n-1)*2^m,1:dur+1) = w(n,:).*sin((pi/I)*
             (2*n-1)*(2^m)*(k)*(t-a_a));
        end %Terms in SMW that share the same harmonic window
    end %Terms in SMW
    s(o-flow+1,:) = sqrt(2/I).*sum(har*f,1)/norm(har,2);
        %Output SMW in row corresponding to the F_0 ranging from flow to fup
end %Frequency ranging from flow to fup
```

Bibliography

[1] S. Mallat, *A Wavelet Tour of Signal Processing, Second Edition*, Academic Press, London, 1999.

[2] H. F. Davis, *Fourier Series and Orthogonal Functions*, Allyn and Bacon, Inc., Boston, 1963.

[3] R. R. Coifman and M. V. Wickerhauser, *Entropy-based algorithms for best basis selection*, IEEE Transactions on Information Theory, **38**, 1992, pages 713–718.

[4] D. L. Donoho and X. Huo, *Uncertainty principles and ideal atomic decomposition*, IEEE Transactions on Information Theory, **47**, 2001, pages 2845–2862.

[5] D. L. Donoho and M. Elad, *Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization*, Proc. of the National Academy of Sciences, **100**, No. 5, 2003, pages 2197–2202.

[6] S. Kadambe, P. Srinivasan, B. Telfer and H. Szu, *Representation and classification of unvoiced sounds using adaptive wavelets*, Proceedings of the International Society for Optical Engineering, **1961**, 1993, pages 325–335.

[7] S. G. Mallat and Z. Zhang, *Matching pursuits with time-frequency dictionaries*, IEEE Transactions on Signal Processing, **41** No. 12, 1993, pages 3397–3415.

[8] W. J. Pielemeier, G. H. Wakefield and M. H. Simoni, *Time-frequency analysis of musical signals*, in Proceedings of the IEEE, Vol. 84, No. 9, 1996, pages 1216–1230.

[9] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia PA, USA, 1992.

[10] M. Unser, *Sampling—50 years after Shannon*, Proceedings of the IEEE, Vol. 88, No. 4, 2000, pages 569–587.

[11] P. McKay, *The first law of harmony*, in The Enchanted Ear Series, The Ottawa Citizen, Monday, October 28, 2002.

[12] B. Fink, *Selected Essays and Readings on The Origin of Music*, Greenwich, 2003.

[13] Y. Meyer, *Les ondelettes: algorithmes et applications*, Armand Colin, Paris, 1992.

[14] C. K. Chui, *An Introduction to Wavelets*, Academic Press, Inc., San Diego CA, 1992.

[15] K. Kennan and D. Grantham, *The Technique of Orchestration, Fifth Edition*, Simon and Schuster, Upper Saddle River NJ, 1997.

[16] S. Isacoff, *Temperament: The Idea that Solved Music's Greatest Riddle*, Random House Publishing, November, 2001.

[17] B. B. Hubbard, *The World According to Wavelets*, AK Peters, Ltd., Wellesley MA, 1996.

[18] D. P. W. Ellis and B. L. Vercoe, *A wavelet-based sinusoid model of sound for auditory signal separation*, International Computer Music Conference, McGill, 1991.

[19] C. A. Elliott, *Attacks and releases as factors in instrument identification*, Music Education J. **23**, No. 1, Spring 1975, pages 35–40.

[20] J. A. Moorer, J. M. Grey, and J. Strawn, *Lexicon of analyzed tones, Part 1: Violin*, Computer Music J., **1**. No. 2, Spring 1977, pages 39–45.

[21] J. A. Moorer, J. M. Grey, and J. Strawn, *Lexicon of analyzed tones, Part 2: Clarinet and Oboe*, Computer Music J., **1**. No. 3, Summer 1977, pages 12–29.

[22] J. A. Moorer, J. M. Grey, and J. Strawn, *Lexicon of analyzed tones, Part 3: The trumpet*, Computer Music J., **2**. No. 2, Spring 1978, pages 23–31.

[23] J.-C. Risset, *Computing Musical Sound* Chap. 13 in Mathematics and Music, A Diderot Mathematical Forum, G. Assayag, H.G. Feichtinger, and J.F. Rodrigues, (eds.), pages 215–231, Springer-Verlag Berlin Heidelberg, 2002.

[24] E. Neuwirth, *The Mathematics of Tuning Musical Instruments—A Simple Toolkit for Experiments* Chap. 14 in Mathematics and Music, A Diderot Mathematical Forum, G. Assayag, H.G. Feichtinger, and J.F. Rodrigues, (eds.), pages 233–241, Springer-Verlag Berlin Heidelberg, 2002.

[25] E. Hernández, and G. Weiss, *A First Course on Wavelets*, CRC Press, 1996.

[26] H. Kawahara *GUI-STRAIGHT ver.23, ATR Human Information Processing Research Laboratories, Faculty of Systems Engineering, Wakayama University, CREST*, March 22, 1998.*