

# **CASSETTE TAPE SIGNAL PROCESSING**

Prof. Hamid Jafarkhani

David Andersson

Adam Heal

Stephen Teran

## **ABSTRACT**

Audio recording has evolved from a 160 year old mechanical process with no playback to a multitude of mastering options. Magnetic cassette tapes were the de facto standard for audio recording for over three decades, creating an abundance of cassette tapes with recordings that exist in no other format. As these tapes deteriorate over time, many of the recordings can be lost to time. However, these tapes can be converted to digital recordings that will not suffer this same degradation over time. Once in a digital format, the audio signal can be processed in an economical manner using digital filters that can remove much of the noise and artifacts prevalent to the magnetic format. This paper discusses a procedure than can be employed by a motivated hobbyist to digitally sample and filter recordings from compact cassette tapes.

# TABLE OF CONTENTS

Physics of sound.....	5
Anatomy of hearing	5
Anatomy of speech	7
A Brief History of Audio Recording.....	7
The Progression of Mechanical Recorders	8
Optical Recording in the Film Industry	10
Magnetic Recording Technology	11
Digital Recording and the Compact Disc	12
Magnetic Recording Formats.....	13
Steel Wire and Tape	14
Reel-to-Reel Recorders	14
4-Track and 8-Track Cartridges	15
Compact Cassette Tapes	15
Preservation in the wake of deterioration.....	16
A National Dilemma	16
Analog Tape Restoration	17
Tape Condition	17
Playback	18
Post Transfer	18
Digital Formats.....	18

Compact Discs and the Red Book Format	19
The Wave Format	20
Lossy and Lossless Codec	20
Analog-to-Digital Conversion	21
Pulse-Code Modulation	22
Digital Filtering	22
Overview of Digital Filtering	23
Design of the Filter	24
Magnitude	25
Phase	25
Specifications	26
Implementation and Testing	28
Ethical Concerns	34
Economics and Manufacturability	35
Conclusion	35
Appendix I – Matlab Code	37
Works Cited	41

# 1 PHYSICS OF SOUND

Sound is mechanical vibration created by a disturbance that propagates energy through air particles or other medium as waves. The waves are longitudinal in that the air particles oscillate parallel to the direction that the energy is transported. The wave is expressed mathematically as a sinusoid, the changing amplitude of the signal indicating pressure compression (crests) and rarefaction (troughs) of the air particles. Waves can be characterized and distinguished from one another by their inherent properties of frequency, period, wavelength, amplitude, intensity and velocity. Complex waves are formed through the constructive or destructive interference of multiple waves. The fundamental frequency of the resulting periodic wave takes on the value of lowest frequency wave of the harmonic series. The interfering waves are harmonics of each other if their frequencies are integer multiples of one another, combining to producing a pleasing sound. If the interfering wave frequencies are not related by a simple whole number ratio, the resulting non-periodic sound will most likely be distasteful. While the physical properties of sound waves can be quantized for their actual values, it is the human auditory perception of them that is of importance.

## 1.1 ANATOMY OF HEARING

Human hearing is an exceedingly complex system. Sound is directed from the environment through our ear canal where it reaches the tympanic membrane (ear drum), a tissue that vibrates in unison with the sound waves. The signal is transmitted via the middle ear bones to the cochlea, a fluid filled corridor that contains the basilar membrane. With over 10,000 nerve cells forming the cochlear nerve, the basilar membrane varies in stiffness along its length, “each nerve

cell only [responding] to a narrow range of audio frequencies, making the ear a frequency spectrum analyzer” (Smith 353). The interactive network between the ear organ and the brain make for a phenomenal human sense. However, the neural transformation that takes place between the outside world and a person’s recognition of the stimulus only closely approximates the true input signal as subtle effects cause a difference in sound intensity and frequency. Psychoacoustics looks into our experience of sound and attempts to correlate our subjective perception with the actual physical properties of the signal we are hearing. *Loudness* is the perceived intensity of sound and is expressed on a logarithmic scale called decibels SPL (Sound Power Level) (Smith 352). Using the decibels SPL scale, we can marginalize the amplitude of sound into a range of one-million from 0 dB, the faintest detectable sound, to 120 dB where sound begins to be uncomfortably loud. The midpoint of that range at 60 dB occurs during human speech and the human ear can detect a change in loudness by every decibel. For every increase in sound power by tenfold, the sound is only thought to be twice as loud. *Pitch* corresponds to our judgment of the predominant or fundamental frequency of the sound wave. Humans can hear frequencies between the range of 20 Hz to 20 kHz but our hearing is not equally sensitive across all frequency levels. A frequency difference between two lower frequencies can be heard much easier than the same frequency difference between two higher frequencies. This is in part due to our nonlinear perception of pitch where the frequency scale is divided into octaves which are intervals of doubled frequency. The information drawn from sound is distributed in this same fashion as the frequency scale, for example, “as much audio information is carried in the octave between 50 Hz and 100 Hz. as in the octave between 10 kHz and 20 kHz” (Smith 358). Sound intensity and frequency also play a dependant role upon one another which ultimately governs our perception of

loudness and pitch.

## **1.2 ANATOMY OF SPEECH**

The act of human communication through speech is a physically involved process from the perspective that many parts of the body contribute to the production and transmission of speech sounds. Our voice stems from our ability to process air through our respiratory system and act upon it with our vocal chords. Air is inhaled with the contraction of our diaphragm, a muscle below our ribcage, the air filling up our expanding lungs. The diaphragm then rises to force air out through our trachea, or airway, during exhalation and allows for sound generation through the operation of the larynx, or voice box, on the forced out air. Two elastic bands of tissue, our vocal chords, housed inside the larynx vibrate to produce sound. With muscular changes in our throat and components in our mouth we create speech. The orientation of our speech organs such as our vocal chords, teeth, lips and tongue allow us to regulate the passage of air and make vowels (vocal sound produced without vocal tract restriction) and consonants (vocal sound filtered by air passage articulation). Human speech transmits complex waves in a 3.2 kHz bandwidth (from 200 Hz to 3.2 kHz for normal speech) as opposed to music that extends to 20 kHz (Smith 358). Our speech bandwidth still represents 8 out of 10 octaves of sound information detectable to the human ear, even though it only represents 16% of the frequency range.

## **2 A BRIEF HISTORY OF AUDIO RECORDING**

Audio recording has evolved over the past 150 years from waveforms scratched onto a cylinder into digital files that transcend corporeal space. The original recordings were experiments in sound, using nothing but the mechanical pressure of the human voice to trace a line on a

rotating cylinder; today's hobbyists can choose from a multitude of various technologies ranging from 4-track magnetic cassette tapes to digital mastering. Fidelity and signal-to-noise ratio were once constraints that defined a technology, while current technologies can reproduce any sound a talented sound engineer can provide and that a given set of speakers can reproduce.

The original recordings were made with machines that replicated the mechanics of the human ear. In the late 19<sup>th</sup> century, electromagnetic technology developed for the telephone was incorporated into the recorders, improving upon the purely mechanical apparatus. At the turn of the century, magnetic recording was developed. Although this technology would eventually become the de facto standard worldwide, it initially was used only in niche markets in Europe. Magnetic tape recording was continually improved upon throughout much of the 20<sup>th</sup> century and only faced competition in the form of modern electronic phonographs. The audio recording landscape over the first 120 years was characterized by continual technological upgrades and steady sound quality improvement.

All of that changed in the 1980s with the advent of digital audio and the compact disc. The compact disc represented a medium with sound quality that never degraded and did not introduce noise into a recording. The increasing availability and affordability of personal computers in the 1990s soon allowed anyone with a sound card to create their own digital masters and produce their own CDs.

## **2.1 THE PROGRESSION OF MECHANICAL RECORDERS**

The original audio recording devices were essentially mechanical ears that engraved the sound onto a rotating surface. Leon Scott's phonautograph, designed in 1857, was the first device that recorded sound. As seen in Figure 1, the phonautograph funneled sound through a cone

into a flexible membrane. The compression of the sound wave vibrated the membrane which oscillated the attached stylus.

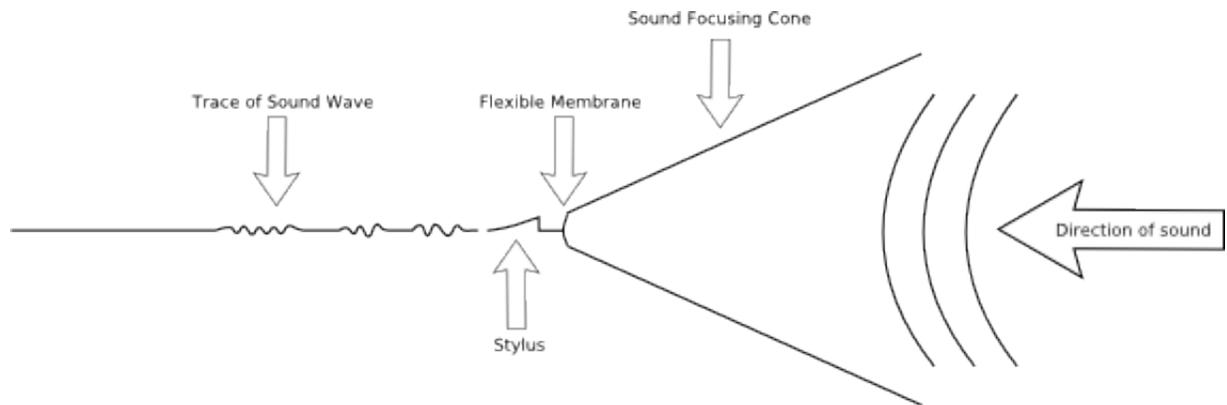


Figure 1 – Diagram of Phonograph Operation

Each component of the phonograph can be mapped directly to a part of the anatomy of the ear. The sound focusing cone is analogous to the outer ear, the flexible membrane to the tympanic membrane (ear drum), and the stylus to the malleus, incus, and stapes of the middle ear. Whereas the bones of the middle ear transfer the acoustic energy from the air into hydraulic pressure in the cochlear fluid, the stylus transfers this same energy into lateral translation that is recorded as a trace on a kerosene smoke-blackened piece of glass. Reproductions of the original traces can still be found, however the accuracy and sensitivity of the machine was too limited for any sound to be reproduced from them with modern techniques (Morton).

Playback of audio recordings was not possible until the development of Thomas Edison's phonograph in 1877. Although Edison originally created a device to record directly from telephone wires, he later incorporated a diaphragm which allowed the phonograph to record acoustically. The storage medium of Edison's choice was a cylinder covered in tinfoil that was embossed in a *hill and dale* (up and down) manner by a stylus attached to the diaphragm. Edison's first attempted recording was the nursery rhyme "Mary Had a Little Lamb." Upon the first suc-

cessful playback, Edison remarked, “I was never so taken aback in all my life...I was always afraid of the things that worked the first time (Miller 336).”

The last major advance of mechanical recording came in the form of an amplified electromagnetic cutting head that was developed by the Western Electric Company in the early 1920s. Columbia used this technology to create clearer sounding records. Purely economic reasons also led the switch to disc-based recordings over cylinder-based recordings; the *stamping* process allowed for the cheap mass-production of prerecorded albums.

## **2.2 OPTICAL RECORDING IN THE FILM INDUSTRY**

The burgeoning film industry of the early 20<sup>th</sup> century was in a transition from silent films to sound films, or “talkies”. Early phonographs were experimented with, but synchronization was a constant problem. The film industry’s need for a dependable way to sync audio during playback led to the development of optical recording systems. In an optical system, the sound wave is recording directly onto the film alongside the picture. Figure 2 demonstrates two different methods of optical recording, variable density on the left and variable width on the right (Iainf, 2006). The variation in each method represents the shift in the amplitude of the audio signal.

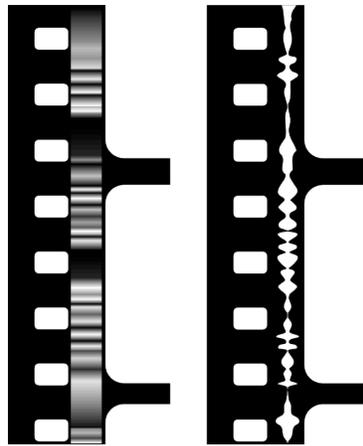


Figure 2 – Optical Soundtracks showing Variable Density and Variable Width methods

With the development of magnetic recording, movie studios experimented with bonding a magnetic strip to the film and recording the soundtrack on the strip. However, sound degradation issues led to the improvement of the optical systems which are still in use today.

### **2.3 MAGNETIC RECORDING TECHNOLOGY**

Magnetic recording was developed concurrently to early phonographs by Danish inventor Valdemar Poulsen. His initial recording, made in 1900, used steel wire as a storage medium, although any magnetizable media is suitable. Although this technology was not accepted initially in the United States due to the sensation that was the phonograph, magnetic recording was adopted throughout Europe. German scientists in the 1930s produced the first practical tape recorded, the Magnetophon.

The basic action of magnetic recording can be seen in Figure 3 (Nave, 1999). An audio signal in electric form (for example, a voice transduced by a microphone) is passed through a wire wrapped around the ferromagnetic tape head. The signal current induces a magnetic field in the tape head with flux that fringes out from the gap in the head. The fringing flux then induces

a magnetic field in the recording media (in this example, magnetic tape coated in an oxide powder) with a magnitude proportional to the signal strength (Fitzgerald).

Playback of a recorded signal is simply the reverse process of recording. Running the recorded tape across the tape head results in flux being induced in the gap by the varying magnetic field in the tape. The flux creates a magnetic field in the tape head which then induces the signal current back into the wire.

All technologies that employ magnetic recording operate in the same manner whether the storage medium is steel wire, magnetic tape, or a platter in a computer hard disk drive.

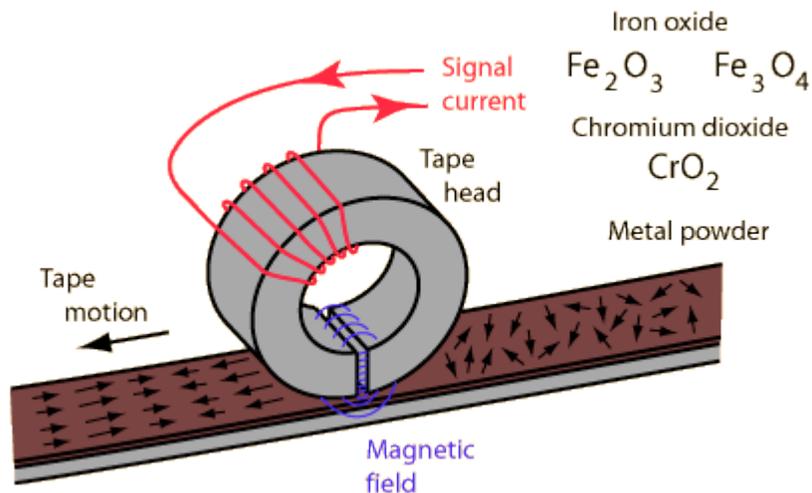


Figure 3 – Cassette Tape Head Recording Action

## 2.4 DIGITAL RECORDING AND THE COMPACT DISC

Until the development of the CD, all audio recording formats suffered from quality degradation over time and due to the mechanics of playback. The Red Book compact disc standard, developed jointly by Sony and Philips Consumer Electronics and released in 1980, stores audio

in a digital format that eliminates signal degradation. Digital formats store the audio signal as a sequence of binary values that are either 1 or 0. Also, there is no physical contact between the storage medium and recording mechanism. Analogous to the phonograph stylus and tape head is the laser and optical pickup combination which reads data from the disc by sensing the intensity of the light reflected by the disc.

Although the CD introduced digital recording to the mainstream consumer in 1982, recording studios had been working digitally with audio for over a decade. The CD also prompted the development of the Digital Audio Tape (DAT), which recorded the signal digitally onto a cassette similar to the standard compact cassette. This combined digital-on-magnetic approach was never widely accepted in the consumer market for audio, but instead became the basis of the platter-based computer hard disk drive.

Widespread adoption of personal computers also spawned a multitude of digital file formats for audio storage. Digital file formats created a new paradigm where recorded audio was no longer bound to the medium it was stored on. Instead, files on computer could easily be moved, modified, copied, or recorded to a new physical medium. Digital Signal Processing (DSP) became a tool that can be used to accomplish a variety of different functions with an audio track that were not easily accomplished (if even possible) with purely analog tools.

### **3 MAGNETIC RECORDING FORMATS**

Although the technology behind magnetic recording varied little since inception, the storage format evolved continuously over 90 years. The first magnetic recorders used steel wire and then steel tape as a medium. Later developments introduced Mylar tape with various magnetiz-

able coatings which replaced the dangerous steel tape. Noise Reduction (NR) technologies increased the fidelity of the recordings, and the development of stereo allowed multiple channels of audio to be recorded on the same tape.

### **3.1 STEEL WIRE AND TAPE**

Poulsen's original magnetic recordings were made on steel wire. The wire was wound around a cylinder, and the recording head tracked the wire as the cylinder was rotated (Morton). Due to the small surface area in contact with the recording head, fidelity was very low. In the early 1930s steel tape (with its greater surface area) replaced steel wire, improving the sound quality of magnetic recorders. At this time, the British Broadcasting Corporation (BBC), the Canadian Broadcasting Company (CBC), and the Reichs-Rundfunk-Gesellschaft (RRG – the German Broadcast Company) were recording their programs on steel tape.

### **3.2 REEL-TO-REEL RECORDERS**

The first plastic-based magnetic tape was produced by BASF in the 1930s. The inspiration for the plastic tape was an iron oxide coated paper tape that was patented by Joseph O'Neill in 1926 (Bruck). The tape was an improvement in fidelity, weight, size, and safety. In 1935, the Magnetophon K-1 was introduced in Germany by AEG as the first practical magnetic tape recorder. The K-1 was a *reel-to-reel* recorder which means that the tape is stored on an open reel and fed onto another reel during playback. All early magnetic tape recorders were reel-to-reel recorders.

In 1939, the technique of AC biasing was accidentally discovered by AEG. Up until this point, when a DC biasing signal was added to an incoming signal there was a slight improvement

in signal quality. However, the addition of an AC signal with a frequency greater than 40 kHz greatly increased the sound quality. AC biasing pushed the incoming signal into the linear response of the tape's transfer function without affecting the frequencies audible to the human ear.

### **3.3 4-TRACK AND 8-TRACK CARTRIDGES**

The next major advance in the area of magnetic recording was inspired by the desire to make recordings more realistic. In the 1930s, experiments in recording two channels of audio simultaneously led to the development of stereo sound. The first commercial stereo recordings were made on reel-to-reel tape by placing two tape heads next to each and recording multiple tracks on the tape. With the advent of the Muntz Stereo-Pak, commonly known as the 4-track cartridge, two sets of stereo tracks were placed side by side on the tape. These tracks were either used directly as two stereo tracks or as four separate monaural tracks.

The 4-track is the first form of magnetic tape recording that does not use the open reel-to-reel design. Instead, the 4-track uses a continuous loop of tape that is housed in a plastic cartridge. The same basic physical construction is also used in the Super-8, also known as the 8-track. The main difference between the 8-track and the 4-track is that the 8-track sacrifices potential fidelity to double the data tracks on the tape (eight tracks must fit where four tracks originally fit).

### **3.4 COMPACT CASSETTE TAPES**

While the 8-track gained its popularity by becoming a must-have accessory in the automobile industry, the compact cassette tape was relegated to the office dictation market (Morton). The compact cassette features a smaller plastic package with an embedded reel-to-reel tape sys-

tem. The fidelity of the compact cassette was considered too low to be taken seriously in the mainstream market with its narrower tape strip. In the late 1960s, Ray Dolby introduced his noise reduction compression techniques that were originally developed for professional recording to the format. Equipped with Dolby NR Type A, the cassette became the de facto standard for audio recording until the arrival of the CD in 1982.

## **4 PRESERVATION IN THE WAKE OF DETERIORATION**

### **4.1 A NATIONAL DILEMMA**

At the close of the cassette tape revolution, when the preeminence of the digital carrier left the analog recording format as a remembrance of a stepping stone technology, recognition of preserving priceless recordings emerged with the understandings that the quality as well as the material archived on magnetic tape would not last forever. In 2000, the U.S. Congress made a commitment to see through the sustenance of those recordings that held historical and cultural significance to our lives in its concern that these materials be continued as part of our heritage. Countless personal and cherished recordings held by everyday citizens would also be desired for advancement through time by their owners. Congress furthered its preservation initiative by organizing The National Recording Preservation Board (NRPB). This entity, in cooperation with the Library of Congress, would help to establish the most effective approach for tape reel to digital file transitions and develop the best methods of copying salvageable media as part of a national plan. The most prominent audio engineers and preservation specialists in the country would come together to discuss current restoration methods and agree upon the best techniques of transferring vulnerable analog recordings to a safe digital haven (NRPB, 2006).

## **4.2 ANALOG TAPE RESTORATION**

The following information for the best practices and procedures for transferring endangered analog tape was developed by the participants of a meeting requested by the Library of Congress and the NRPB that was held in January, 2004. These participants made up a select but diverse pool of audio engineers and preservation specialists. The information was outlined in a report published by the Library of Congress and the NRPB.

## **4.3 TAPE CONDITION**

A nondestructive visual inspection of the magnetic tapes condition and material type will expose the user to what types of physical problems they may have to deal with. The tapes composition will determine what measures must be taken for a successful transfer. Specific physical conditions of the tape have been outlined, each one needing to be addressed with care. Brittleness affecting acetate and paper tape must be recognized for proper handling. Splices, which pose the biggest obstacle to quality transfers, must be repaired and cleaned of any residues before the tape is processed. Cupping, most common to improperly stored acetate tape, is a “deformation in which the tape, when viewed end-on, appears curled instead of flat” (NRPB 19). The best solution for mitigating cupping is to store the tape in its B-wind configuration (oxide sides exposed on outside and the backing facing inward toward the reel) for up to six months. Damage or wear on the edges of the tape can be combated by using a narrower track head although this practice will compromise the signal. If the carrier includes a plastic leader, the user may want to replace it with a paper leader while repairing splices as plastic leaders build up electrostatic charge over the tapes use. Oxide loss on the tape, called blocking, and magnetic tape shedding may also occur but cannot really be fixed, only prevented by proper storing measures. Debris and mold

accumulation on the surface of the tape can be carefully removed by special vacuuming techniques. A sticky tape surface will render a disrupting squealing sound during playback but can be remedied by baking the material at low heat in a convection oven.

#### **4.4 PLAYBACK**

Playback equipment must be configured or calibrated to meet the needs of the medium being recorded. The speed of the playback, which can nominally be set to 7.5 inches per second, can be adjusted for optimum signal collection with pitch being a byproduct of this setting. A user's ear and calibrated test equipment should actively seek to detect abnormalities in the signal. "Tools are available to assist in distinguishing transient, event-based anomalies (e.g. clicks and pops) from global anomalies that may affect the entire tape (e.g., bandwidth and dynamic-range limitations)" (NRPB 22).

#### **4.5 POST TRANSFER**

Once the signal has been transferred, frequency and spectrum analyzers can be used to obtain knowledge on the frequency information of the recording and also be used to locate artifacts or abnormalities like hum or rumble. Software based tools can also be used to identify these unwanted characteristics.

## **5 DIGITAL FORMATS**

Although the compact disc first introduced digital recording to the mainstream audience, the prevailing format in the current market is the MP3. The physicality of the disc format has become a limiting factor in its usefulness. The CD has become a high bandwidth, high quality

(and correspondingly high cost) form of digital audio transmission. The market penetration of various personal digital audio players has mandated the use of lossy compression codecs that result in greatly reduced file sizes with relatively little perceived quality loss.

## **5.1 COMPACT DISCS AND THE RED BOOK FORMAT**

The standard compact disc measures 120 mm in diameter, is 1.2 mm thick, and is made of polycarbonate plastic. Typically, an aluminum layer is then applied to one side of the disc and then spin-coated with lacquer. Data is stored on the CD by a series of “pits” and “lands” that are read by a 780 nm (near infrared) laser. These two features reflect back a different intensity of light that is read by a photodiode. A transition from one feature to another—either a pit to a land or a land to pit—represents a logical 1, while the lack of a transition—either a pit to a pit or a land to a land—represents a logical 0 (NRZI or Non-Return-to-Zero, Inverted encoding).

The logical format for data storage on an audio CD is the Red Book Format. In fact, the CD Digital Audio logo, as seen in Figure 4, can only be used on CDs that follow the Red Book standard, which is derived from the color of the cover of the specification released jointly by Sony and Phillips (Jnavas, 2007). There are several error prevention and correction techniques employed that help account for the minute feature size of the disc (pit width is 500 nm and length varies from 850 nm to 3.5  $\mu\text{m}$ ). Eight-to-Fourteen Modulation (EFM) is used prevent a single pit from being coded in a sequence of lands, and vice versa, while cross-interleaved Reed-Solomon coding (CIRC) adds a redundant parity bit to every 3 data bits.



Figure 4 – Compact Disc Digital Audio Logo

The audio file itself is stored as two tracks of 16-bit PCM (Pulse-Code Modulation) encoded at a frequency of 44.1 kHz. PCM is explored further in section 6.1.

## 5.2 THE WAVE FORMAT

Wave files store audio analogous to CDs using PCM. However, wave files, which often have the .wav extension, are not bound to a physical medium, thus reducing the overhead required for storage. The structure of a Wave file consists of the RIFF file header, which includes the type of file (Wave), the length of the file, the sampling frequency, and the bit-depth, followed by the PCM encoded audio signal. Unlike CDs, Wave files can be monaural, stereo, or contain any number of channels.

## 5.3 LOSSY AND LOSSLESS CODECS

The major disadvantage of the straightforward Wave format is the size of the file. One solution to this problem is to compress the audio signal. Audio compression can be achieved using either a lossy or lossless codec (**coder-decoder**). Lossy codecs, such as the MP3 (MPEG-1 Audio Layer 3), use a psychoacoustic algorithm that discards information that is generally undetectable by the human ear. These algorithms are termed *lossy* due to the fact that part of the sig-

nal information is discarded. The signal reproduction with a lossy codec is dependent on the bit rate used during compression.

Lossless codecs, such as FLAC (Free Lossless Audio Codec) use compression algorithms that are optimized for audio that function similar to the standard .zip file. For example, FLAC uses linear prediction to reduce sample size, then encodes the new sample representation using Golomb-Rice encoding.

## **6 ANALOG-TO-DIGITAL CONVERSION**

The analog-to-digital conversion process involves reading the signal from the magnetic tape, capturing a value at each sampling period, and encoding these samples in the digital file format. The procedure used for this project required a cassette tape player with a line-out jack. Other output jacks, such as the headphone jack, often output an amplified signal since it is expected to drive speakers. The signal was then passed through a computer sound card. Since the sound card used did not offer a line-out jack, the microphone jack was used. However, the first attempt to record the signal suffered severely from clipping. This problem was resolved by disabling the preamp on the microphone. The open source tool Audacity was used to record the signal.

Using the same methodology behind the sampling rate of a CD, 44.1 kHz was used for this project. With a Nyquist frequency of 22.05 kHz, this sampling rate effectively captures the entire frequency range of human hearing. A bit depth of 16-bit was used because this was the highest bit depth possible with available equipment. The purchase of another sound card capable of recording 32 bit was considered, but due to the quality present in the original magnetic record-

ings, it was determined that 16-bit was sufficient.

## 6.1 PULSE-CODE MODULATION

Pulse-Code Modulation (PCM) is the encoding that is used to store the audio signal in CDs and Wave files. Using 16-bit PCM, each sample is stored as a signed integer with a value ranging from -32768 to +32767. This value represents the signal for the entire period until the next sample is taken; PCM is a software representation of a zero-order hold (ZOH). Figure 5 shows an example of unsigned 4-bit PCM (Ktims, 2006).

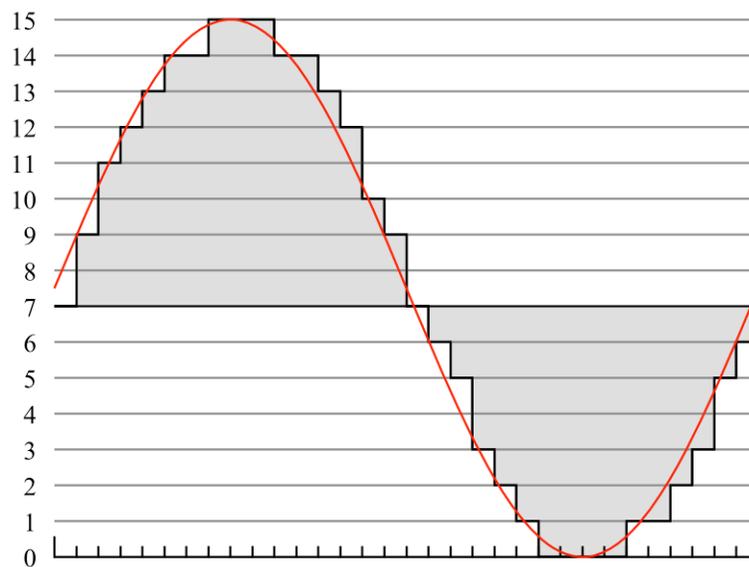


Figure 5 – Pulse-Code Modulation Example

In the above figure, the red function is the analog signal and the gray shaded region represents the digital signal encoded using PCM.

## 7 DIGITAL FILTERING

## 7.1 OVERVIEW OF DIGITAL FILTERING

Digital filtering is the process by which a stream of digital information is passed through a filter in order to modify some aspect of the information. To achieve digital filtration, a digital signal is convoluted with a filter represented by a mathematical model. In time-domain this process can be confusing and unintuitive, but in frequency-domain it becomes very straightforward. Time-domain representation of a digital signal is how one perceives the signal as time progresses, whereas frequency-domain representation of a signal is a spectrum of all the different oscillation frequencies contained in the given signal. The bridge connecting these two domains is called the Fourier transform, and in digital domain is represented mathematically as

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-i\omega n}.$$

where  $x[n]$  is the digital signal in time-domain and  $X(\omega)$  is the signal in frequency-domain. As previously stated, it is highly desirable to do as much design as possible in the frequency-domain, because of the level of intuition that comes with it.

Digital filters, when designed properly, allow us to target very specific frequency regions of a signal and either amplify or suppress them, which becomes very useful for the specific application of noise reduction. The best thing about designing digital filters in frequency-domain is that the convolution used in time-domain to mix the filter and the signal becomes a much simpler multiplication process. To illustrate this, consider a random signal in frequency-domain with a spectrum as shown in the figure below. Now consider filtering this signal with a low-pass digital filter, which is a filter designed to pass only low frequencies. The two signals are simply multiplied together in frequency-domain, and the resulting output is shown below in Figure 6.

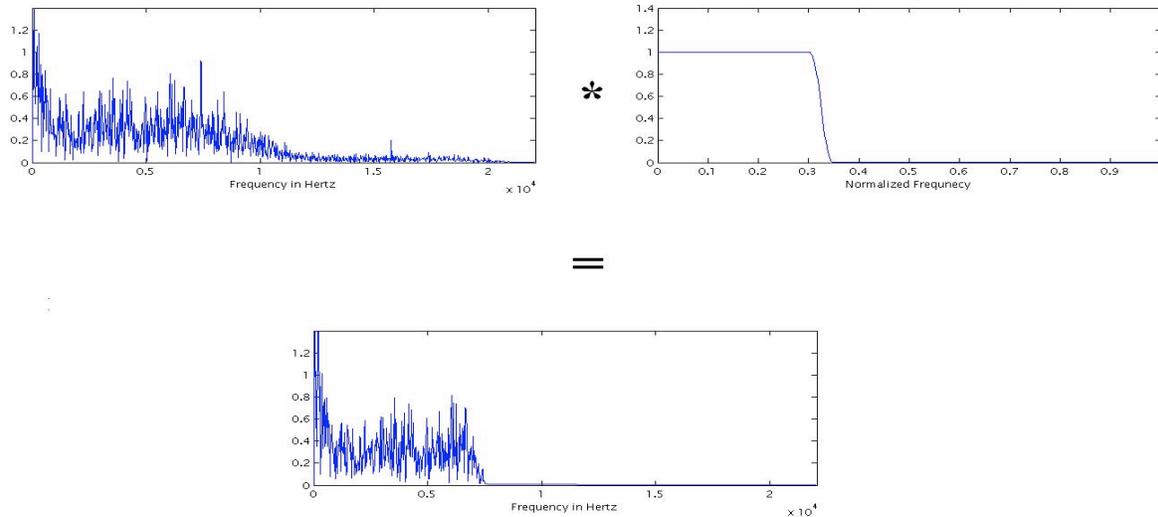


Figure 6 – Filtering Operation in Frequency Domain

In this picture, the left picture shows the digital signal in frequency-domain, and the right picture shows the digital filter in frequency-domain. The bottom picture shows the result of the filtering process, where the two signals are multiplied together. Notice that where the filter is 1, the signal is completely preserved and where the filter is 0 we lose all the corresponding information for the signal. This is the cornerstone property for all filtering, both digital and analog. By properly designing a filter, one can selectively remove unwanted frequencies while preserving wanted ones.

## 7.2 DESIGN OF THE FILTER

There are many different ways in which a filter can be designed, so many in fact that there exist volumes upon volumes of literature on the subject. At the core of the subject, however, is the mathematical representation of a digital filter by a mathematical equation. This equation can take many different shapes and forms, from small and simple to massive and complex. The equation of the digital filter in time-domain can be represented in frequency-domain through

the use of the discrete-time Fourier transform, which was shown earlier. The frequency-domain representation of this equation contains two very important properties of digital filters: the magnitude and phase.

### 7.2.1 MAGNITUDE

The magnitude of a digital filter tells us exactly which frequencies are amplified, and which ones are suppressed. This is illustrated by the picture of the digital filter from the previous section in this paper. In this image, the magnitude of the filter is 1 up until a frequency of  $0.3\pi$  radian/second, and then falls down to magnitude 0 by the time we reach a frequency of  $0.35\pi$  radian/second. When designing a filter, the first thing one should consider is which frequencies are desired and which are not. From this point, one can set requirements for the magnitude of the filter for the entire frequency spectrum to selectively amplify and suppress different frequencies, just like in the picture above. The regions where frequencies are passed (i.e. where the magnitude is 1) are referred to as the passband regions. The regions where frequencies are suppressed (i.e. where the magnitude is 0) are referred to as the stopband regions. The regions where the filter transitions from one magnitude to another (for the illustration above, consider the region between  $0.3\pi$  and  $0.35\pi$ ) are referred to as the transition regions. The parameters of the magnitude of a digital filter are all specified in terms of these three regions, and specifying these parameters is the first step to designing the proper filter.

### 7.2.2 PHASE

The second aspect of a digital filter is its phase. The phase of a filter determines how different frequencies are delayed when perceived in time. For digital audio processing, it is crucial

that one maintains what is referred to as linear phase for the filter to work properly. A linear phase ensures what is referred to as constant group delay. This means that all frequencies perceived in time have the same delay applied to them; meaning that the overall signal sound is maintained. If the phase of a filter is not linear, different frequencies will have different group delays, which will in turn cause audio distortion, with severity proportional to how non-linear the phase is. Consider an audio signal passed through a filter with non-linear phase. Before filtering, all frequencies of the audio signal arrived at the appropriate time, as desired. However, passing the signal through this nonlinear filter could for example speed up the arrival of high-frequency components of the signal, distorting the sound, which is undesirable. It is thus extremely important to maintain a linear phase when processing audio signals.

### 7.2.3 SPECIFICATIONS

Having reviewed these two crucial parameters, one must choose an appropriate filter that meets all of the requirements. For this project, the goal is to reduce the noise in an audio signal. Most of this noise is contained in the higher frequencies, so it is understood that a low pass filter must be used to filter out the high frequency noise, while leaving the desired lower frequencies where the sound is located intact. Also, as previously discussed, the filter implemented must have a linear phase to avoid audio distortion, which would completely defeat the purpose of this project.

Digital filters are generally divided into two categories: finite impulse response (FIR) and infinite impulse response (IIR) filters. As the names suggest, the difference between these two filter types is that FIR has a finite response to an impulse applied to the input, whereas the IIR filter can respond indefinitely to an impulse. The benefit of an IIR filter is the small size of this

filter in terms of computations. The downside, however, is that the phase of IIR filters are non-linear, making them undesirable for the field of audio signal processing. FIR filters, on the other hand, can be designed to have a linear phase, but this comes at the expense of filter size. FIR filters are generally an entire order of magnitude larger than IIR filters. Below, find an illustration of an FIR filter compared to an IIR filter.

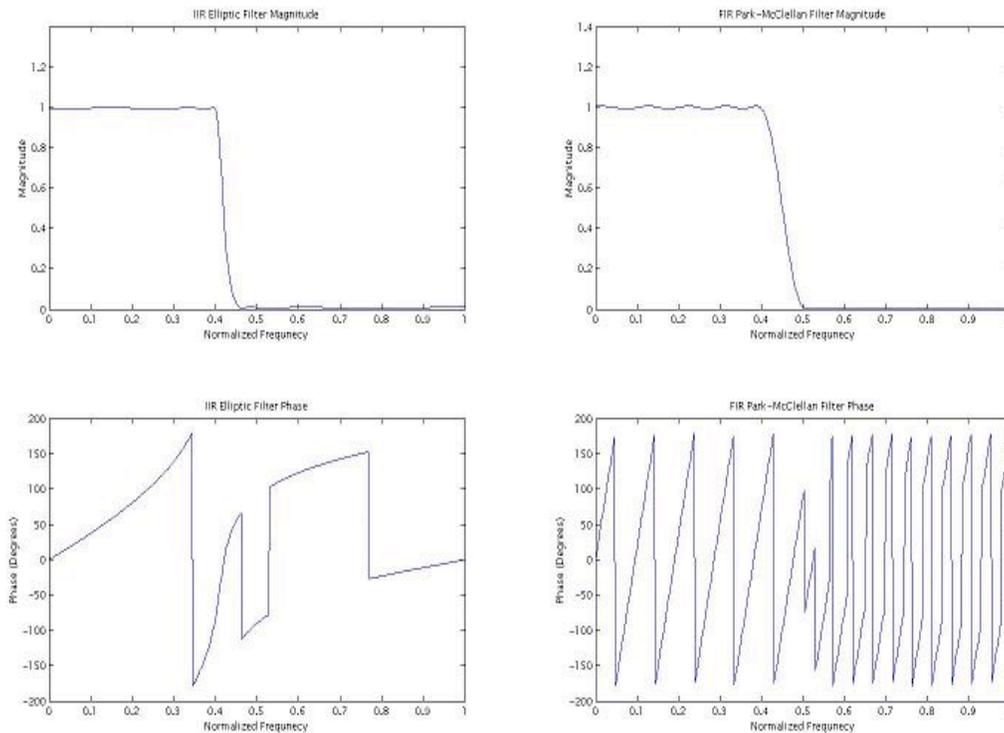


Figure 7 – IIR filter vs FIR filter

In Figure 7 above, the left column depict an IIR filter’s magnitude and phase, on top and bottom respectively. The right column contains a filter of same specification with a FIR implementation. Notice that while the magnitudes of the two filters appear almost identical, their phases look very different. The phase of the IIR filter is clearly non-linear in the passband region, containing what looks like an exponential curve, leading the group delay to be nonlinear.

The FIR filter, on the other hand, maintains a perfectly linear phase in the passband region. Sudden jumps occur from 180 degrees down to -180 degrees, but this does imply a break in linearity since 180 degrees equals -180 degrees on a unit circle.

The price paid for this linearity is, as discussed previously, the size of the filter. The equation necessary to implement the IIR filter is 6<sup>th</sup> order, whereas the FIR filter is 42<sup>nd</sup> order. This is obviously a massive difference, but FIR is necessary for a linear phase, so this will be the filter type of choice for the rest of the project.

The final step in determining an appropriate filter for audio processing is determining how to implement the FIR filter, and then specifying the desired properties of the filter to achieve the desired output. There are different techniques for implementing a FIR filter, and in this project the windowing approach has been used. This technique starts by assuming an ideal filter with a corresponding ideal, infinitely large, equation representing this filter in time domain. To implement this equation in a realistic digital system, a window of the ideal equation is taken to reduce the size. This will inevitably diminish the ideality of the digital filter, but depending on the size of the window chosen one can obtain reasonably close approximations of an ideal filter. To pick the ideal window, this project utilizes the Parks-McClellan algorithm to determine the lowest order window that will meet the desired specifications. This algorithm allows for better performance for lower order filters than other windowing techniques.

### **7.3 IMPLEMENTATION AND TESTING**

For noise removal of audio signals in this project, a completely software-based approach was chosen. Specifically, Matlab, and its accompanying signal processing toolbox was employed to create the digital lowpass filter required for noise reduction. To acquire the digital

audio signal an old cassette was sampled using the sound card and a home stereo system at 44.1 kHz. To more carefully analyze the noise spectrum of the tape, a portion of “silence” was sampled on the tape, where only unwanted noise was present. This noise can then be plotted using Matlab and is shown in Figure 8. Notice that there is a significant amount of noise in the low frequencies, from 0 to 5 kHz. There is also a noise peak around 15 kHz, but one would expect this to be attributed to a particular type of noise, true only to this specific example. It must, none the less, be removed to improve the quality of the sound, something which is easily achieved due to its high frequency.

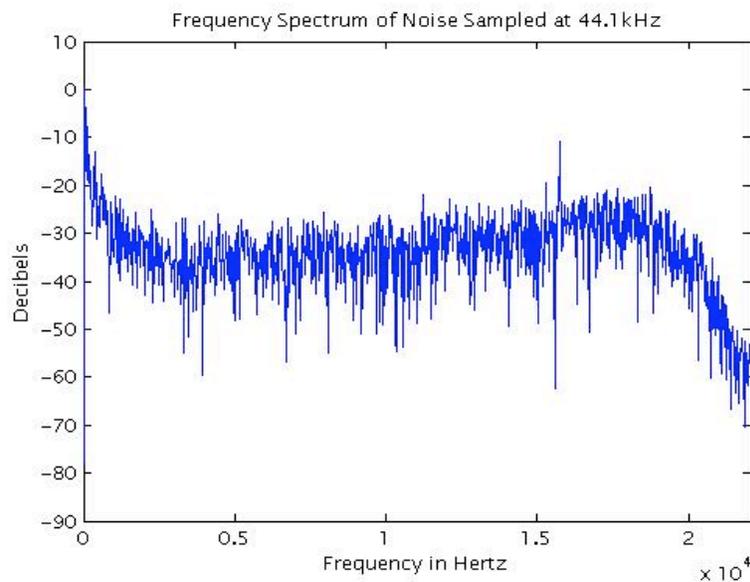


Figure 8 – Noise Sample Frequency Spectrum

Following the plot of the noise, a plot of the audio signal is produced so frequency distributions can be compared and a reasonable filter can be designed to remove as much of the noise as possible. In Figure 9 below is a frequency plot of a sampled voice clip.

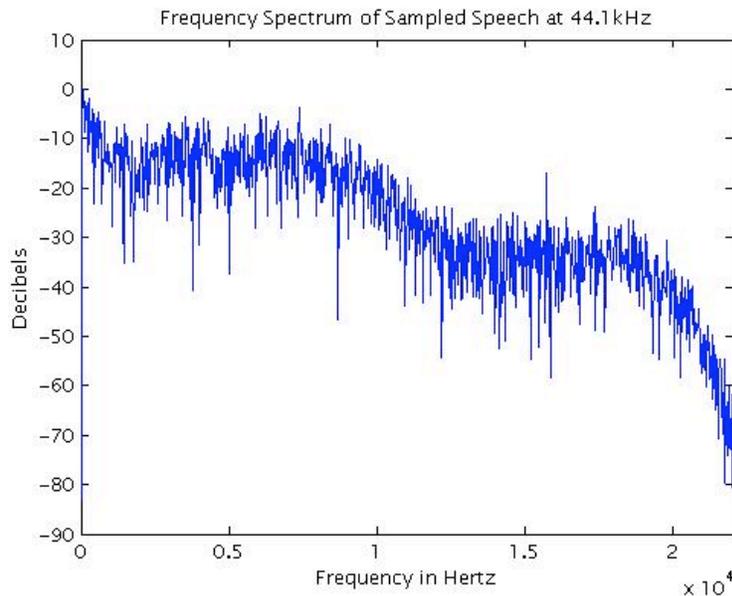


Figure 9 – Speech Frequency Spectrum

Comparing Figure 8 with Figure 9 will provide good intuition regarding where to place the cutoff frequency of the lowpass filter. For this specific example, notice that the majority of the information regarding speech is contained in frequencies less than 10 kHz. It is safe to say that our cutoff frequency should not be any higher than 10 kHz, but to obtain optimal results the best method for picking a cutoff frequency is simply trial and error. The plots are a good guide line, but since this audio clip is ultimately intended for human hearing, using human hearing to discern the best possible cutoff frequency is preferred in the end.

Following analysis of the signal, the appropriate filter is ready to be designed. As previously mentioned, the Park-McClellan algorithm is used in this project to reduce computational time. In addition to the cutoff frequency, ripple in the passband and stopband attenuation must also be considered. It is desirable to have minimal ripple in the passband, and for this filter the ripple is set at 0.01 dB. The stopband attenuation implies how far down the stopband is pushed.

This value should be as large as possible to ensure that the stopband does indeed suppress unwanted frequencies to an acceptable degree. In this filter the stopband is set 60 dB below the passband. After all the requirements for the filter are specified and the appropriate matlab code is executed, a filter as depicted below in Figure 10.

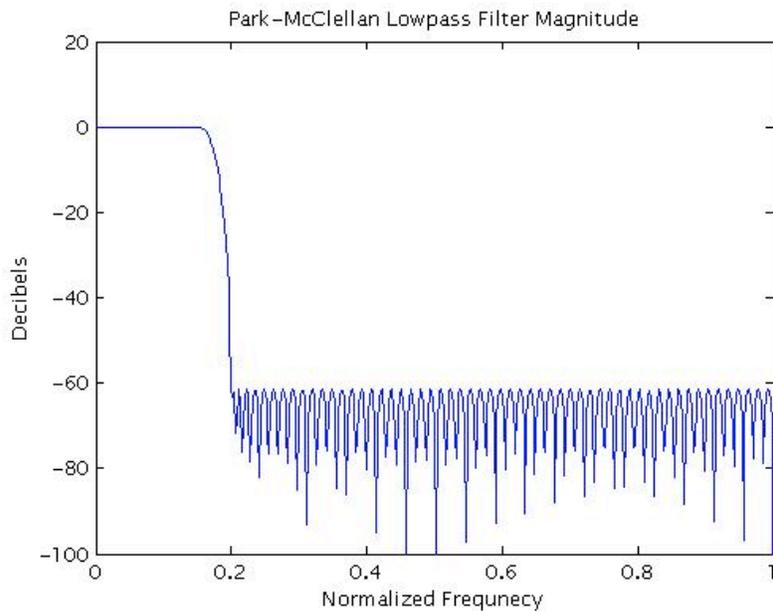


Figure 10 – Park-McClellan Lowpass Filter

In this picture we notice that the cutoff frequency is at  $0.2\pi$ , which is the specification in digital domain. This translates into an analog frequency of approximately 5 kHz. Running the signal through this filter will reduce a significant amount of noise and the end result is plotted in frequency domain, shown in Figure 11 below.

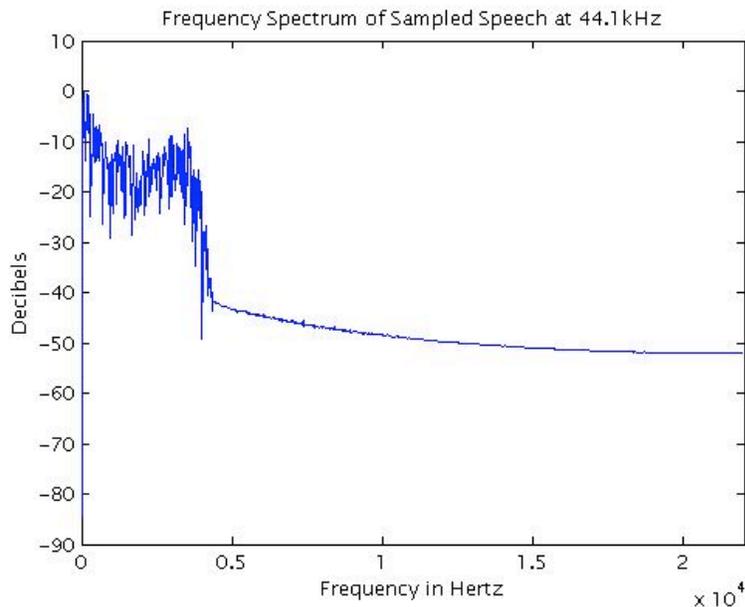


Figure 11 – Filtered Speech Sample

As expected, the lower frequencies are preserved, while the lowpass filter has effectively removed the higher frequencies. This is a very effective method for removing the high frequency noise. The major unresolved problem with this signal is that there is still a good deal of unwanted noise in the lower frequencies. For the scope of this project, this problem is largely irresolvable because the implemented filter can only discriminate between varying frequencies. Noise at the same frequency and the desired sound cannot be removed with a filter like this. Instead, an alternate approach to attempt and reduce some of the low frequency noise is through the use of a time-averaging filter. The basic concept of this filter is to take the average of a select number of values in time for a signal, thus “smoothing the signal” and reducing some sporadic noise. To illustrate this see Figure 12 below; this shows a plot (in time domain) of the speech signal before and after applying the filter. Notice that the signal on the right looks a bit smoother, which effectively reduces some of the noise. The drawback to this is that the overall

quality of the audio signal is reduced, and the sound ends up being muffled, as if someone covered the speaker with a pillow. A time averaging filter is not ideal for every application, but if the goal is to reduce noise, even at the cost of the desired signal quality, it may be appropriate.

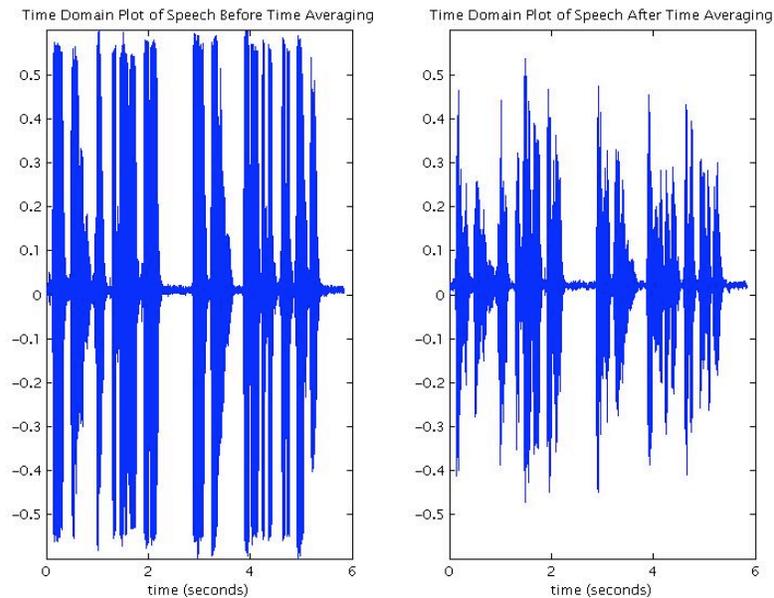


Figure 12 – Speech Sample Before and After Time Averaging Filter

As a final step, a sample of music is filtered using the same lowpass filtering technique as described previously. The process is the same, and a picture of the filtering process can be seen on the next page in Figure 13. Note the spikes at higher frequencies of the original signal, which are the source of the noise in this file. By passing the data through a lowpass filter, these spikes are removed, resulting in a better sounding audio clip.

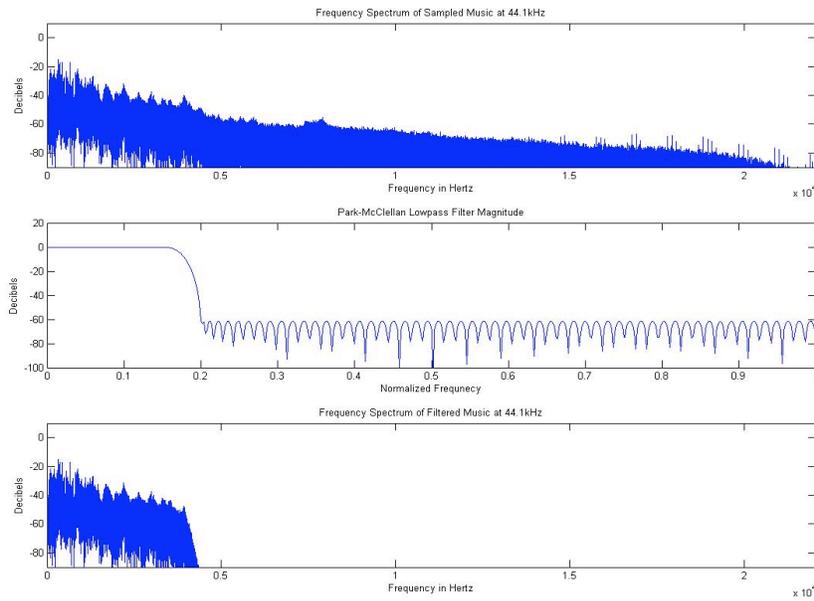


Figure 13 – Frequency Range of Music before and after Lowpass Filter

## 8 ETHICAL CONCERNS

In the internet age, piracy of digital music is a major concern. However, the preservation of existing audio does not necessarily fall under piracy. The techniques presented in this paper are aimed towards personal archival use. Although prerecorded music may be preserved in this manner, it does not represent opposition to CD sales. In almost all cases, a professionally done digital master will sound better than a filtered sample of a magnetic tape recording, negating the argument of piracy. One must also consider that actual act of copying that occurs during the analog-to-digital conversion. Although a copy is made, it is for personal archival use of an owned cassette tape which falls under fair use.

## **9 ECONOMICS AND MANUFACTURABILITY**

Many tools and services already exist today to help realize the analog-to-digital transfer process. As people become more aware of the necessity to convert their vulnerable media, these outlets will be acknowledged more on the market and become more familiar to consumers.

There is enough capability in most equipment found at home to assemble a system to record magnetic tape data onto a PC. Open source software tools can be easily downloaded and implemented for filtering or other equalization operations on the recorded files. Ion's Tape2PC cassette deck and software kit makes it easy for anyone to execute the transfer process through a familiar USB hook-up. Services are available to do the work for you at a professional level, allowing one's media to be mastered to its full potential. There seems to be enough representation already with respect to equipment available to help make reproductions in a cost effective manner.

However, there is not enough communication with the public concerning the state of deterioration of cassette tapes. The market potential of any equipment, services, or guidebooks could potentially draw large financial gains based on the quantity of existing cassette tapes. While it is possible that repurchasing music that has been released in a digital format is more appealing than converting and processing old cassette tapes, there is no solution for personal or archival recordings. The NRPB was formed because there are many recordings that only exist in a magnetic format.

## **10 CONCLUSION**

Digitally sampling and filtering recordings on compact cassette tapes protects them

against further degradation while also removing several artifacts and high frequency noise. The technique employed for this project demonstrates how this process can be accomplished with tools that freely and/or commonly available. While the results of this process will never approach the quality of a digitally mastered recording, they do effectively preserve the original audio signal. The filters used were designed for the particular samples and will not necessarily work for all recordings. Future work could be done to generalize the filters and create an interface to allow user-defined filter parameters. Another approach is to design an adaptive filter that samples a “silent” portion of a recording to build a noise profile which is then used to define the filter parameters.

## APPENDIX I – MATLAB CODE

```
%Plots a frequency spectrum a noise sample.

x1 = wavread('backgroundnoise3.wav');
y1 = fft(x1,2048);
y1 = fftshift(y1);

%Noise Frequency Spectrum

f = [-22050:22050/1024:22050-(22050/1024)];

mag1 = abs(y1);
mag1(1025) = 0.0001;
db1 = 20*log10(mag1);
subplot(2,2,1);
plot(f,db1-max(db1));
axis([0 22050 -90 10]);
title('Frequency Spectrum of Noise Sampled at 44.1kHz');
xlabel('Frequency in Hertz');
ylabel('Decibels');

%Sample Speech Frequency Spectrum

x2 = wavread('voicesample3.wav');
y2 = fft(x2,2048);
y2 = fftshift(y2);

mag2 = abs(y2);
mag2(1025) = 0.0001;
db2 = 20*log10(mag2);
subplot(2,2,2);
plot(f,db2-max(db2));
axis([0 22050 -90 10]);
title('Frequency Spectrum of Sampled Speech at 44.1kHz');
xlabel('Frequency in Hertz');
ylabel('Decibels');

%PM Filter Design

f = [0.15 0.2];
m = [1 0];
rp = 0.01;
rs = 60;
rp = (10^(rp/20)-1)/(10^(rp/20)+1);
rs = 10^(-rs/20);
delta = [rp rs];

[N,f,m,weights] = firpmord(f,m,delta);
h = firpm(N,f,m);
[db,mag,pha,grd,w] = freqz_m(h,[1]);

subplot(2,2,3);
plot(w/pi,db);
title('Park-McClellan Lowpass Filter Magnitude');
```

```

xlabel('Normalized Frequency');
ylabel('Decibels');
axis([0 1 -100 20]);

y3 = filter(h, [1], x2);

y4 = fft(y3, 2048);
y4 = fftshift(y4);

f = (-22050:22050/(length(y4)/2):22050-22050/(length(y4)/2));

mag4 = abs(y4);
mag4(1025) = 0.0001;
db4 = 20*log10(mag4);

subplot(2, 2, 4);
plot(f, db4-max(db4));
axis([0 22050 -90 10]);
title('Frequency Spectrum of Sampled Speech at 44.1kHz');
xlabel('Frequency in Hertz');
ylabel('Decibels');

wavwrite(y3, 44100, 'SpeechOutput1.wav')

%Time Average Filter

M = 100;
B = ones(M, 1)/M;

y5 = filter(B, 1, y3);
y5 = y5*2;

figure (2);

t = [0:1/44100:(length(x2)/44100)-1/44100];

subplot(1, 2, 1);
plot(t, y3);
title('Time Domain Plot of Speech Before Time Averaging');
xlabel('time (seconds)');
axis([0 6 -0.6 0.6]);

subplot(1, 2, 2);
plot(t, y5);
title('Time Domain Plot of Speech After Time Averaging');
xlabel('time (seconds)');
axis([0 6 -0.6 0.6]);

wavwrite(y5, 44100, 'SpeechOutput2.wav')

%Song Noise Reduction

%Sample Speech Frequency Spectrum

x2 = wavread('eydie_2.wav');
y2 = fft(x2, 2048);
y2 = fftshift(y2);

```

```

f = [-22050:22050/1024:22050-(22050/1024)];

mag2 = abs(y2);
mag2(1025) = 0.0001;
db2 = 20*log10(mag2);
subplot(3,1,1);
plot(f,db2-max(db2));
axis([0 22050 -90 10]);
title('Frequency Spectrum of Sampled Music at 44.1kHz');
xlabel('Frequency in Hertz');
ylabel('Decibels');

%PM Filter Design

f = [0.15 0.2];
m = [1 0];
rp = 0.01;
rs = 60;
rp = (10^(rp/20)-1)/(10^(rp/20)+1);
rs = 10^(-rs/20);
delta = [rp rs];

[N,f,m,weights] = firpmord(f,m,delta);
h = firpm(N,f,m);
[db,mag,pha,grd,w] = freqz_m(h,[1]);

subplot(3,1,2);
plot(w/pi,db);
title('Park-McClellan Lowpass Filter Magnitude');
xlabel('Normalized Frequency');
ylabel('Decibels');
axis([0 1 -100 20]);

y3 = filter(h,[1],x2);

y4 = fft(y3,2048);
y4 = fftshift(y4);

f = [-22050:22050/1024:22050-(22050/1024)];

mag4 = abs(y4);
mag4(1025) = 0.0001;
db4 = 20*log10(mag4);

subplot(3,1,3);
plot(f,db4-max(db4));
axis([0 22050 -90 10]);
title('Frequency Spectrum of Filtered Music at 44.1kHz');
xlabel('Frequency in Hertz');
ylabel('Decibels');

wavwrite(y3,44100,'SongOutput1.wav')

%Time Average Filter

M = 20;
B = ones(M,1)/M;

```

```

y5 = filter(B,1,y3);

wavwrite(y5,44100,'SongOutput2.wav')

%FIR vs IIR filter

%Lowpass Filter Parameters
wp = 0.4;
ws = 0.5;
Rp = 0.1;
As = 40;

%IIR Filter
[n1,Wp1] = ellipord(wp,ws,Rp,As);
[b1,a1] = ellip(n1,Rp,As,Wp1);
[db1,mag1,pha1,grd1,w1] = freqz_m(b1,a1);

subplot(2,2,1);
plot(w/pi,mag1);
axis([0 1 0 1.4]);
title('IIR Elliptic Filter Magnitude');
xlabel('Normalized Frequency');
ylabel('Magnitude');
subplot(2,2,3);
plot(w/pi,pha1*180/pi);
title('IIR Elliptic Filter Phase');
xlabel('Normalized Frequency');
ylabel('Phase (Degrees)');

%FIR Filter
f = [wp ws];
m = [1 0];
rp = (10^(Rp/20)-1)/(10^(Rp/20)+1);
rs = 10^(-As/20);
delta = [rp rs];

[n2,f,m,weights] = firpmord(f,m,delta);
h = firpm(n2,f,m);
[db2,mag2,pha2,grd2,w2] = freqz_m(h,[1]);

subplot(2,2,2);
plot(w/pi,mag2);
title('FIR Park-McClellan Filter Magnitude');
xlabel('Normalized Frequency');
ylabel('Magnitude');
subplot(2,2,4);
plot(w/pi,pha2*180/pi);
title('FIR Park-McClellan Filter Phase');
xlabel('Normalized Frequency');
ylabel('Phase (Degrees)');

```

## WORKS CITED

Behrman, Alison. Speech and Voice Science. San Diego: Plural Publishing Inc., 2007.

Bruck, Jerry, Al Grundy, and Irv Joel. "An Audio Timeline." 17 October 1999. Audio Engineering Society Historical Committee. 26 February 2008

<<http://www.aes.org/aeshc/docs/audio.history.timeline.html>>

Fitzgerald, A. E., Charles Kingsley, Jr., and Stephen D. Umans. Electric Machinery. New York: McGraw-Hill, 2003.

Iainf. "Optical-Film-Soundtrack.svg." 12 July 2006. Online image. Wikimedia Commons. 25 February 2008. Licensed under Creative Commons Attribution 2.5 license.

<<http://commons.wikimedia.org/wiki/Image:Optical-film-soundtrack.svg>>

Jnavas. "CDDAlogo.svg." 3 February 2007. Online image. Wikipedia. 10 March 2008. Copyright Philips. <<http://en.wikipedia.org/wiki/Image:CDDAlogo.svg>>

Ktims. "Pcm.svg." 10 May 2006. Online image. Wikimedia commons. 13 March 2008. Licensed under the GNU Free Documentation License. <<http://en.wikipedia.org/wiki/Image:Pcm.svg>>

Miller, Carol Poh. Landmarks in Mechanical Engineering. West Lafayette: Purdue University Press, 1997.

Morton, David. "Overview History of the Technologies for Recording Music and Sound." Recording History. 2006. 24 February 2008.

<http://www.recording-history.org/HTML/musictech1.php>

National Recording Preservation Board (NRPB). Capturing analog sound for digital preservation: report of a roundtable discussion of best practices for transferring analog discs and tapes.

Washington, D.C.: [Council on Library and Information Resources](#) and (U.S.) Library of Congress, 2006.

Nave, Carl Rod. "tape9.gif." 1999. Online image. HyperPhysics. Georgia State University. 25 February 2008. <<http://hyperphysics.phy-astr.gsu.edu/hbase/audio/tape2.html#c1>>

Smith, Steven W. The Scientist and Engineer's Guide to Digital Signal Processing. San Diego: California Technical Publishing, 1999.