Automatic Repair of Physical Flaws in Recorded Music

A thesis submitted in partial fulfillment of the
requirements or the degree of Bachelor of Science in Physics
at the College of William and Mary

by

Jori S. Byrne-Diakun

Williamsburg, Virginia
April 2008

Abstract:

Recording music in any form often produces artifacts that are not pleasing to the ear. Reducing these artifacts is the focus of millions of dollars in research for music companies every year. However, despite this research, there is still a considerable volume of work that requires special attention, and historic non-commercial recordings are often never cleaned up enough to be sold, or even provided, to the general public because of the great amounts labor each recording demands. This paper describes the beginnings of a search for an automated process to clean up the damage to this piece of human history.

Automatic Repair of Physical Flaws in Recorded Music

Table of Contents

1. Introduction

Our project's long term goal is the production of an automated program that will clean up degraded music signals without human oversight until the final steps. This research is important to both the worlds of music and science and has a wide range of applications, scientific and otherwise. Music is an integral part of our society, and of our history as a culture. It is one of the signs that a culture has progressed to what we in the modern world term as "civilized."

This project is primarily aims to provide a practical means to preserve these priceless and unique works of art which have degraded due to age. In the National Archive today, thousands of old magnetic tape recordings are degrading in storage and there are simply not enough people or resources to save them. With the advent of an automated method of cleaning these works to listenable quality, we can save a priceless collection and return it to the public. In the world of economics, this technology would be very useful in two areas in particular. Firstly, this technology would open up a sizable quantity of refurbished music to the market. Secondly, this technology would be useful to help avoid the common problem of signal degradation, which occurs with the over replication and conversion of sound files.

This thesis details the first steps in developing this technology: the qualitative musical analysis and the conversion of that musical knowledge into tangible results through the manual manipulation of the signal. This thesis has two main sections. The first is the qualitative musical analysis required to approach this project, which includes the *Background* and *Musicology* sections. The second is the beginning of the signal theory section, which includes the *Basic Theory*, *Spectrographic Analysis*, *Wavelet*

Analysis and De-Noising Signals, and *Filters and De-Noising Signals* sections. The conclusion section is used both to present findings and to provide an outline for the next stage of the research required to produce the desired project aims.

2. Background

In 1939, the University of Wisconsin took Gunnar Johansen, a piano virtuoso of the highest order, as the first Artist-in Residence in the United States. Mr. Johansen was a collector of musical oddities, both unique and beautiful, including an amazing collection of pianos. During his tenure, he owned his own label called Artist Direct, for which he had recording equipment and a recording studio in the basement of his Blue Mounds home. The recording equipment was housed in his basement, near a waterfall that ran though the property, as were the recordings themselves. Mr. Johansen produced a huge volume of recorded works, ranging from works by Johann Sebastian Bach to the complete works of the romantic piano phenomenon, Franz Liszt. Mr. Stevens, an engineer in the NDE Lab at William and Mary, came into the possession of these magnetic tape recordings upon Mr. Johansen's death in 1991. Magnetic tape is susceptible to environmental damage and the recordings had degraded before the music was digitalized, meaning that the works required substantial cleaning up.

The most recent phase of the process of preserving this timeless piece of artistic history is cleaning them up to make them worth listening to. The recordings can be cleaned using traditional programs by having persons who know the repertoire listen to the recordings and manually clean the recordings. However, due to the sheer volume of work, a more efficient way of cleaning the recordings would save many recordings that might otherwise be lost or never attended to. The purpose of this research is to find a way to process the music using computers. Computers, however, cannot evaluate a great work of art, nor understand artistry, but they can analyze mathematical patterns, which humans can provide by understanding the music-math link. This project is a first step towards developing that link.

3. Musicology

Music is one of the single most complex entities in existence. The living breathing nature of a score is unmatched in complexity even by the sweeping beauty of a painting or the stationary brilliance of a sculpture. Every time a performer plays a work of music, the music changes, evolves, and comes to life with the performer's own experiences. As such, every performer produces a work that is different than any other performance and the original score it came from. It would be easy to handle any one musical work if it were always static, or even if it were static in the way that it evolved over time. Music defies this simplicity. Furthermore, it is this very unpredictability that makes music so appealing to listeners. A "perfect" performance, exemplified by a static score played by a computer, produces a negative reaction from listeners to the extent that they search out living performers. Even more interestingly, listeners, even with no musical background, can listen to the same piece played by two different performers and have a completely different aesthetic response the work even if they have no idea why it happens.

In order to begin analyzing the physics of cleaning up music, it is important to look at and understand the music with an educated ear. Only then can the music be evaluated and edited in an aesthetic and beneficial way. There were several musical issues, which were likely to become physics issues, which were evident from the onset of the project. These needed to be considered during musical analysis to prevent them from creating issues that would interfere with the physics later. One of the major issues we were immediately aware of was the concept of rubato. Because rubato is used consistently by most soloists to shape phrases, no performer will produce music that conforms exactly to the score. This is not a serious issue if a human is interactively analyzing of the signal, particularly if the analyst is a musician. However, because a computer is not good at handling artistic concepts, it was clear that there would be a problem teaching the computer to understand an artist's rubato not as musical flaw, but rather an acceptable variance in the signal. The second issue we needed to address musically was base pitch variance. Most instruments are not perfectly tuned. Their intonation is often dictated by the intonation of a base pitch, most commonly the A above middle C or the B-flat above middle C. The A above middle C is defined as having a frequency of 440Hz, but most performers, again, do not tune exactly to A-440 before recording, and this can make analysis of the recording tricky for a computer if the computer is looking discrepancies from expected pitches. The third major area of concern we had, which was not applicable to the current work, though it would be applicable to general application of this research, and that is the practice of vibrato. The signal variances introduced by vibrato could be mistaken for signal distortion by a computer. All of these areas of musical practice needed to be accounted for and understood before we could proceed to the more quantitative analysis of the music.

We began the process of analysis by selecting a test performance that had been subjected to a great deal of damage, both during and after recording, to work with. We chose this work from Gunnar Johansen's piano recordings of the complete works of Franz Liszt, selecting "Polonaise No. 2 in E Major" as our test score. Before even beginning to work with the damaged recording, we familiarized ourselves with the score and multiple different performances of the work [1-9]. We began our research using the score and six of the eight recordings. The last two recordings were brought in at a later date to clarify a particular passage that was unclear in most recordings we possessed.

After learning the score well, we began to listen to the entire damaged work, as performed by Gunnar Johansen. The purpose of this preliminary listening was two-fold. First, we were looking to analyze the types of flaws that were audible in the recording so that we would know what types of repair methods we would need to use to approach cleaning up the recording. Second, we were interested in the idiosyncrasies that were unique to Johansen's playing style because those characteristics cannot be mathematically analyzed when looking purely at a score. After a period of about two week of listening to the recording, we began to dismantle the signal in MATLAB, looking at smaller sections of the music. Once the signal was spliced into small chunks, we had the option of listening to a more focused section of the sound file at full and reduced speeds, using built in audio software. This allowed us to find and catalog particular areas of interest which had artifacts representative of those that we wanted to eliminate.

For the purposes of the musicological section of our analysis, we parsed the score down and identified four main sections of interest which represented several different artifacts we wanted to analyze. The first section, which is represented in Figure 1 as a spectrogram, was denoted "Excerpt1." Excerpt1 was interesting because it provides two principal types of artifacts, not including the noise which was a flaw that was present in all sections of the Johansen recording. The first was an overload spike, which was likely produced during the recording of the piece. In most modern recording studios, there is a person who monitors the input signal during the recording process. Johansen recorded his music by himself, and since he was playing the piano, he could not have adjusted settings when his music varied in intensity. This means that intense moments, such as a forzondo at the top of a run or other sudden burst of sound could have overloaded the signal processing ranges of the recording device causing the signal to not clearly record. Another possible explanation for the overload sound is a post-recording flaw such as a scratch on the magnetic tape on which the recording was made. There was also a blurring effect in the sound near the end of the secondary run. We believe that this may have occurred during recording due to the combination of the rapidity of the run and the natural reverb of the piano at lower frequencies.

The second section we focused on is represented as a spectrogram in Figure 2. It is labeled “Excerpt2” and it was principally interesting because of two click events. The first is only semi-audible at full speed playback. The second is fully audible, though for most people we tested the click passes by without notice because it occurs in such a short time frame. The reason that we call the second click “fully audible” is because it does produce a general impression of a disjoint and crackly quality to the music, even if the event itself is not immediately audible to all listeners. These clicks are most like the result of post-recording damage to the magnetic tape, probably a stretch or scratch.

The third section that was analyzed is labeled as “Excerpt3” and is shown in Figure 3 as a spectrogram. The principal artifact in this excerpt was an overtone that was present in the entire excerpt. We were originally considering the overtone as a post-recording event due to a scratch on the tape. However, we have shifted our opinion with more recent work, and we now believe that the overtone was probably a part of the overtone series of the chords being played that got magnified in intensity by the recording device.

The final section we focused on was an ideal example of noise damage. This section, represented in Figure 4, is labeled as "Excerpt4." The signal was garbled by external sounds in the entire recording. We slowed the recording down and could hear some sounds that were local to the environment of the recording. One such sound we identified was most likely the sound of Gunnar Johansen's fingers on the keyboard while he was playing. There were also other ambient sounds present which may have been sounds from the rest of the house where the music was being recorded.

Once we finished the qualitative musical analysis of the excerpts, we moved in the realm of physics to begin looking at frequency-time representations of the signal to quantitatively locate the damage to the signal.

4. Basic Theory

The theory surrounding signal processing is fairly well defined, though it is not always easily applied. There are many different representations of frequency-time relations that have been explored by physicists and mathematicians. However, we focused primarily on two types of analysis: Fourier Analysis [10] and Continuous Wavelet Analysis [11].

Fourier Analysis and Spectrograms:

For the purposes of signal analysis we are principally interested in the FFT, or Fast-Fourier Transform, which is defined in its most general below.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2i\pi}{N}nk} \qquad \text{where: } k = 0,\ldots,N-1. \tag{1}$$

In order to separate the frequency and time components in a typical signal, we first reduce the signal matrix down to a vector and then apply a FFT to the signal. It is important to note that FFT's are limited because they only will yield meaningful results when the signal vector's length is a power of 2. The spectrogram function then applies the sample rate, hanning window, overlap, and other specified parameters to the FFT in order to produce a spectrographic plot of the signal. In MATLAB, the output figure is oriented with the normalized frequency on the horizontal axis and the time on the vertical axis. Conventionally, spectrograms are displayed with the frequency on the vertical axis and the time on the horizontal axis, so we often need to reorient the axes.

Continuous Wavelet Analysis:

For the purposes of our signal analysis, MATLAB handled the majority of the computations, but in order to readjust the settings, we needed to have a basic understanding of the manipulations of Continuous Wavelet Analysis. Below is the general form of the mother wavelet (2) and the general form of the transform(3).

$$m_k = \int t^k \psi(t) dt \tag{2}$$

$$X_{\omega}(a,b) = \frac{1}{\sqrt{|(b)|}} \int_{-\infty}^{\infty} x(t) \psi(\frac{t-a}{b}) dt \tag{3}$$

The transform can be calculated using the FFT. The principal purpose of wavelets as we have used them is to filter out noise from signals.

5. Spectrographic Analysis

The purpose of spectrographic analysis is to provide a clear visual presentation of the information carried in a signal. The basic key to reading a spectrogram is relatively straight forward [11]. The colors represent intensity. In the case of the spectrograms in Appendix A, red represents the most intensity and blue represents the minimal intensity (following the visible light spectrum). The horizontal slashes represent individual pitches, which can be converted from pure pitch to note values by using Figure 5. Smears can represent extended pitches or blurring. Vertical bars are noise lines, which can be caused by the overtone scale or ambient noise. Differentiating between the two is best done by listening, but a computer can understand if programmed because the expected overtone series progresses regularly.

When we used spectrographic analysis, we were looking to find the four primary types of artifacts heard in our excerpts and also to observe the effects of noise. The first artifact was blurring, which were heard in Excerpt 1. This occurs when an older recording device cannot keep up with the speed of the performer and it sounds like the notes in runs are smeared. Figure 1 shows this artifact as it showed up in Polonaise No. 2 in E Major. The second artifact was an overload or feedback artifact, which was also heard in Excerpt 1. These events occur when the performer plays to loud for the recording device to handle, and thus the pitch is lost in a burst of noise. In Figure 1, this can be seen in the first box. The red pitch at the top of the run is the actual note, but the frequencies around that pitch are also high intensity. This means that the recording system recorded noise around the high intensity pitch. The combination of these two effects, the high intensity pitch and the noise, produces an overload artifact. The third artifact we were looking for occurred in Excerpt 2. Cracks and shifts, or stretches, are post recording events. When a tape rips, stretches, or loses data due to other external damage, the sound that is heard when the music is rendered is damaged as a result, and this can usually qualified by a blip in the sound or a stretching of tempo that is not taken by the performer. In Figure 2, we can see a series of cracks, only two of which are audible. The last major artifact we were looking for was an overtone or extra pitch group. These events can occur from scratches on the tape or from reverb during the recording, it is event specific. We can only detect overtones and extra pitches by using Figure 5 and a score, automated or not. In Excerpt 3, we heard an overtone, which we found in Figure 3 (see Figure 3 caption).

Noise was treated separately because it is present in all of the recordings. Some noise can be expected in all recordings. If there is no noise, often a signal can seem too perfect. When a listener goes to a concert hall, there is always ambient noise. The purpose of a recording is to recreate the experience sans the need to go to a concert hall. As such, some noise is important to provide an auditory guard from pure tonality, which is unnerving to listen to. However, too much noise detracts from the experience of listening to music because the pitches become ambiguous. After multiple tests, we found that while spectrograms are very useful for finding artifacts, dealing with noise can be more efficiently handled by wavelets.

6. Wavelet Analysis and De-noising Signals

Wavelet Analysis is an alternative representation of the frequency-time relation of a signal. We were particularly interested in using one-dimensional wavelets to clean up the noise in the Johansen recording. For our test we worked with Excerpt 1, for one primary reason: due to the variable harmonic patterns in the excerpt, a filter would not be an effective tool for cleaning out noise. Any filter, excluding a variable filter, we could design would either miss noise or filter out music as well. A filter would, however, be extremely effective for Excerpt 4, where there is more harmonic stability.

MATLAB is an effective tool for wavelet analysis because it has a built in toolbox for working with wavelets and in particular SWT De-noising. SWT De-noising uses Stein's Unbiased Estimate of Risk in order to formulate a threshold for de-noising. The formulation for the threshold is shown by the equation below, where T is a fixed form threshold, or in other words a fixed form “one dimensional de-noising orientation function” [12], and n is the fixed signal length.

$$T = \sqrt{2\log_e(n\log_2(n))} \qquad (4)$$

The SWT De-noising tool box was our primary experimental design station for working with noise and wavelets (see Figure 6). We were working with a modified version of Excerpt 1, in which the length parameters had been adjusted to fit the wavelet specifications. The new signal is detailed in Figure 7. Our first attempts to de-noise Excerpt 1 using SWT de-noising were complete failures. Figure 8 shows our spectrogram after the first attempt, in which all non-decimated detail coefficients wavelets were killed. The noise was effectively cleaned out, but we also cleaned out the majority of the treble signal as well. The music itself sounds very clear of noise; the only draw back it that it is very free of melody as well. The next logical step to try to correct this error was to only partially kill the coefficients, which also ended with a poor resultant signal. The new signal, Figure 9, had a huge new blur in the middle of the spectrogram. Furthermore, all of the individual pitches had a sound quality that more resembled a honky-tonk than a pianoforte.

In order to try to get a cleaner signal we needed to separate the functions of the detail coefficients. We started by killing d5 while leaving the other coefficients alone. The resultant signal is shown in Figure 10. The d5 coefficient, in terms of noise, seemed to represent the non-auditory range of frequencies. The spectrogram for d5 was de-noised in the highest frequency regions, but little change was observed in the lower frequency regions. Since we have less interest in the non-auditory realm than the auditory realm, we moved on to kill d4. This produced a signal that was noisier than the d5 iteration and blurrier in the auditory realm, which was immediately evident when we listened to the new signal. Figure 11, the spectrogram for the new signal where d4 was killed, shows that the d4 iteration failed to produce the desired results. The next coefficient to kill was d3, which resulted in a distorted signal. Analyzing Figure 12, we found that d3 served no useful purpose in reducing the noise of the signal, so we discarded that iteration as well.

When we listened to the signal where we killed d2, shown in Figure 13, we finally heard the first real noise-level reduction in our wavelet trials. The noise was reduced in the auditory region without serious disruption of the music signal. Furthermore, the music was clearer than the original recording for the first time. In order to finish testing of the effects of the coefficients, we continued with the progression and killed the d1 coefficient. This produced the cleanest signal of all the iterations. The sound quality was noticeably better than that of the original signal. However, once the noise was removed, there were a few new artifacts in the signal that immediately became noticeable to the naked ear. We are not sure if these artifacts were produced by the noise removal process or were imbed in the signal already and could not be heard until the noise was removed, further research is required.

After the independent effectiveness of the iterations killing d1 and d2, the next logical step was to kill them both at the same time, as in Figure 15. This, surprisingly, was not effective in cleaning up the signal. The signal resulting from killing both d1 and d2 was noisier than the independent killing of either coefficient. Our final attempt to make the d1-d2 combinations work in conjunction ended in failure as well. We attempted to kill only d1 and then simply minimize d2 instead of killing it. Though the resultant signal was better than the original signal (see Figure 16) on whole, the cleaned signal was not as sharp as the iteration killing d1.

From killing d1 alone we were able to produce a cleaned signal for Excerpt 1. However, there is further work with this signal that will need to be done in order to completely clean it up, as will be detailed in the *Conclusion* section.

7. Conclusions

We did successfully clean up one of the signal sections. Wavelet de-noising was fairly effective in cleaning up the signal and shows promise for future research as a method of de-noising music signals. However, as noted in the wavelet analysis section, cleaning out the noise with wavelets is either uncovering more artifacts or it is producing some new artifacts while de-noising the signal. The possible advantages of wavelet fingerprints over spectrograms are worth further investigation in later research.

We began work recently using filters to clean noise and extra pitches out of the music, using Excerpt 4. Despite the failures of our initial attempts to use a variable filter to clean up the signal, the system shows promise if it can be streamlined more extensively. Filter methods should also be effective in cleaning up stray pitches and overtones, as in Excerpt 3. A variable band-pass filter may be able to clear up the signal. It would be worth investigating the possibility of linking a variable band-pass filter to a digital map of the score. Furthermore, we could eliminate the problems of vibrato and tuning, mentioned in the *Musicology* section, by providing f_{3db} such that some fluctuation is allowed in the signal to account for predictable variations between the recording and the musical score.

One issue which still needs much further investigation because it was a continual problem throughout the analysis phase of this project is rubato. Specifically, further research needs to find a way to define the differences between rubato and physical damage to the recording. We suggest wavelet fingerprinting as a possible avenue for this continued research, because wavelet fingerprints may reveal a pattern in the artist's rubato such that damage to the recording will stand out from the patter of rubato.

Acknowledgements:

I would like to take the time now to be personal and thank a few very important persons in this whole process, without whom this project would have never happened. Firstly, I would like to thank Professor Hinders for all of his help and guidance, both on matters relating directly to my thesis and on matters relating to surviving the insanity of the last two semesters. I would like to thank Jonathan Stevens for all of his help over this last year, without whom there would have been no basis for my thesis at all. I would like to thank, Professor Fujimoto and Jonathan Mott, whose help this semester with ear training has been invaluable to working with minute differences in recording quality. Without their help, I would have been unable to effectively keep my ear trained well enough to make significant strides in this research area.

References

1. Liszt, F. “The Piano Music of Franz Liszt: Great Performer’s Edition.” Editor: Earl Wilde. G. Schirmer Inc.

2. Liszt, F. “Liszt: Sonata, Ballades and Polonaises.” Performer: Hough, S. Hyperion, UK. December 12, 2000.

3. Liszt, F. “Liszt Polonaise No. 2 in E Major.” Performer: Mamriev, A. Van Cliburn Foundation. 2007.

4. Liszt, F. “Liszt: Ballades, Legends & Polonaises.” Performer: Howard, L. Hyperion, UK. November 15, 1993.

5. Liszt, F. “Liszt: Polonaise No. 2 in E major.” Performer: Paperno, D. Cedille Records. 1998.

6. Liszt, F. “Polonaise No. 2 in E Major (Liszt).” Performer: Kudoyarov, R. Classical Records. 2005.

7. Liszt, F. “2 Polonaises, S223/R44: No. 2. Polonaise in E major.” Performer: Dube, J. Naxos. April 19, 2005.

8. Liszt, F. “Liszt: Polonaise in E Major, Mephisto Walz No. 1, etc & Dvorak: Slavonic Dances Nos. 2, 3, 6, 8.” Performer: Khaykin, B. Music Online. March 14, 2008.

9. Liszt, F. “Polonaise No.2 in E, S.223 No.2.” Performer: Richter, S. Decca. November 30, 2007.

10. Vetterli, M. “Fast Fourier Transforms: A Tutorial Review and a State of the Art.” *Signal Processing*. **19**: 259–299.

11. Howard, D., Angus, J. “Acoustics and Psychoacoustics.” Focal Press. 2001.

12. Matlab Help Site: function definition for wavelets. http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/index.html?/access/helpdesk/help/toolbox/wavelet/thselect.html&http://www.mathworks.com/cgi-bin/texis/webinator/search/?db=MSS&prox=page&rorder=750&rprox=750&rdfreq=500&rwfreq=500&rlead=250&sufs=0&order=r&is_summary_on=1&ResultCount=10&query=%22stein%27s+unbiased+estimate%22&submitButtonName=Search. 2007.

13. Olson, H. “Music, Physics, and Engineering.” Dover Publications. 1967.

14. Randel, D. “The Harvard Dictionary of Music.” Harvard University Press. 2003.

Appendix A: Figures

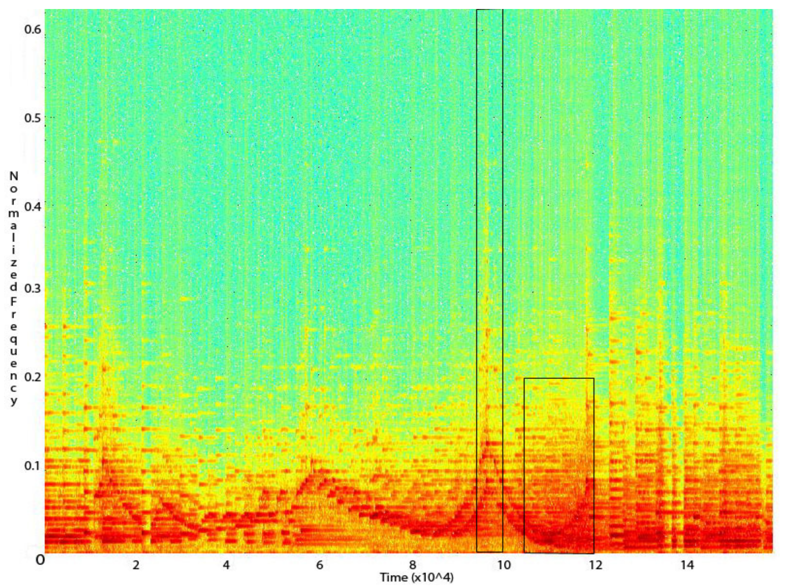

Figure 1

This is the spectrogram of “Excerpt1.wav” before anything was altered from the original Gunnar Johansen recording. There are several artifacts of note ignoring the general ambient noise. Firstly, there is a smearing of pitches in the second swing of the cadenza run. This probably was caused by original recording equipment having a hard time managing the rapidly successive pitches. Secondly, there is an artifact at the peak of the run which is an overloaded tone. That is, again, most likely a recoding error and was probably caused by the recording device not being adjusted to handle the intensity of the pitch in question.

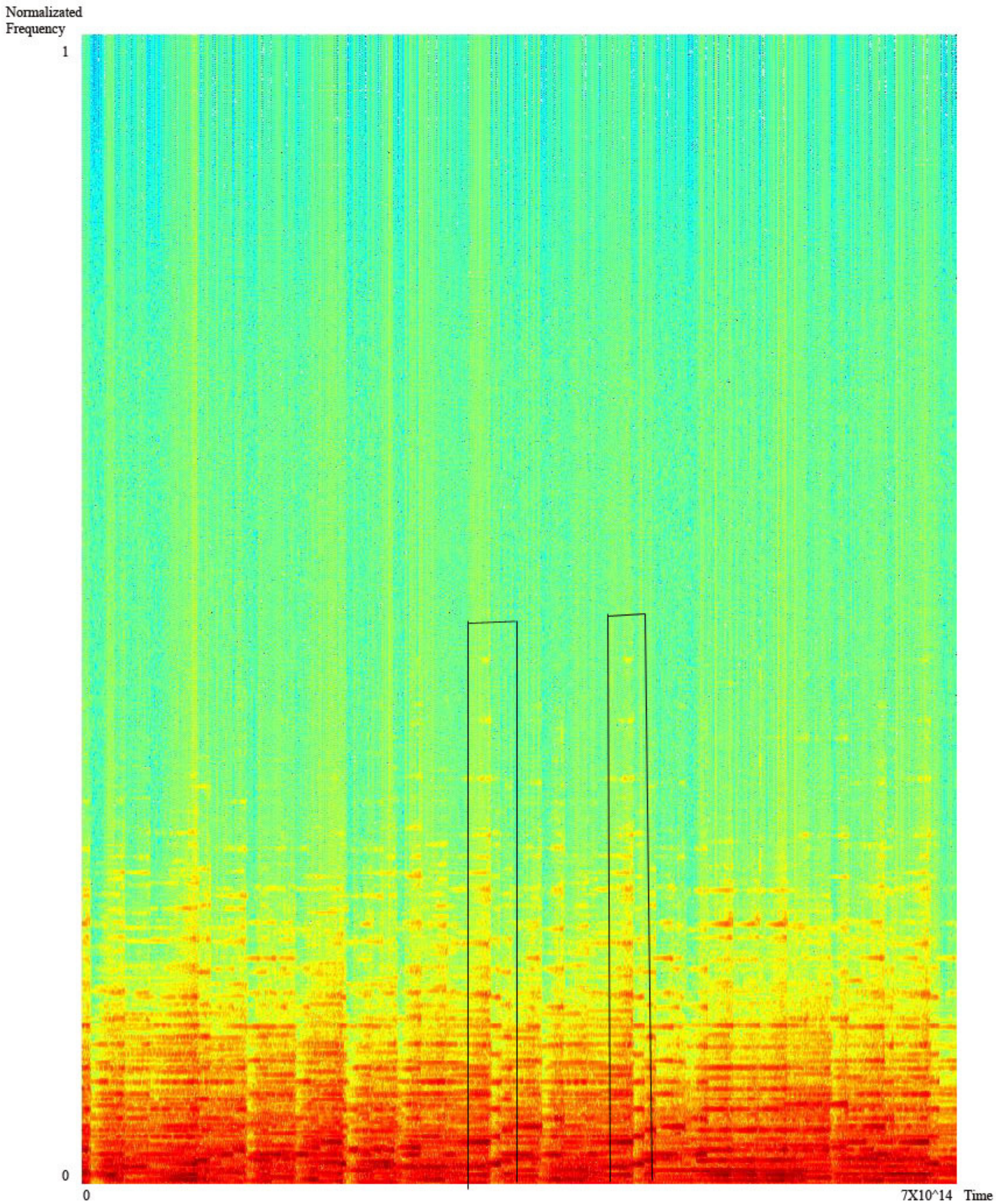

Figure 2

This is the spectrogram of "Excerpt2.wav" before anything was altered from the original Gunnar Johansen recording. The first artifact boxed is the semi-audible click which can be heard if you slow the recording down. The second artifact boxed is the audible click. There are more audible clicks in the slowed recording than are boxed in this spectrogram, but these are the most prominent.

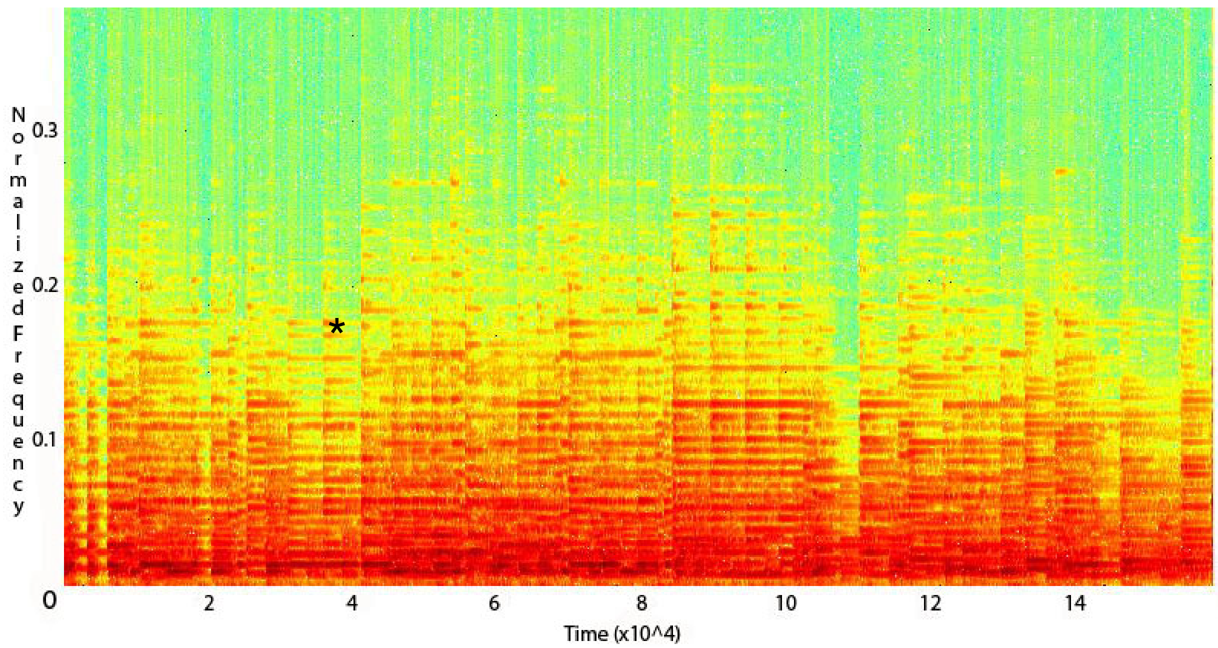

Figure 3

This is the spectrogram of “Excerpt3.wav” before anything was altered from the original Gunnar Johansen recording. It is very obvious that this particular spectrogram is of a heavily chorded section of the score. The horizontal striations are held pitches which form chords. The artifact we are interested in here is the overtone that is audible in the recording. This pitch is in the upper register. Given the start time of the overtone and the pitch, we are pretty certain that the artifact is the one associated with the red (high intensity) bars to the right of the astir.

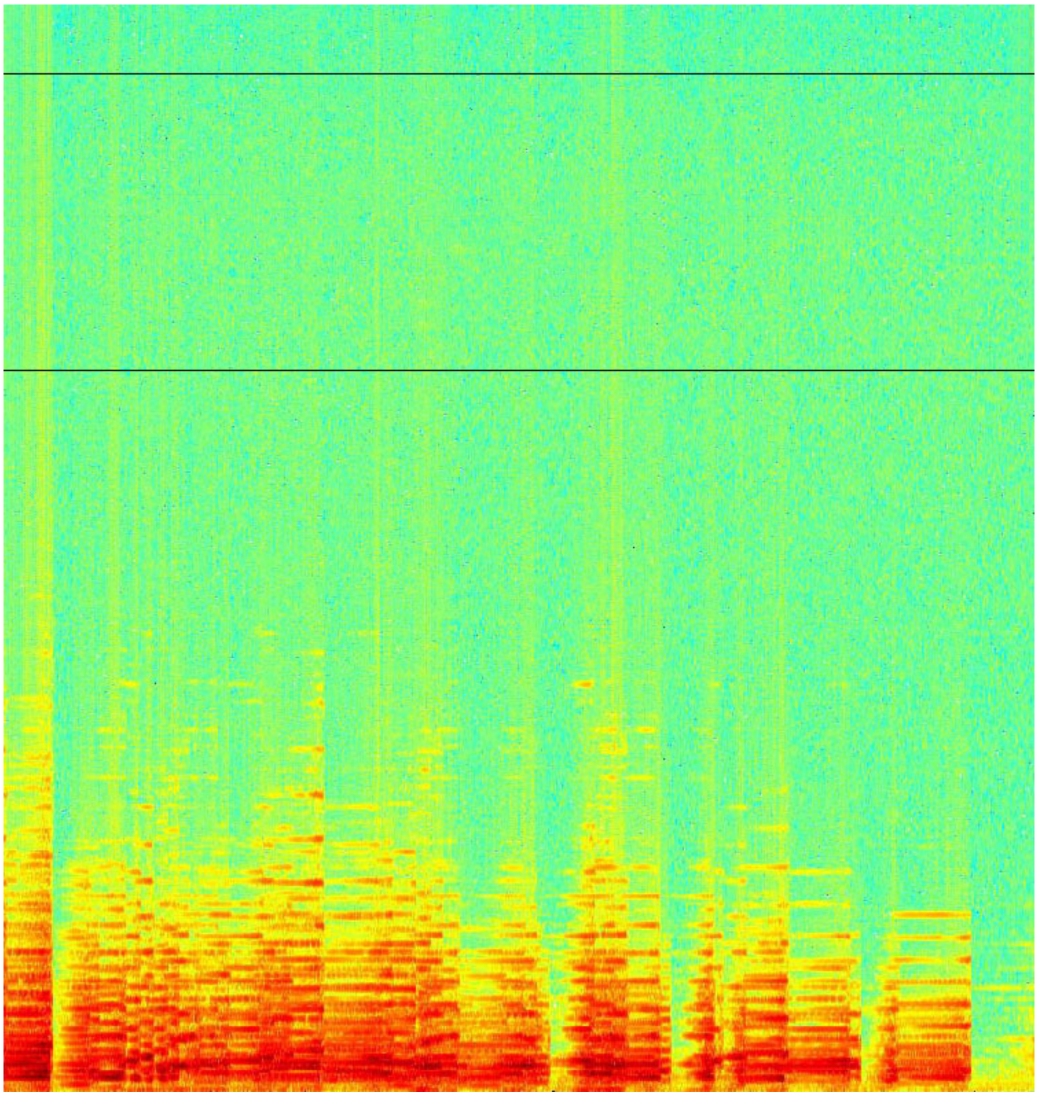

Figure 4

This is the spectrogram of "Excerpt4.wav" before anything was altered from the original Gunnar Johansen recording. There is a lot of noise in this signal and there are quite a few high intensity pitches outside of the pitch values in the score. We boxed the high frequency noise as to not disturb the spectrogram, but the noise is universal. *Note*: humans cannot hear the noise in the boxed frequency range, it is just a demonstration of the noise in the signal.

Figure 5: Equal Temperament Scale

This diagram is called the equal temperament scale [13]. The purpose of such a diagram is to relate musical pitch to frequency for purposes of analysis. This is the diagram that we used to take sections of the score and convert them to frequency charts. When analyzing a spectrogram, we could take the spectrogram and write music with it using this relation.

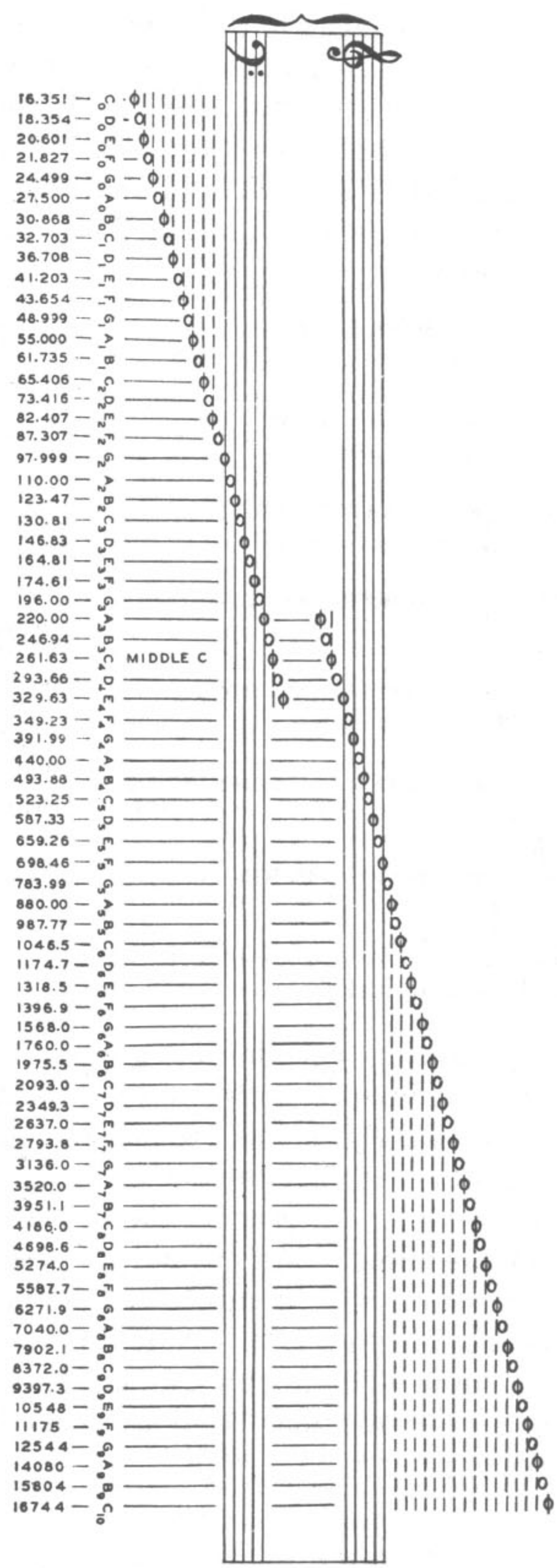

FIG. 2.9. The frequencies of the notes in the scale of equal temperament in the scale of C from 16 to 16,000 cycles.

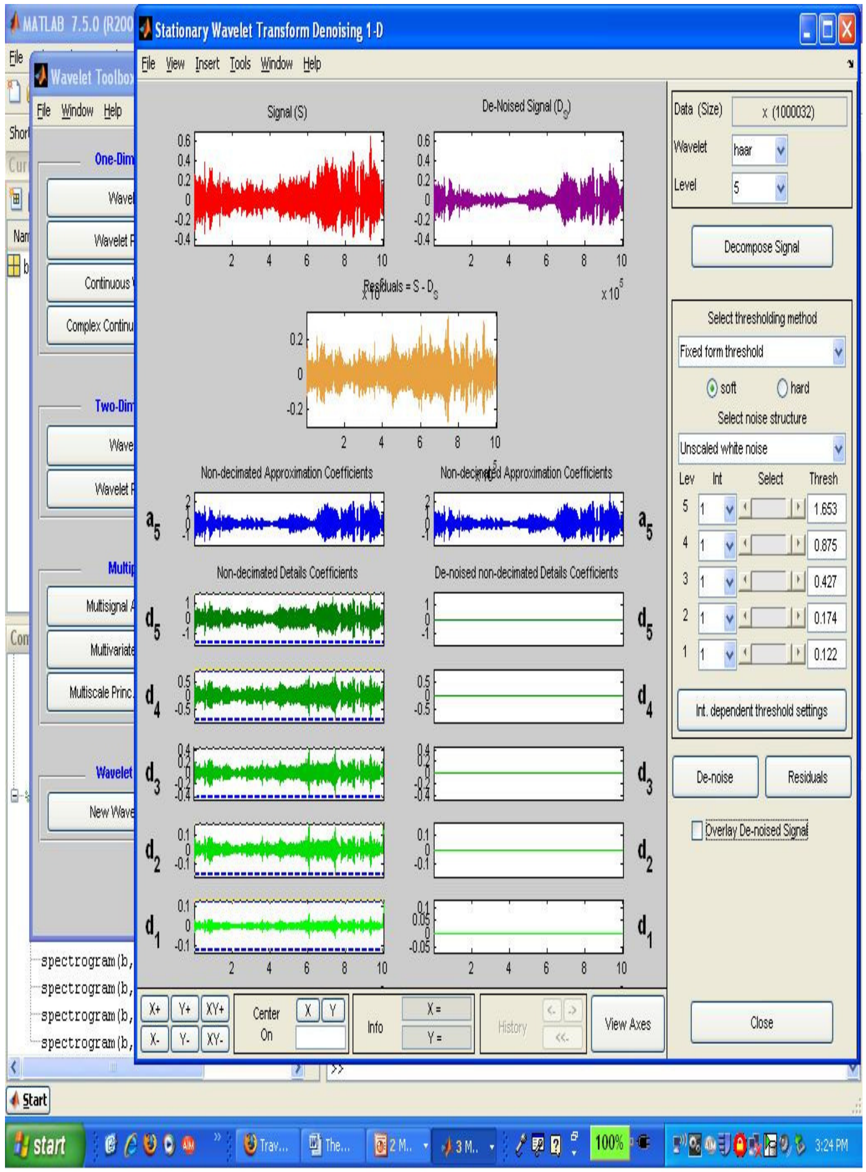

Figure 6

This figure is spread out over the next seven pages. It is principally to show the experimental set up for the majority of our wavelet de-noising. This is the toolbox for the SWT de-noising function. We first imported the signal and then cleaned it up using the MATLAB algorithms and reiterated the new signal in the MATLAB command module to produce a cleaned spectrogram. Above is the set up associated with Figure 8

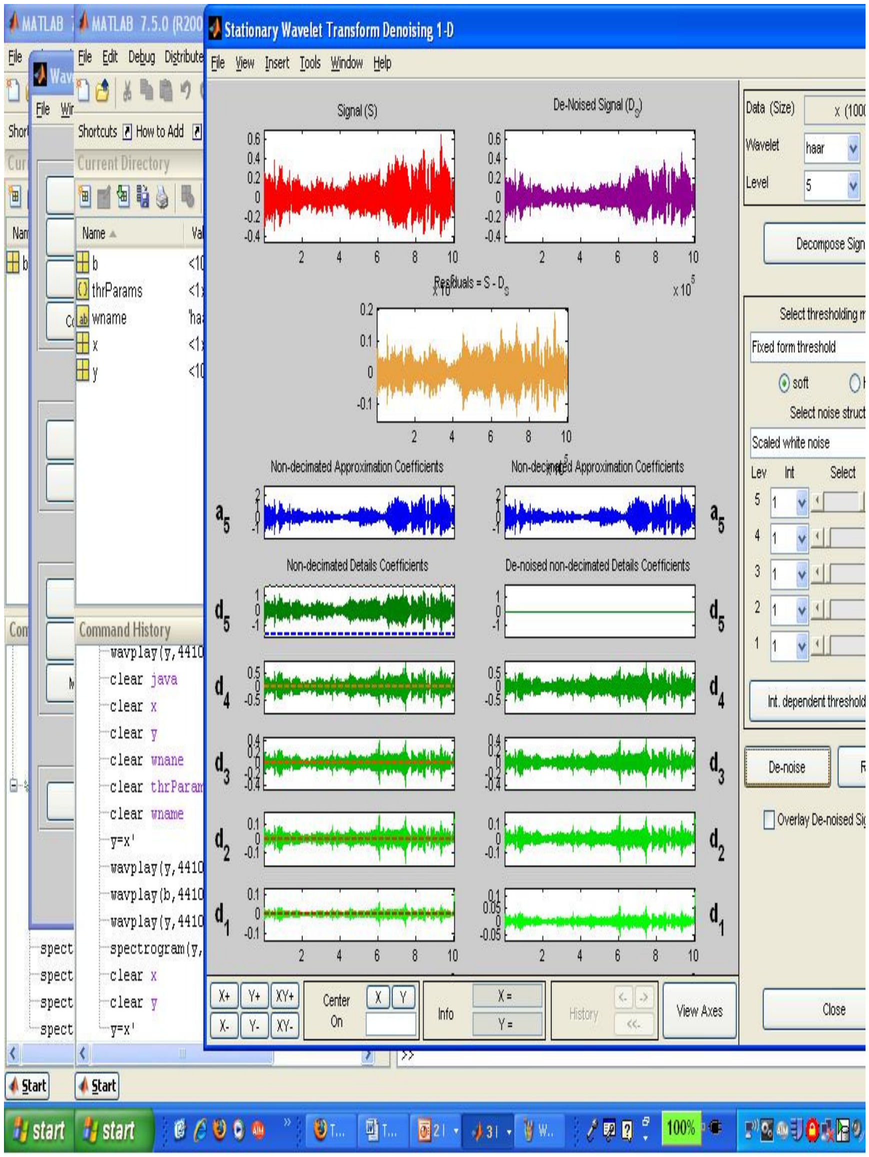

Figure 6

Above is the setup associated with Figure 10, killing only d5.

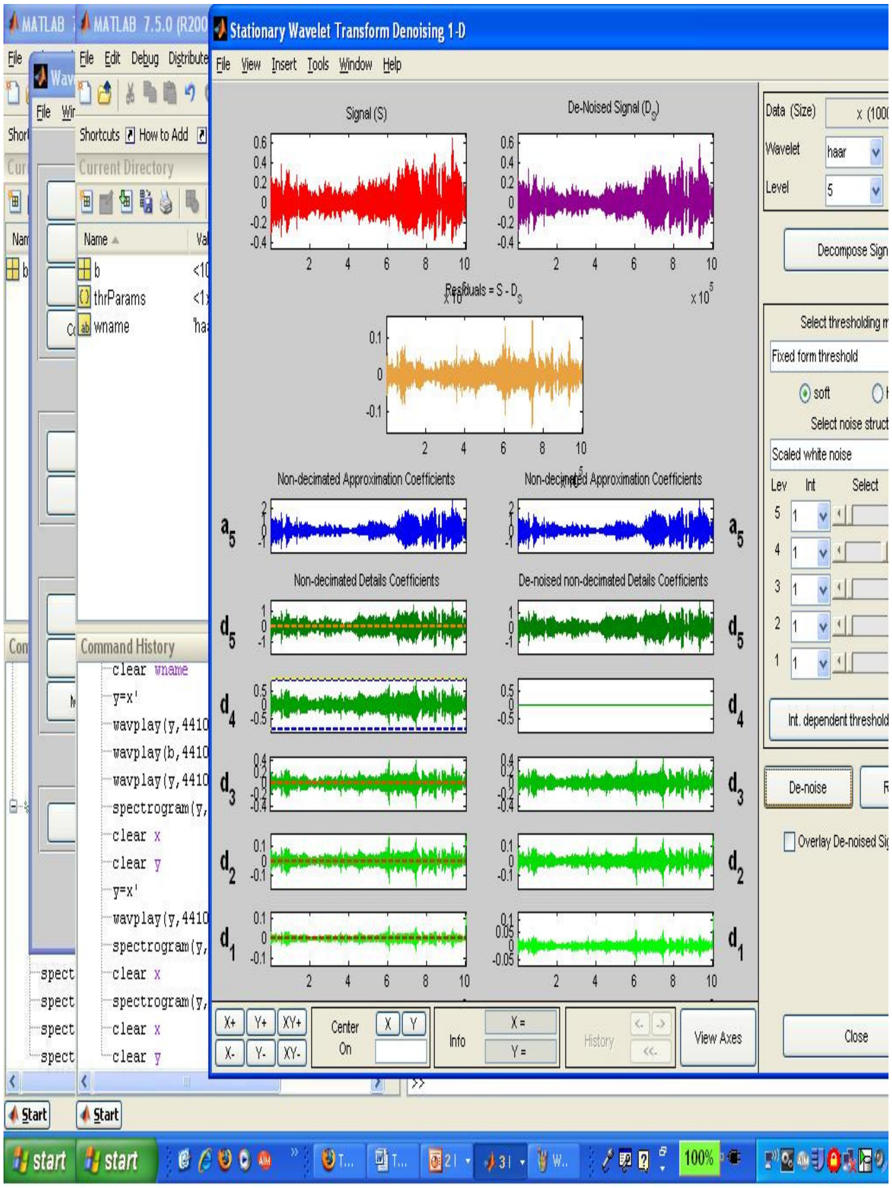

Figure 6

Above is the setup associated with Figure 11, killing only d4.

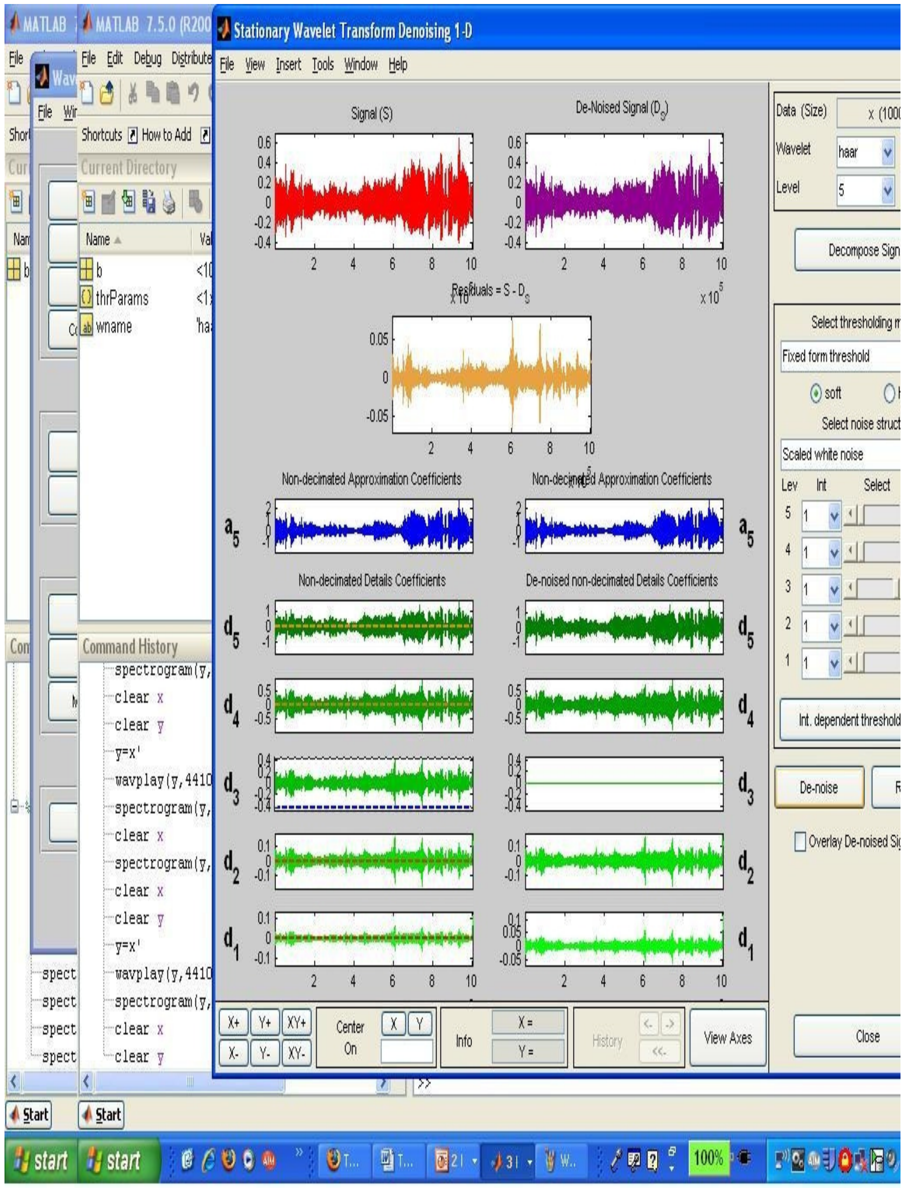

Figure 6

Above is the setup associated with Figure 12, killing only d3.

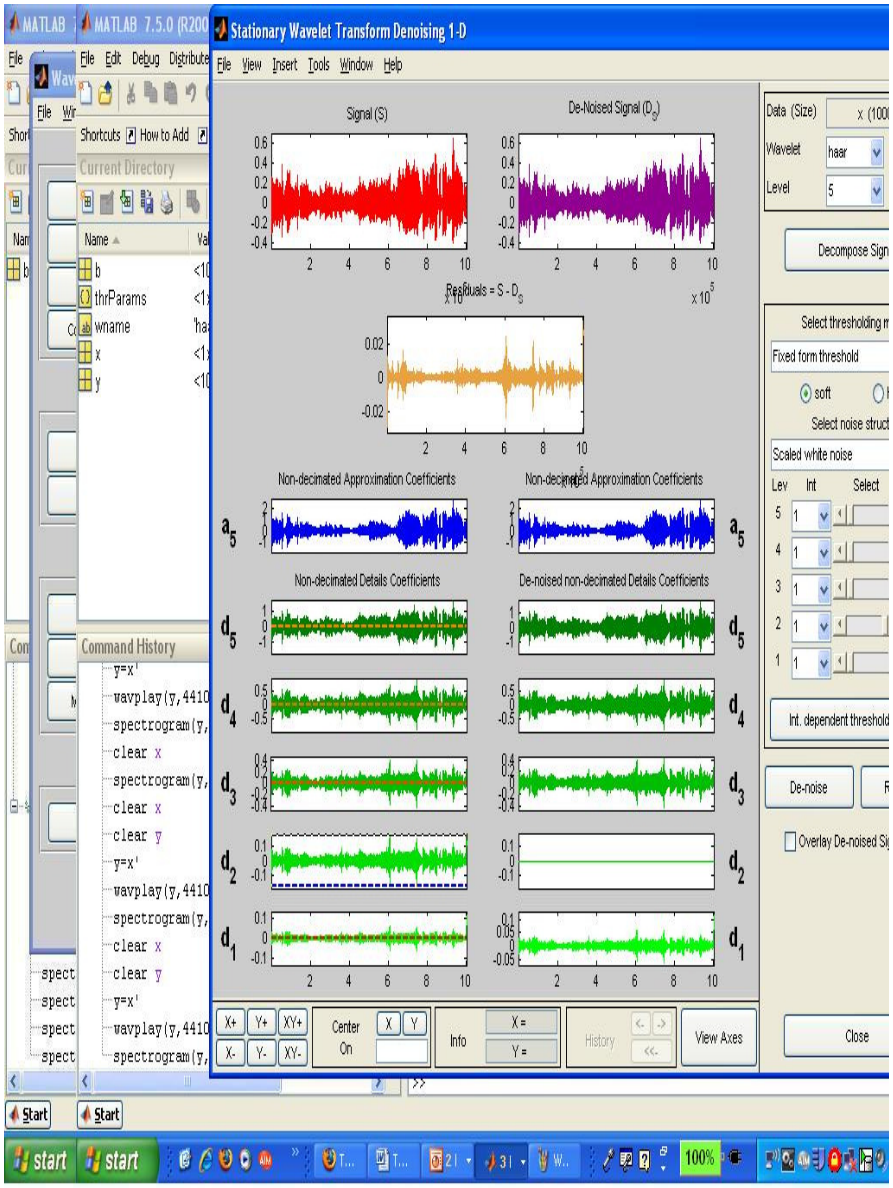

Figure 6

Above is the setup associated with Figure 13, killing only d2.

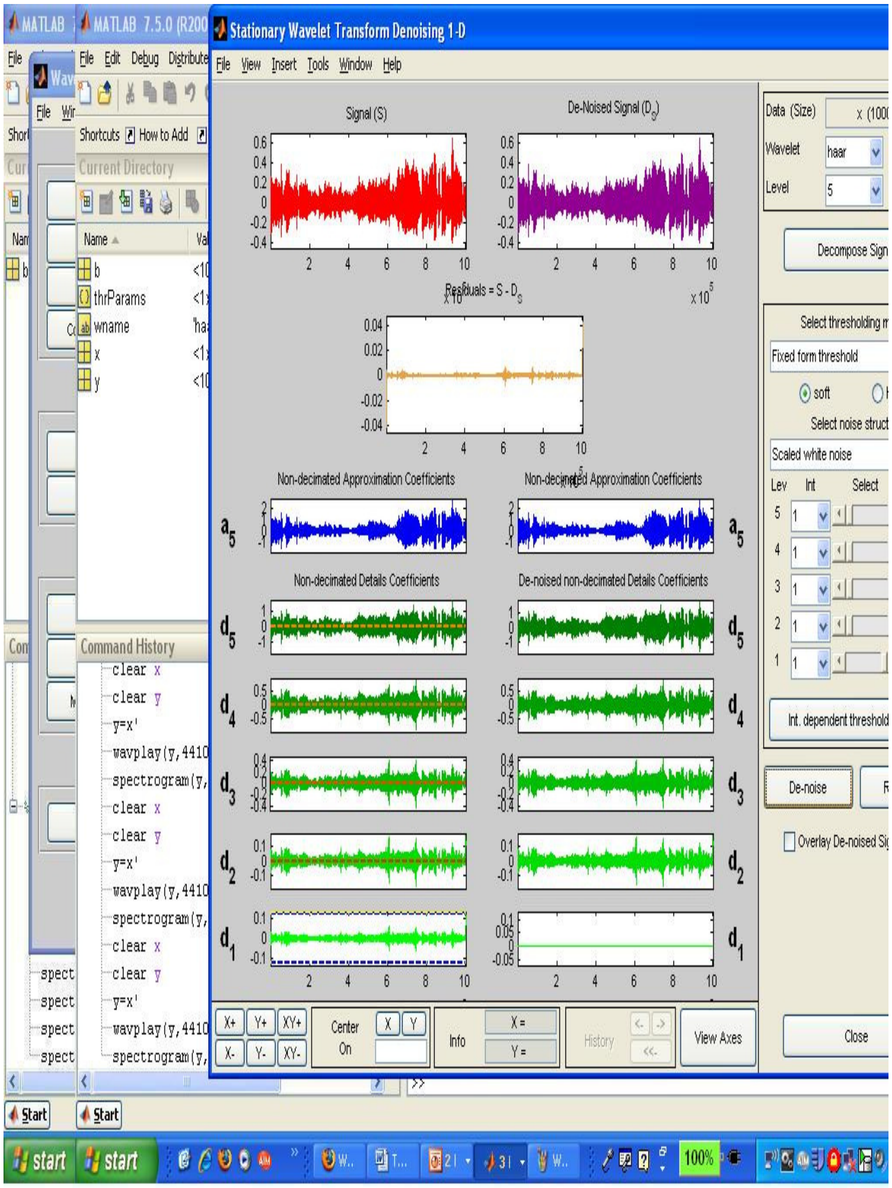

Figure 6

Above is the setup associated with Figure 14, killing only d1.

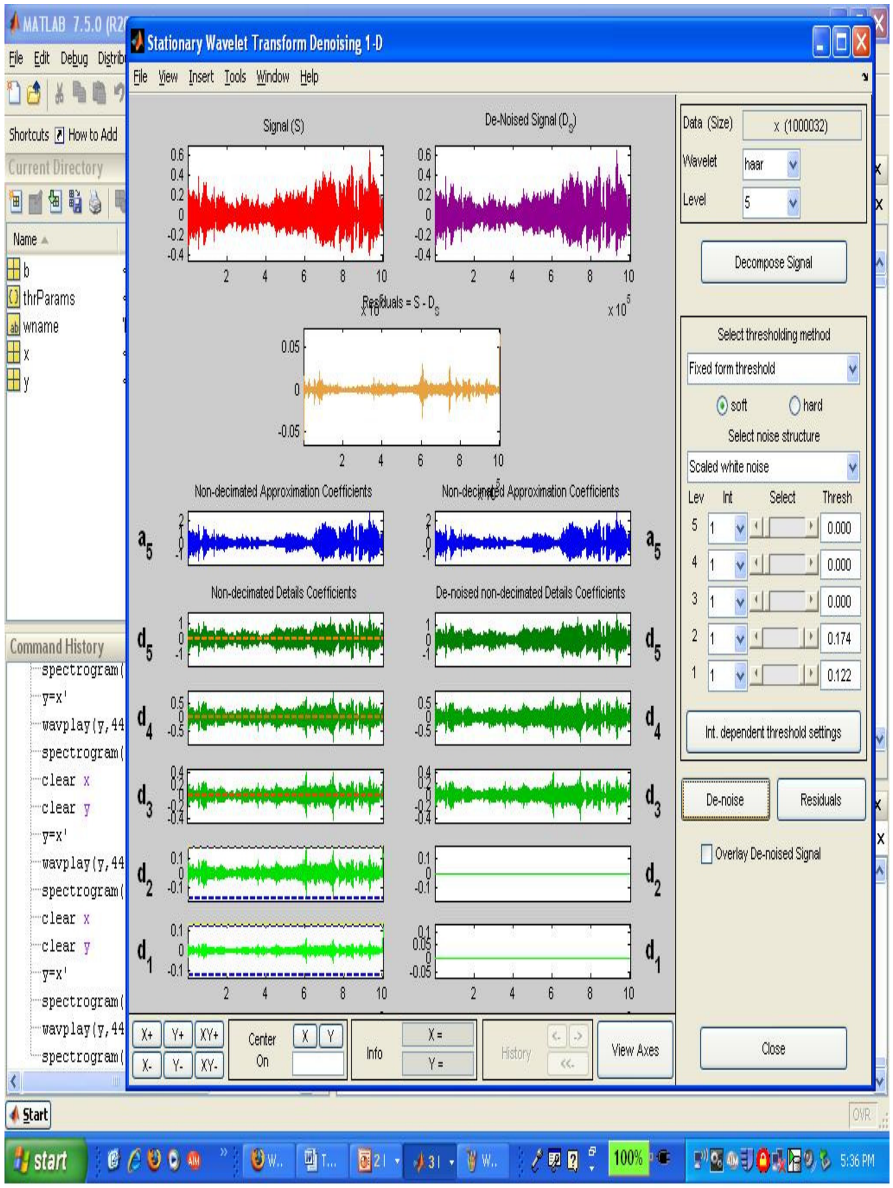

Figure 6

Above is the setup associated with Figure 15, killing d1 and d2.

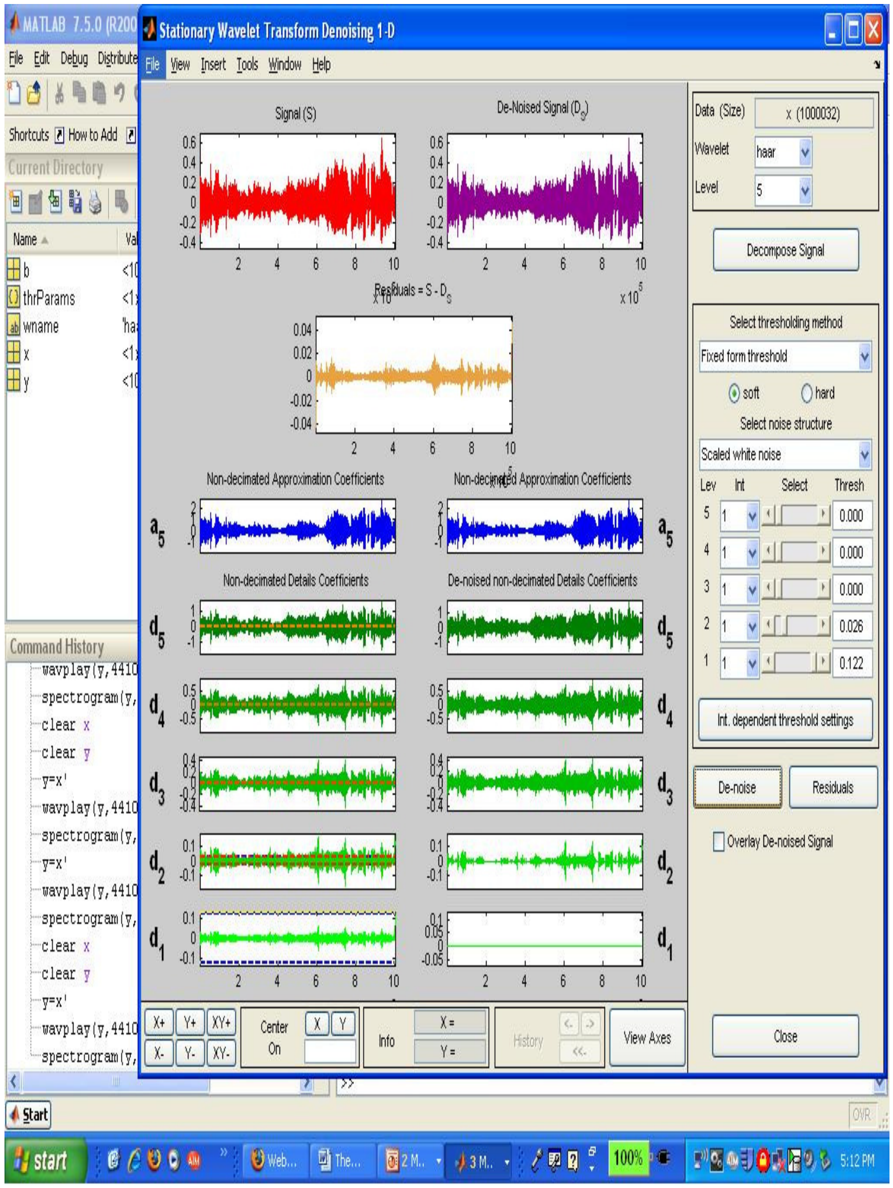

Figure 6

Above is the setup associated with Figure 16, killing d1 but only partially killing d2.

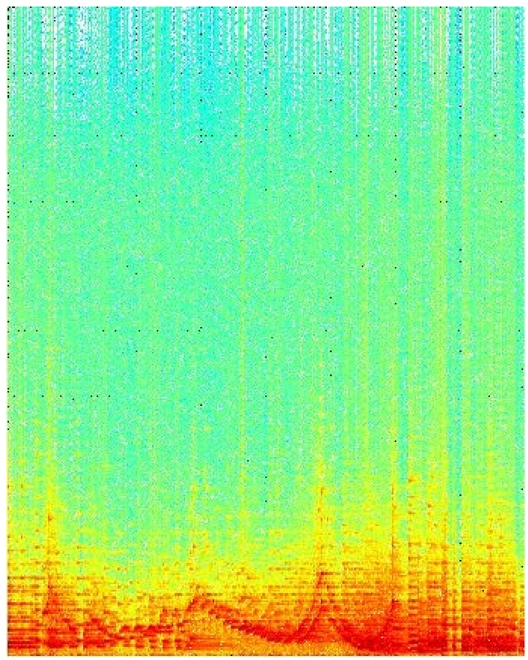

Figure 7

This is the sound clip and associated spectrogram for Excerpt 1. This figure is reiterated from Figure 1 because the time scale has been adjusted to fit wavelet analysis. This clip was the first target section for cleaning. We used the wavelet toolbox to clean the signal. The build-in function of MATLAB for de-noising wavelets was used to attempt to clean up the noise in this signal.

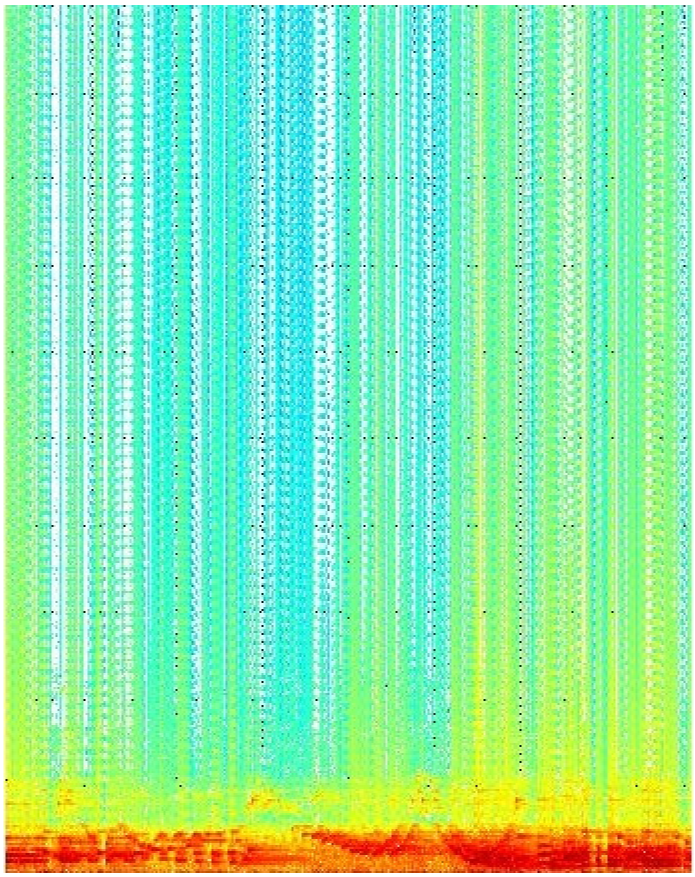

Figure 8

Killing all five coefficients (d5-d1) did an excellent job of killing the noise. Unfortunately, it also killed the treble region of the music signal, which meant that this "holistic" cleaning of noise failed to produce a "better" signal.

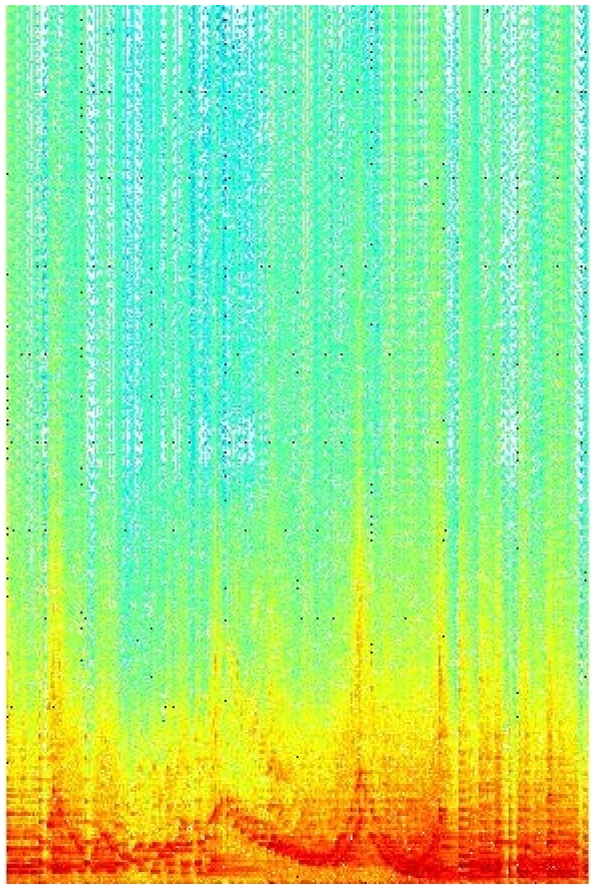

Figure 9

The next logical step was to try partially killing the coefficients in an attempt to keep the music signal but not the noise. Unfortunately, this created a "tin" sounding signal, which is even worse from an aesthetic perspective than cutting out the majority of the treble signal. Notice the blur in the region around matrix index 3750000 over the beginning of the run. This blur is uglier, musically speaking, than the worst of the pre-cleaned noise in this passage.

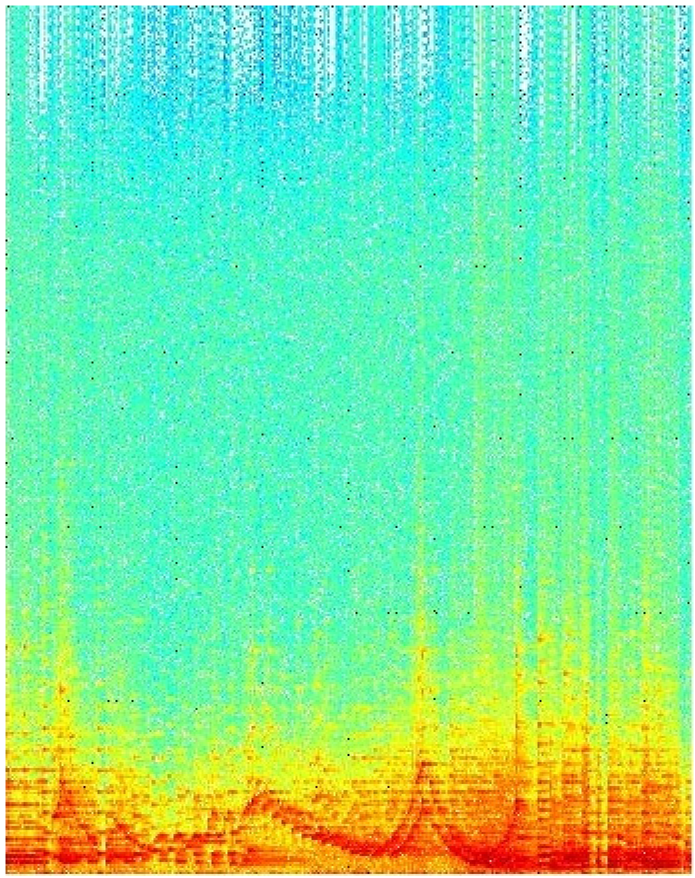

Figure 10

In order to find which coefficient was having the greatest impact on the signal, we iterated the process of killing on coefficient and leaving the rest untouched five times, one for each coefficient. This is the iteration for the d5 coefficient. Killing the d5 coefficient kills the noise above the auditory region.

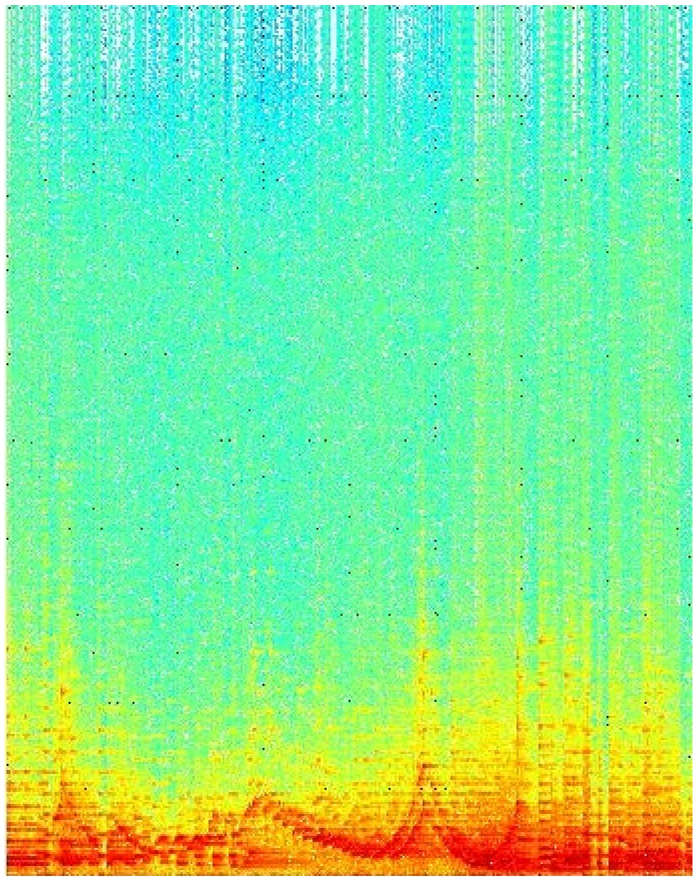

Figure 11

The d4 iteration had a similar effect as that of the d5 iteration, but not as effective in the cleaning ambient noise. The only positive difference in the d4 iteration is that it seems to have less of an effect on the auditory signal than does d5.

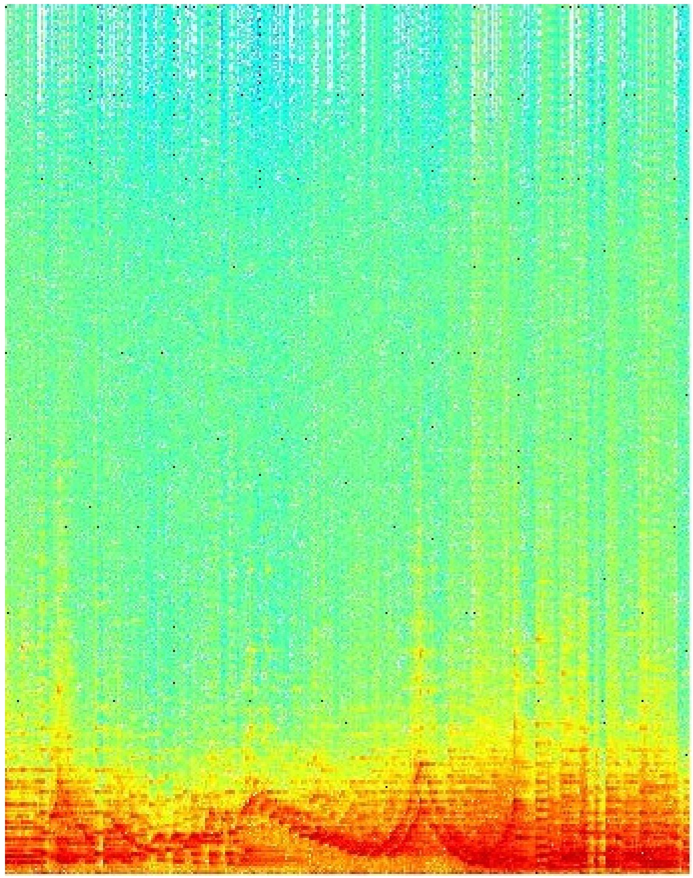

Figure 12

For the purposes of our experiment, the d3 iteration seems to be of little interest. Not only does it have little effect on the ambient noise, but it also has intensified noise in the auditory region, making all of the pitches sound close to overload.

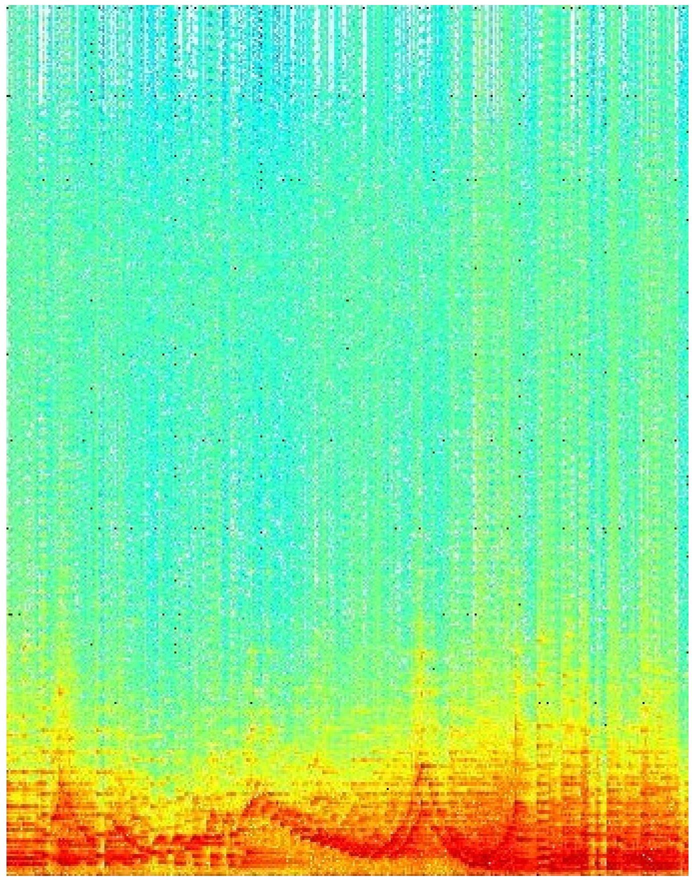

Figure 13

When we listened to the d2 iteration, even before producing the spectrogram, we heard the first real noise improvement in our wavelet trials. The noise was reduced both in the auditory and non-auditory regions without serious disruption of the music signal.

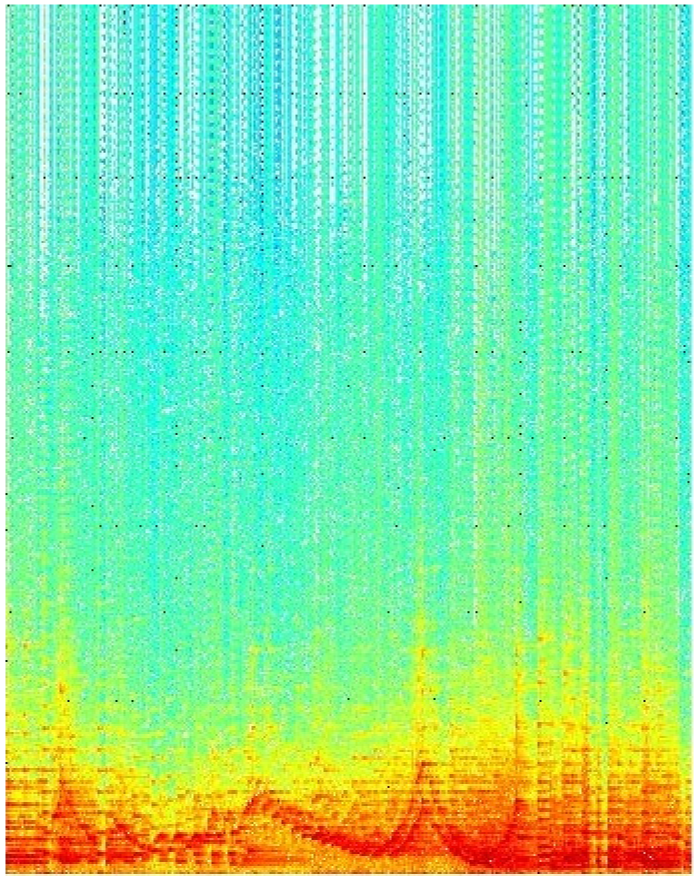

Figure 14

The cleanest signal from killing one of the detail coefficient iterations occurred in the d1 iteration. The sound quality of this iteration was noticeably better than that of the original signal.

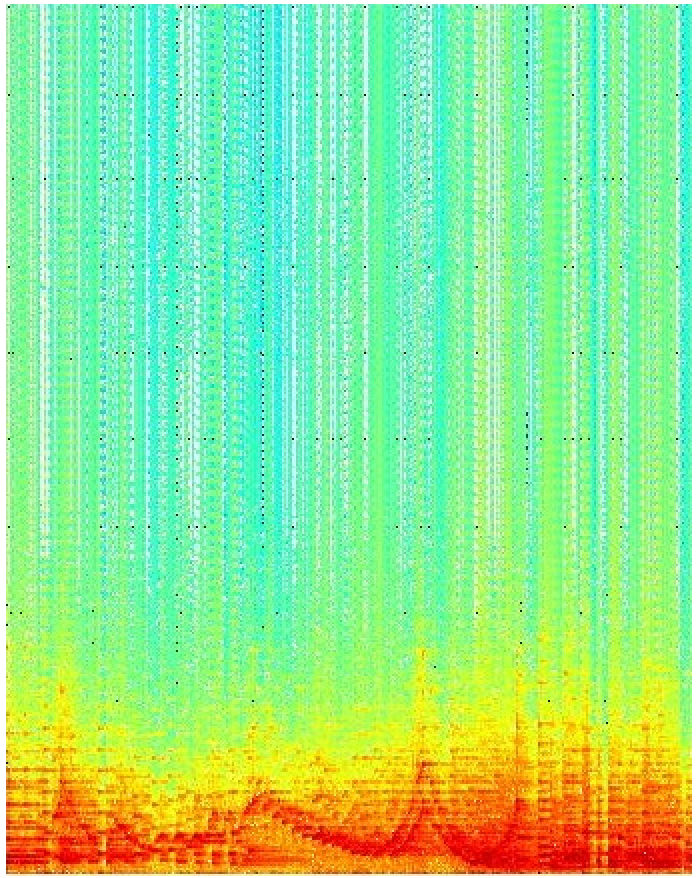

Figure 15

The next logical step, after the effectiveness of the iterations of d1 and d2, was to combine the two together. However, killing both was not effective in cleaning up the signal. Though some auditory distortion was cleaned, the music signal sounded fuzzier than the original signal.

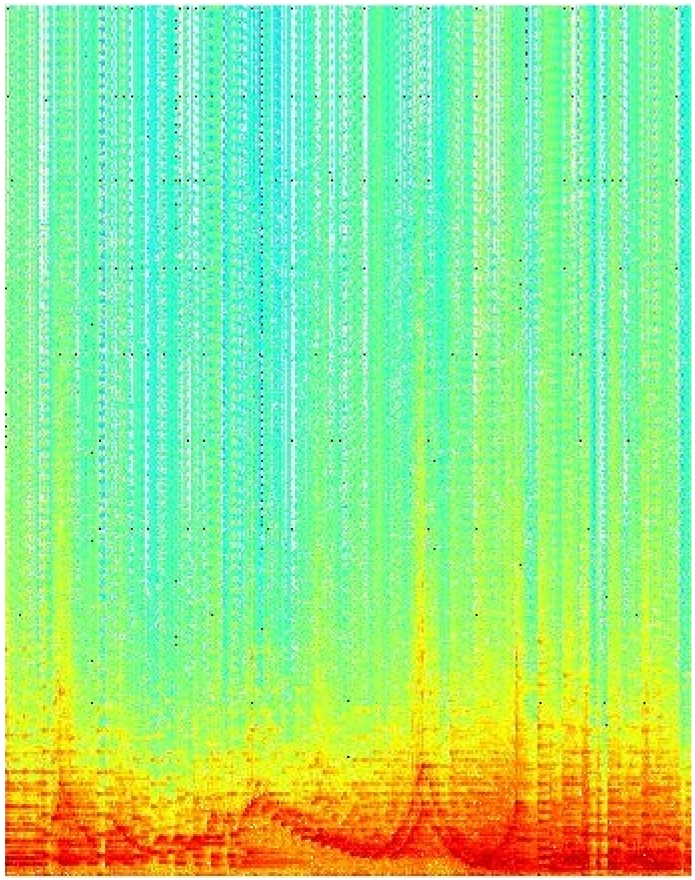

Figure 16

Our final attempt to make the d1-d2 combinations work in conjunction ended in failure. Though the signal was better than the original signal, on whole the cleaned signal was not as sharp as the iteration of d1.

Appendix B: Music Glossary for Physics Readers

For more information, we recommend "The Harvard Dictionary of Music"

Rubato- "In performance, the practice of altering the relationship among written note-values and making the established pulse flexible by accelerating and slowing down the tempo; such flexibility has long been an expressive device. Two varieties of rubato are usually discussed. In the first, the underlying pulse remains constant while the rhythmic values are minutely inflected…the second type is the more common present day understanding of rubato. Changes in tempo and rhythmic figuration are made in all parts at the same time without any compensation; the original tempo is simply resumed at the performer's discretion…" [13]

Vibrato- "A slight fluctuation of pitch used by performers to enrich or intensify the sound. In modern string playing, vibrato is produced by rocking the left hand, usually from the wrist, as the note is played; in modern wind playing, it is effected by regulating the air flow into the instrument or by varying the tension of the lips or the pressure of the mouth on the reed or mouthpiece…" [13]

Forzando- "Forcing, forced, i.e., strongly accented." [13]