HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and Communications Engineering
Laboratory of Acoustics and Audio Signal Processing

Sira González González

Digital Audio Antiquing

Final project.

Espoo, June 7, 2007

Supervisor: Professor Vesa Välimäki

HELSINKI UNIVERSITY OF TECHNOLOGY	ABSTRACT OF THE MASTER'S THESIS

Author: Sira González González

Name of the thesis: Digital Audio Antiquing

Date: June 7, 2007 **Number of pages:** 54

Department: Electrical and Communications Engineering

Professorship: S-89 Acoustics and Audio Signal Processing

Supervisor: Prof. Vesa Välimäki

Historical music recordings have many different types of degradations. These deficiencies are introduced both in recording and reproduction techniques, worsening the quality of the final music file.

In this thesis, a study of these degradations is conducted, analyzing their sources and characteristics. The implementation of the defects is carried out using different audio signal processing methods each time, taking into account the nature of the degradation, the available data and computational resources. Finally, these synthetic degradations are applied to a CD-quality music file in an attempt to simulate phonograph, gramophone and LP recordings.

Keywords: acoustic signal processing, degradations, digital filtering, historical recordings, interpolation, music, restoration

Acknowledgements

This thesis work has been carried out in the Laboratory of Acoustics and Audio Signal Processing of TKK, Helsinki University of Technology.

First of all, I would like to thank my supervisor, Prof. Vesa Välimäki, for giving me this opportunity and for his help during the whole working process. Also, I wish to thank Ossi Kimmelma and Jukka Parviainen for providing me much with helpful material and for their patience while explaining all their code to me.

Finally, I would like to thank all the people who have been put up with me during these past months.

Otaniemi, June 7, 2007

Sira González González

Contents

Abbreviations

CD	Compact Disc
DA	Digital Analog
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
IFFT	Inverse Fast Fourier Transform
LMS	Least Mean Square
LP	Long Playing
pdf	Probability density function
PPCC	Probability Plot Correlation Coefficient
Rp	Pass-band ripple
Rs	Stop-band attenuation
SNR	Signal to Noise Ratio
Wp	Pass-band corner frequency
Ws	Stop-band corner frequency

Chapter 1

Introduction

Since the sound waves were first stored on a piece of paper by the Phonautograph in 1857, the evolution of recording methods has reached unthinkable levels. Meanwhile, the recorded music was undergoing a transformation which was getting the most out of the possibilities offered by the developing techniques. The advance made in audio signal processing has allowed methods to reduce or increase the dynamic range of the signal to can adapt it to the available in the recording medium or to use electronic music to replace an existing instrument or to create new sounds. These changes have introduced new characteristics in the audio signal both in time and frequency domain.

A huge study has been made for restoring old recordings, in an attempt of having them available with a reasonable sound quality due to the increase of interest in historical and nostalgic material. Nevertheless, there have been none or little interest in the simulation of the degradations present in these recordings, most of it only for testing restoration algorithms.

Sibelius Museum, in Turku, Finland, had the idea of exhibiting the history of recording showing to the audience a song in phonograph, gramophone, LP and CD quality. By this way, people could listen and thus understand the progressive evolution made from one technique to another. The initial idea of the exhibition was to have available two different songs in all the qualities. One of the songs would be an old song they had already in the four mediums and the other one would be an actual song in which the degradations of historic recordings would be simulated. The two chosen songs were ‘Valse Triste’ from Sibelius and ‘In the Shadows’ by The Rasmus. ‘Valse Triste’ was selected because of the importance of Sibelius in Finnish music history and the reason for the second one was that The Rasmus was the first Finnish group selling more than Sibelius. When the ‘Laboratory of Acoustics and Audio Signal Processing’ of TKK (Helsinki University of Technology) was proposed for the fulfillment of the defects simulation the interest of the project was analyzed. The simulation of all degradations historical recordings have starting from CD-quality seemed to be an interesting topic to learn about the evolution in recording methods as well as in the recorded music. In fact, the chosen song for the simulation have some signal processing typical nowadays, like the use of a compressor to make the original dynamic range suitable for the recording medium and the presence of high frequencies due to unexpected changes in the waveform.

Some similar projects were done in the past, like the one implemented for the ‘Music Exhition of Heureka Science Center’ in 2005 by Jukka Parvaianen and Ossi Kimmelma (Kimmelma, Parviainen and Välimäki, 2005). Besides, a free plug-in, iZotope Vinyl, can be downloaded to simulate a ‘vinyl’ recording (iZotope, 2007). However, there were still some tasks to improve in both of them. The work done for Heureka exhibition had good results, but the way of carry out the simulation was not really scientific, adjusting the parameters based on the listening test more than extracting them from the historical recordings. On the other hand, one of the drawback of the available plug-in is that is focused on ‘vinyl’ recordings, being not the most appropriate for other simulations. Apart from that, some problems can be found when reproducing with iZotope Vinyl some of the degradations. For example, while low frequency pulses appear to be additively superimposed upon the undistorted waveform, in this tool the audio waveform is just deleted and replaced with zero samples.

The final goal of this project is to process a CD-quality audio signal to simulate the quality of a wax cylinder, a gramophone disk and a vinyl disk would provide. For that, the algorithms needed to carry out each degradation present in old recording should be implemented.

Along this thesis, the process made for the digital audio antiquing is explained. In Chapter 2 an overview of the history of recordings is presented, Chapter 3 is based on the explanation of the different degradations that can be found in historical recordings and Chapter 4 in their simulation. Finally, in Chapter 5 the parameters used for the final audio files are presented and the conclusions are made in Chapter 6.

Chapter 2

Overview of the History of Recording

Sound recording history started when Frenchman Léon Scott discovered the way to display voice waveforms on a piece of paper with a device called phonautograph in 1857 (see figure 2.1). It consisted of a sharp point attached a diaphragm situated at the end of a cone-shaped speaking horn. The sharp point was touching the surface of a piece of paper, so when someone shouted down the horn the vibration of the diaphragm was reflected in the soot of paper as a squiggly line (IEEE Virtual Museum, 2007).

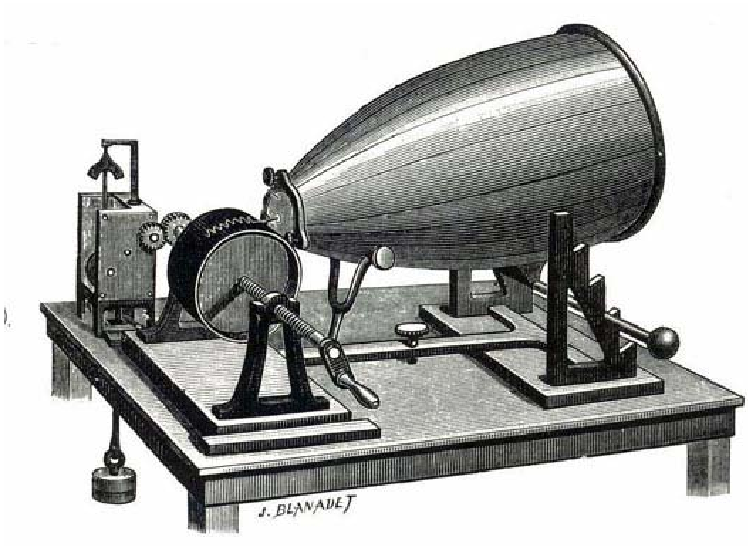


Figure 2.1: Leon Scott's phonautograph.
(http://members.aol.com/antiquephono/phonauto.htm)

However, it was not until 1877 when a device for storing and reproducing sound was invented, the phonograph (figure 2.2). Created by Thomas Edison, this device was able to store sound in a grooved tin foil rotating cylinder using a steel point stylus which cut vertically into it. An inverted horn whose final part was closed by a metallic diaphragm transmitted the movements caused by the sound to the stylus. For the reproduction, a system similar to the storing one in reverse order was used. The needle was placed at the beginning of the recorded groove and when turning the cylinder the oscillations of the needle arouse in the membrane some vibrations which were turned into sound waves amplified and channeled by a horn (Colección F.B., 2007b). Further developments of the invention replace the tin foil drum with a beeswax cylinder, which provided much better sound quality. The commercialization of the phonograph with this wax cylinder began in 1888, the same year as the invention of the gramophone by Emiler Berliner. The phonograph presented important difficulties for commercial exploitation, having no inexpensive method for duplicating the cylinders until 1902 (Godsill and Rayner, 1998).

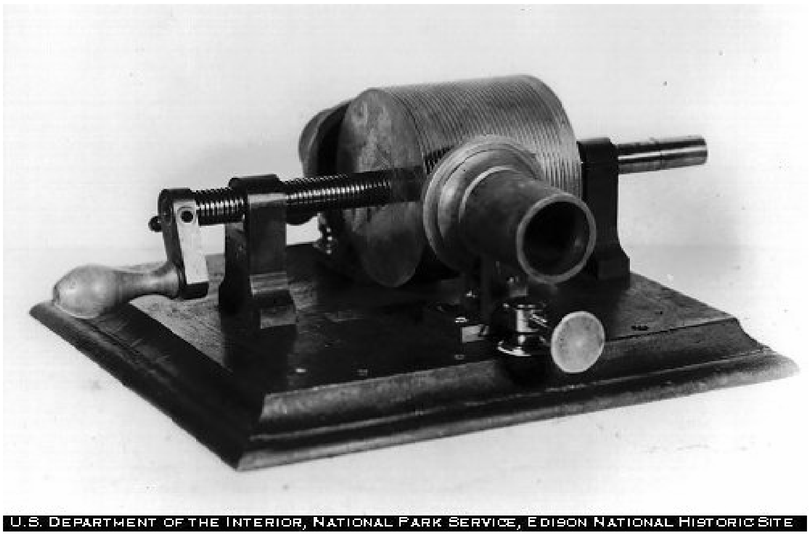


Figure 2.2: Original Edison Tin Foil Phonograph.
(http://history.sandiego.edu/gen/recording/pix2/29110013.jpg)

The gramophone (figure 2.3) solved some of the problems the phonograph had. This device used a flat disk record instead of a cylinder, using an acid etching process to cut grooves into the surface of a polished zinc plate. Berliner discovered a method for creating a metal negative master disc from which hundreds of shellac disks were pressed. Moreover, the flat disk allows more precise mechanic arrangement as the disks were controlled by the fitting of the center hole. The sound was registered in the spiral grooves of the flat record, but instead of using vertical movements of the needle tip, lateral movements were used (figure 2.4), making the friction between the needle and the groove minimized and the membrane vibrations more precise (Colección F.B., 2007a). Shellac was used for 78 rpm recordings until vinyl was invented in the mid-twentieth century. When a motor driven gramophone was introduced, with a more sensitive soundbox and a new wax disk recording process, the gramophone became a huge commercial success (Godsill and Rayner, 1998).

Figure 2.3: Gramophone player
(http://www.wpclipart.com/music/gramophone_2.png)

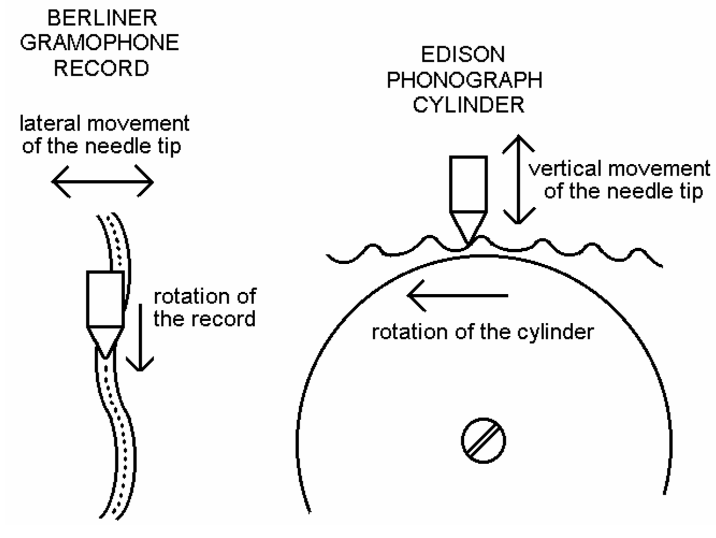


Figure 2.4: Gramophone and phonograph recording techniques.
(http://www.soundfountain.com/amb/ttadjust.html)

In both phonograph cylinder and gramophone discs the sound was amplified by a horn which was attached to the membrane, being a difficult and unreliable procedure. This system limited the range and type of music, having a poor bandwidth (164 Hz – 2000 Hz) which did not allow the appropriately recording of some instruments. The arrival of electrical recordings in 1925 using the Western Electric system developed at AT&T's Bell Labs over the previous 10 years drastically improved the quality of the recording process. Electrical recordings involved an electrical microphone and amplifier which actuates a cutting tool, improving the poor frequency response of the system (20 Hz – 5000 Hz) and reducing the surface noise (Godsill and Rayner, 1998). This technique made possible to record whole orchestras and symphonies.

In 1931 Alan Blumlein of EMI patented stereo, and a few years later the first stereo test dics were produced. However, it was not used commercially until a quarter of century later. In 1935 AEG (Germany) exhibited at the Berlin Radio Exposition its "Magnetophon" Model K-1, the precursor of the magnetic tapes (Museum of Sound Recording, 2003). By 1945 the first stereophonic tapes arrived, but it was not until 1947 when the modern form was developed. The magnetic tapes provided finally the needed practical way to edit recordings, including the ability to 'cut and splice' the tape for restoration purposes. This recording technique had the important advantage of not having the problems of the disk and cylinder recordings, like clicks, thumps, etc. Meanwhile, the disc recordings were much improved and in 1948 the first vinyl LP (see figure 2.5), a 33 rpm long playing record, was introduced in America. At the same time, the seven inch 45 rpm discs were introduced for popular music. These together replaced the 78 rpm disc, although at the end the 33 rpm long playing records were the predominant ones. In 1958, the first stereo disc was cut. Around this time analogue electronic technology for sound modification by filtering, limiting, compressing and equalization was being introduced, which allowed the filtering of recordings for reduction of surface noise and enhancement of selective frequencies (Godsill and Rayner, 1998).

Figure 2.5: A 1980's record player playing an LP disk

The CD (Compact Disc) was developed by Sony and Philips since the late 1970s and it was available on the market in late 1982. The stereo sound is recorded on a plastic disc in which the digital information is stored as a series of tiny indentations read by a small laser

(Wikipedia, 2007a). The CD brought digital audio processing techniques into homes, including algorithms for decoding, error correction, oversampling, digital filtering and DA converters. The standard parameters used for the recording process include a sampling rate of 44100 Hz (providing a bandwidth up to 20 kHz) with 16 bits resolution. However, higher sampling rates and resolutions are used in recording studios, where the quality of the signal is a fundamental characteristic.

Chapter 3

Degradations in Historical Recordings

Many different kinds of degradations can be frequently found in audio sources. They can be divided into two general groups: global degradations and localized degradations. Global degradations affect all samples of the waveform and contain background noise, wow and flutter and certain types of non-linear distortion. Localized degradations are discontinuities which affect only certain samples in the waveform (Godsill and Rayner, 1998).

3.1 Global Degradations

As it was explained above, every sample of the waveform is distorted in some way due to global degradations. There are distinct types of this kind of degradations, being explained as follows.

3.1.1 Frequency Response Deviation

The frequency response of the loudspeaker or the horn used in music reproduction is an important feature in the final sound. In audio, a high quality system should have a flat frequency response at least in the margin of audiofrequencies (20 Hz - 20.000 Hz). In the ideal case, all the frequencies have the same treatment in the system, so the listened sound is the original one (Wikipedia, 2007b).

However, in practice, the response for different frequencies is not usually the same. A system with an inappropriate response will affect the final quality. For example, if high frequencies are emphasized the final sound will be more trebly and if they are lost it will sound dark.

Although the quality of the actual loudspeakers can be quite good, giving almost the same emphasis to all the frequencies, the same can not be said about the horns used for the gramophone or phonograph music reproduction. The emphasis some frequencies receive with these horns and the attenuation of another makes the final sound have a characteristic touch up.

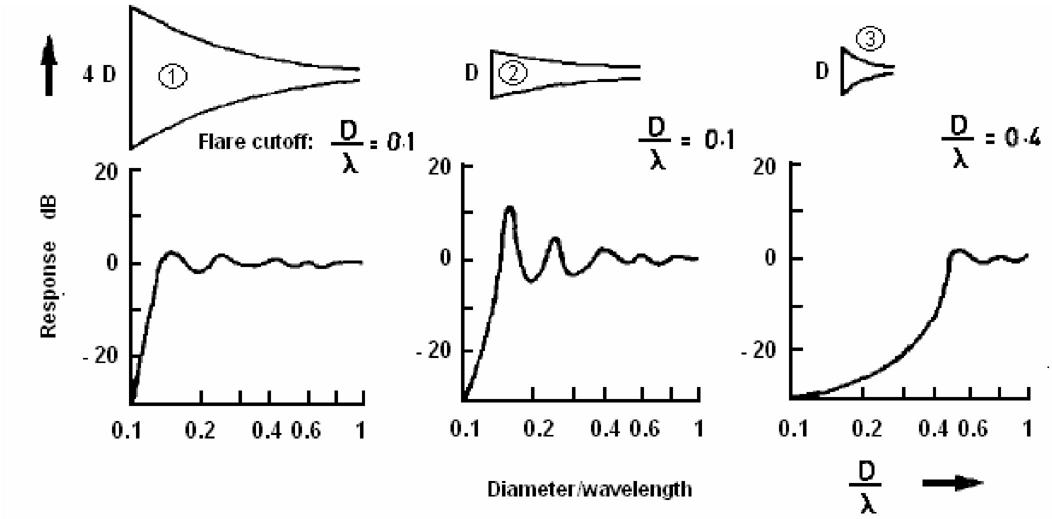


Figure 3.1: Frequency response for different phonograph horns.
Adapted from (Fesler, 1983)

In the figure 3.1 the unequal response some horns have for different frequencies is observed, being thus the final sound distorted. Horn 1 (the largest) is the one with the flattest response, but the election of the reproduction horn depends on more factors. For example, horn 1 might not be appropriate to use with a particular recording diaphragm because it may not cancel its resonant peaks and valleys. In that case, a smaller horn could theoretically cancel the resonant effect of the diaphragm if it is the inverse mirror of the diaphragm (Fesler, 1983).

3.1.2 Signal Bandwidth

The frequency range a human can hear is from 20 Hz to 20.000 Hz. However, it is unusual to find a person who can hear over the entire audible range, varying the range of hearing greatly among individuals. When growing up the human losses sensitivity to sounds of high frequency, so that an adult may have difficulties hearing sound beyond 10.000 Hz or 12.000 Hz (Rossing, Moore and Wheeler, 2002).

The signal bandwidth the used recording technology can store is an important characteristic. Nowadays the CD allows the storage of all the frequencies below 20.000 Hz (the sampling rate used is 44.100 Hz), being the full range of human hearing fulfilled. Nevertheless, in the old recordings the bandwidth of the stored audio signal was narrower, being the lowest and the highest frequencies lost. Although it is difficult to notice the differences in it between a CD and an LP because of the human sensitivity to high frequencies, listening to an older device like a phonograph or a gramophone where the bandwidth is much narrower than the heard frequency range, the lost sounds are quite noticeable.

The bandwidth of the analog recordings (LP, gramophone and phonograph) is limited by the size and speed of the record and the size and shape of the stylus. The low frequencies limit is determined by the compliance of the stylus. On the other hand, the maximum high frequency is defined by the rubbing with the record groove when the stylus attempts to move very rapidly. The hardness of the stylus should be in keeping with the hardness of the material used for the recordings, being different for the wax cylinder than for the vinyl records. Apart from the hardness, the size of the stylus is an essential characteristic for the final signal bandwidth, determining the size of the groove in a record and thus the recordable frequency range. The order of the speed of the record is established by this frequency range (Laser Turntable, 2006).

The differences between the stylus used for the recording and reproduction process have a hand in the final listened bandwidth. The stylus is not expected to follow the surface details which are much smaller than the width of its contact area, so any improvement in resolution obtained by increasing the contact area may purchase at the cost of a reduction in the available signal bandwidth. On the other hand, choosing a smaller stylus would sacrifice resolution for a wider bandwidth (Lesurf, 2001). Moreover, the frequency response of the media is degraded by frequent playback if the cartridge is set to the track too heavily.

3.1.3 Dynamic Range

The dynamic range is defined as the ratio of maximum signal value to the softest or to the noise level. The dynamic range of human hearing is about 120dB, but less can be sufficient for a good quality (CDs have approximately 96 dB). In fact, taking into account the background noise and the dynamic range offered by a CD it can be said the noise is inaudible.

In the case of old recording, the noise level is higher than in the CD and it gets worse due to being scratched easily when playing. The best dynamic range for a gramophone disk can be around 70 dB (Encyclopedia Britannica, 2007).

3.1.4 Distortion

Distortion is a kind of degradation which includes many non-linear defects like amplitude related overload (e.g. clipping) or groove wall deformation and tracing distortion (Godsill and Rayner, 1998).

There are multiple sources from which the distortion can stem from. The stylus used in cutting and reproduction process of the audio signal is an important part for the listened distortion. In this way, the distortion caused by the different stylus can be classified into: cutting distortion, tracing distortion and tracking error (Muraoka, Onoye and Eargle, 1975).

Cutting distortion is mainly due to the shape of the chosen cutting stylus. Overload, a kind of this distortion, is made when the stylus can no longer trace the modulation in the groove (Kogen, 1968), being this limit dependent on frequency (figure 3.2).

Other type of distortion, known as tracing distortion, is made due to the discrepancy between the shape of the pickup and cutting stylus (Trendell, 1977). The difference between the two shapes makes the replay stylus be incapable of following a groove cut by a distinct stylus. Moreover, after some use the needle develops "flat" spots at the points of contact with the groove, not being able to follow in a reliable way the groove modulation, losing high frequency output and creating distortion (Bauer, 1977).

The difference between the cutting angle of the cutting stylus and the tracking angle of the pickup stylus causes the tracking error (Muraoka, Onoye and Eargle, 1975).

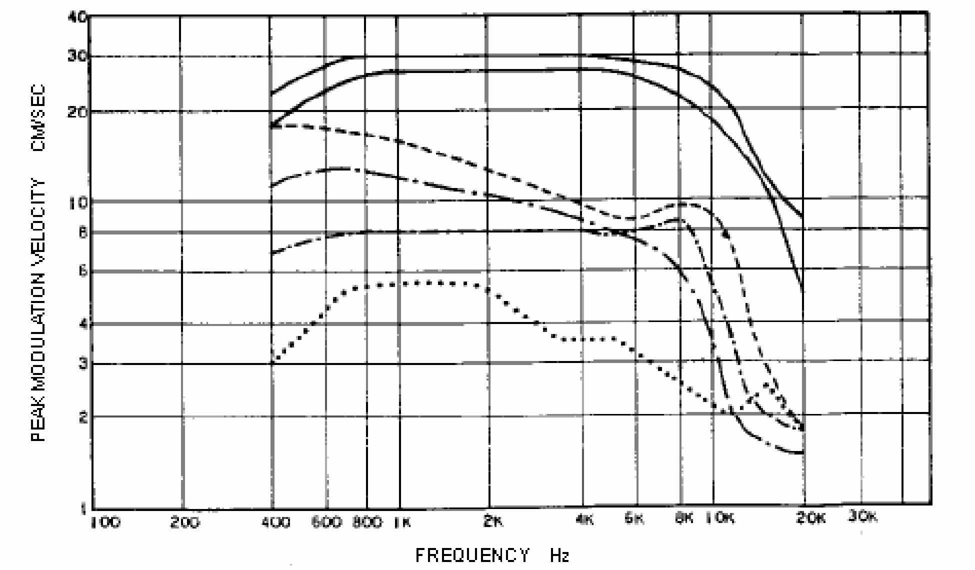


Figure 3.2: Tracking ability vs frequency for several cartridges.
Adapted from (Kogen 1968)

3.1.5 Wow and Flutter

Wow and flutter are pitch variation defects which were not present in the original performance (Godsill and Rayner, 1998). Wow refers to slow shifts in pitch (i.e. speed) and flutter to faster fluctuations. These fluctuations are in usually periodics both cases. They can occur during the recording or playback process, leading to undesirable changes of all of the audio file frequency components (Czyzewski, Maziewski, Dziubinski, Kaczmarek and

Kostek, 2004). The audible effect of wow and flutter is most noticeable on signals which contain pure tones, being the threshold of hearing them around 0,5%, which is the speed variation compared to the ideal value. They can be characterized by means of the pitch variation function $p_w(t)$ or equivalently by the time warping function $f_w(t)$ (Godsill and Rayner, 1998). If $x(t)$ is the undistorted signal and $f_w(t)$ is the pitch variation defect expressed as a time warping function, the heard signal is given by:

$$x_w(t) = x(f_w(t)) \tag{3.1}$$

The relationship between $f_w(t)$ and $p_w(t)$ is expressed as:

$$p_w(t) = \frac{d(f_w(t))}{dt} \tag{3.2}$$

Wow and flutter can be produced by many different sources. A possible mechanism by which wow occurs is the variation of the rotational speed of the recording medium. Another cause is eccentricity during disc and cylinder sound recordings or when copying or reproducing, a wrongly punched gramophone disc is an example of this kind of degradation.

Depending on the source, a mechanical correction could be used for solving the problem, like in the case of a gramophone disc which is not punched correctly. Unluckily, this kind of solutions are normally impractical and some signal processing approach should be used instead (Godsill and Rayner, 1998).

Further details about the causes of pitch variation defects in motor driven record players (see figure 3.3) can be found in the work made by Furst (1946), where the studied sources of the degradation are classified into three groups: those originating in the turntable itself, those originating in the driving motor, and those originating in the driving mechanism necessary to transfer the motion of the motor to the turntable.

3.1.6 Hiss

When listening to an analogue system a characteristic kind of noise, usually perceived as ‘hiss’, is always present, in a more or less noticeable way. This random additive background noise stems from different sources: electrical circuit noise, irregularities in the storage medium and ambient noise from the recording environment (Godsill and Rayner, 1998). Despite its different origins, this degradation is considered as one single noise process, even though the ambient noise from the recording environment could be considered part of the original audio recording. The random nature of this noise makes it be present in all the frequencies, making a filtering or equalization procedure is not appropriate for its restoration.

Moreover, this degradation can have non-stationary characteristics, being found in several early 78rpm and cylinder recordings with highly variable coloured noise. These variations can be noticeable within each revolution of the playback system, resulting in the typical ‘swishing’ effect connected with some early recordings (Godsill and Rayner, 1998).

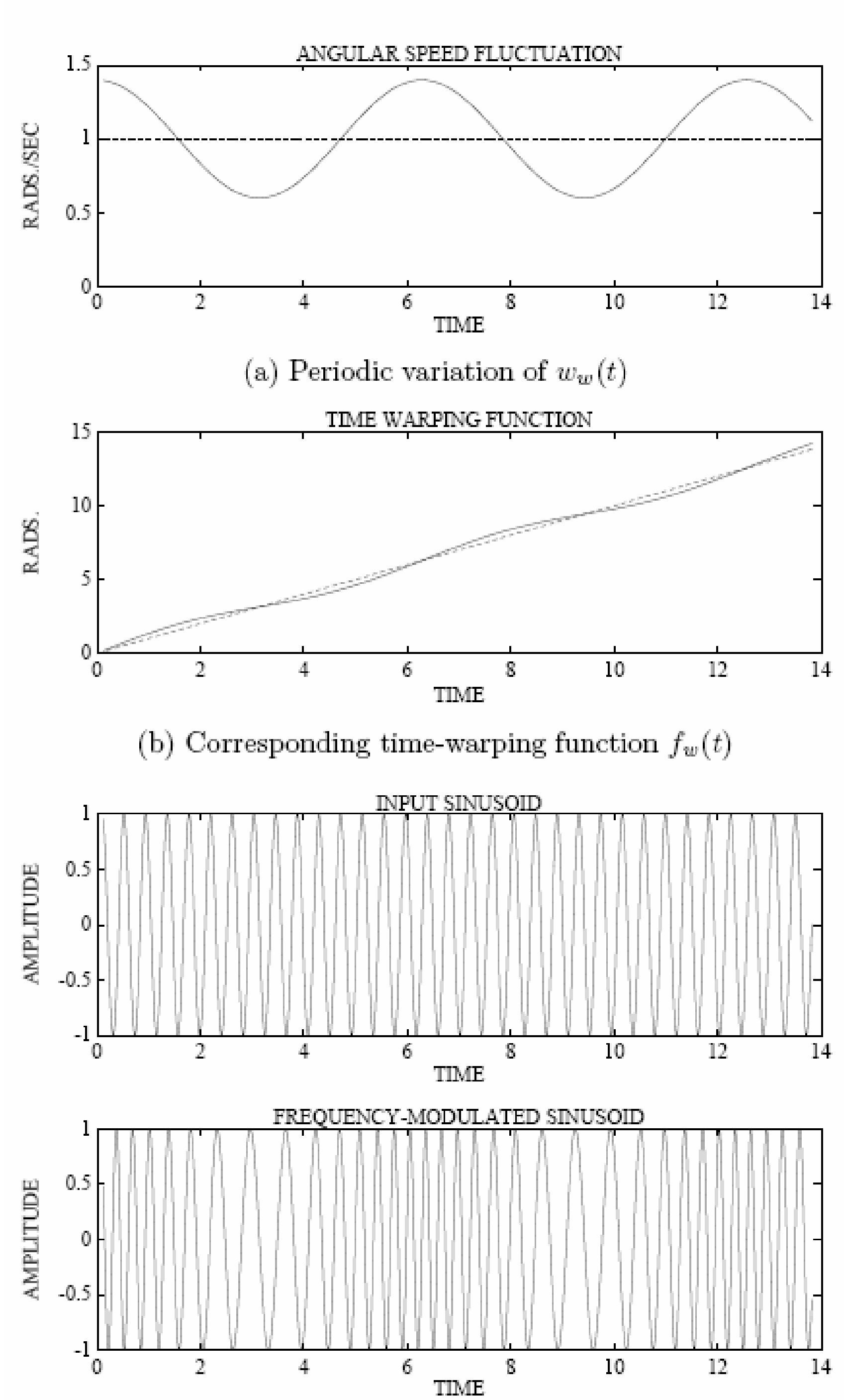


Figure 3.3: Frequency modulator of sinusoid due to motor speed variations.
Adapted from (Godsill and Rayner, 1998)

3.2. Localized degradations

Localized degradations are those which only affect certain samples of the waveform. These can be classified into tracking errors, clicks and low frequency pulses (thumps).

3.2.1 Tracking errors

Tracking errors occur when a big discontinuity in the playing groove make the needle jump to an adjacent groove. If the needle goes back to a previous groove, there is the possibility than when arriving to the same point the same error is produced, being the same part of the song played many times, till the needle is able to continue by itself or some external help is made to stop the repetition.

In case the needle continues playing the next groove, a part of the audio signal will be missed, corresponding to a period of the rotational speed.

3.2.2 Clicks

Clicks are impulsive disturbances random in time and amplitude. Known as one of the most common problems in historical musical recordings, the duration of these bursts of interference are usually less then 1ms. Although this degradation can be quite annoying, it only affects approximately 10% of the samples, being thus possible to hope a successful restoration (Esquef, Biscainho, Diniz and Freeland, 2000).

There are distinct types of ways by which clicks can be noticed by the listener, including tiny ‘tick’ noises common in all recording medium (even in modern digital sources) and ‘scratch’ and ‘crackle’, typically related to most analogue disc recording methods.

The sources by which clicks occur can be really different. The most common reasons in analogue disc recordings are the specks of dirt and dust adhered to the grooves (figure 3.4) and granularity in the material used for pressing the disc. Small scratches in the surface are common too, being another type of click source. In digital recordings, however, clicks can be the consequence of poorly concealed digital errors and timing problems (Godsill and Rayner, 1998).

Clicks, as almost all localized degradation (Godsill and Rayner, 1998), can be well modeled by an additive model, where the degradation is added to the audio signal. This additive model can be expressed as:

$$y_t = x_t + i_t n_t \tag{3.3}$$

where x_t is the audio signal, n_t is the corrupting noise and i_t is the indication of when the defect is present. An example of a click-degraded signal is shown in figure 3.5.

Figure 3.4: Dirty surface of a vinyl disk.

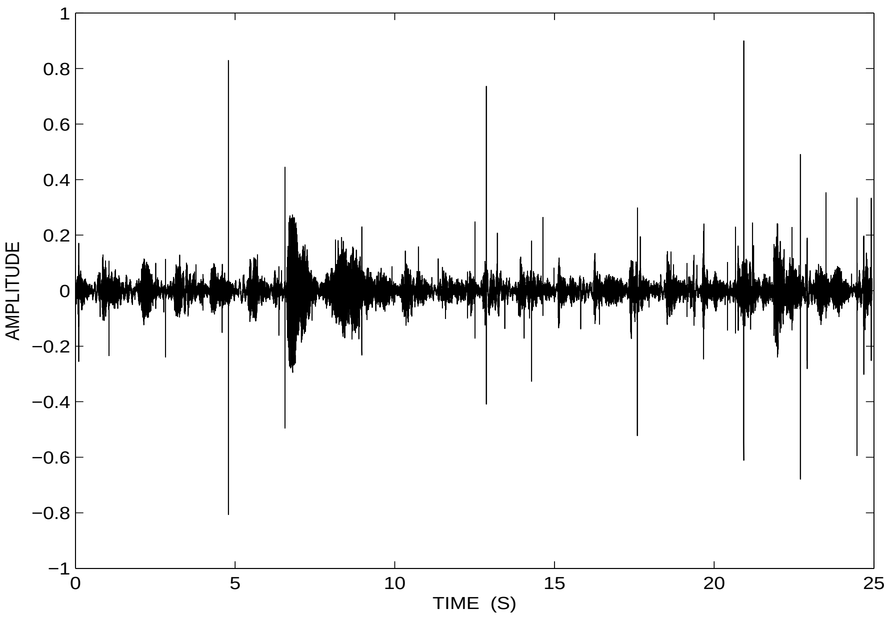


Figure 3.5: Click-degraded audio waveform taken from vinyl disk.

3.2.3 Low Frequency Pulses (Thumps)

When a severe damage is done on the groove walls of a disc or cylinder, the effect is the degradation known as 'low frequency pulses' or 'thumps', one of the most annoying disturbances in historical musical recordings. Some mechanisms by which thumps can occur are deep scratches on the disc surface and the joints resulting when parts of a broken disc are fixed with glue. During the reproduction of the musical recordings, the stylus-arm is exited by these discontinuities and its impulsive response is additively superimposed upon to the undistorted audio signal in such a way that a low frequency pulse is produced (see figure 3.6) (Godsill and Rayner, 1998). In this way, this type of distortion can be generally described by a short and strong discontinuity similar to a click with a duration usually shorter than 2 ms followed by a long and decaying transient of low frequency content which normally lasts more than 50 ms (Esquef, Biscainho and Välimäki, 2003).

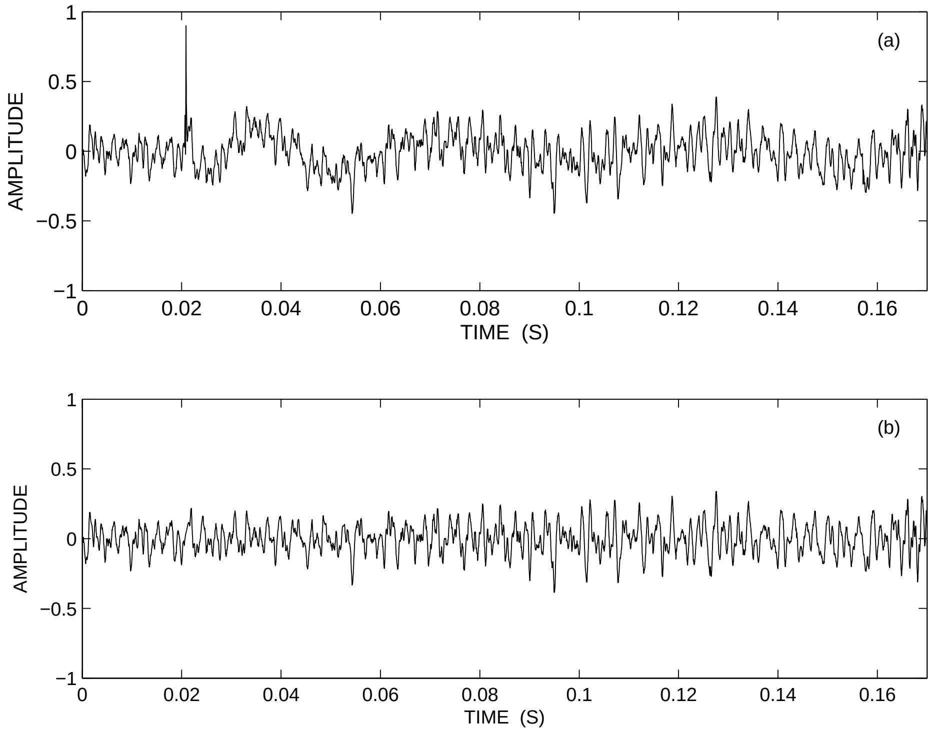


Figure 3.6: (a) Audio waveform corrupted by a low frequency pulse. (b) Restored audio waveform of the corrupted one presented in (a).

Chapter 4

Digital Simulation of Degradations

The distinct types of degradations can be implemented using different techniques, and the most suitable for each one taking into account the available music and computational resources was tried to be chosen.

4.1 From Stereo to Mono

The first step in audio antiquing is to have the desired music file in monophonic to can reproduce the quality of the old recordings, when the stereophonic sound was still not developed. Even though some of the disk recordings were stereo, the earliest ones (even for the LPs) were monophonic. Due to the material which was provided to us, it was decided to do the LP file monophonic too, despite the fact that some simple changes in the Matlab code would allow the required degradations for a stereo file.

There are different issues to take into account when a music file is reduced from stereo to mono. The two available channels in the stereo music file allow the performance of some audio effects. These effects can include the inversion of the phase in one of the channels or the introduction of a delay between them. In these cases, the conversion from stereo to mono should be done carefully, knowing where the effects are and deducing in each case the best method not to lose information (not more than strictly necessary). For example, a way for the detection of a delayed channel could be done with the evaluation of the peak in the correlation of the two signals. Anyway, although at the beginning a solution for these cases was studied, they were finally rejected because of their complexity.

Nowadays, most of the music is produced “mono-compatibly” to avoid the problems some devices with only one speaker (e.g. a simple radio) would have with its reproduction. The most broadly extended method is working out the average of the two channels, which is a simple way to do the conversion and has good results in a high percentage of the cases. Because of that, this is the implemented solution.

4.2 Frequency Response Deviation

The frequency response of the horn used in gramophone and phonograph recordings reproduction has an important role in the final sound. To calculate it different programs are available, like ‘Hornresp’ or ‘AJ-Horn’. Between them, the program ‘Hornresp’ was used for the simulation. The main reasons for this election were that it was free and easy to use. Having the measurements of the desired horn and the place in which it would be positioned (radiating into free space, half space, quarter space or eight space) the acoustical impedance of the horn can be calculated by this program (McBean, 2007).

Once the impedance is estimated, the next step is to work out the transmission coefficient to know how much sound will be transmitted through the horn at different frequencies. From (Fletcher and Rossing, 1991) the formula to calculate the transmission coefficient is:

$$t = 1 + r = \frac{2Z(w)}{Z(w) + Z_0} \qquad (4.1)$$

where r is the reflection coefficient, $Z(w)$ is the acoustical impedance of the horn and Z_0 is the characteristic impedance of the pipe to which the horn is attached, expressed as:

$$Z_0 = \frac{\rho c}{S} \qquad (4.2)$$

where ρ is the density of the medium inside the pipe (in this case air), c the velocity of sound in that medium and S the cross-sectional area of the pipe. The formula for a lossless cylindrical pipe (‘if the walls of the pipe are rigid, perfectly smooth and thermally insulating then the presence of the tube wall has no effect on wave propagation’, (Fletcher and Rossing, 1991)) has been chosen to work with, simplifying the calculation and providing reliable results.

When the transmission coefficient of the horn is known, i.e. when the frequency response is known, it is time to implement the filter. The data given by the ‘Hornresp’ program is in base a logarithmic frequency scale. A technique for calculating a FIR filter coefficients knowing the frequency response is to use the *IFFT* (Inverse Fast Fourier Transformation) of a uniformly sampled frequency response. In this case, the frequency response obtained is not equally spaced, so an interpolation is made to have the values in frequency at the desired points.

In figure 4.1 the steps for obtaining the frequency response are shown. Starting from a schematic diagram of the horn, its impedance is calculated by ‘Hornresp’ and finally the transmission coefficient is worked out with the explained formulas.

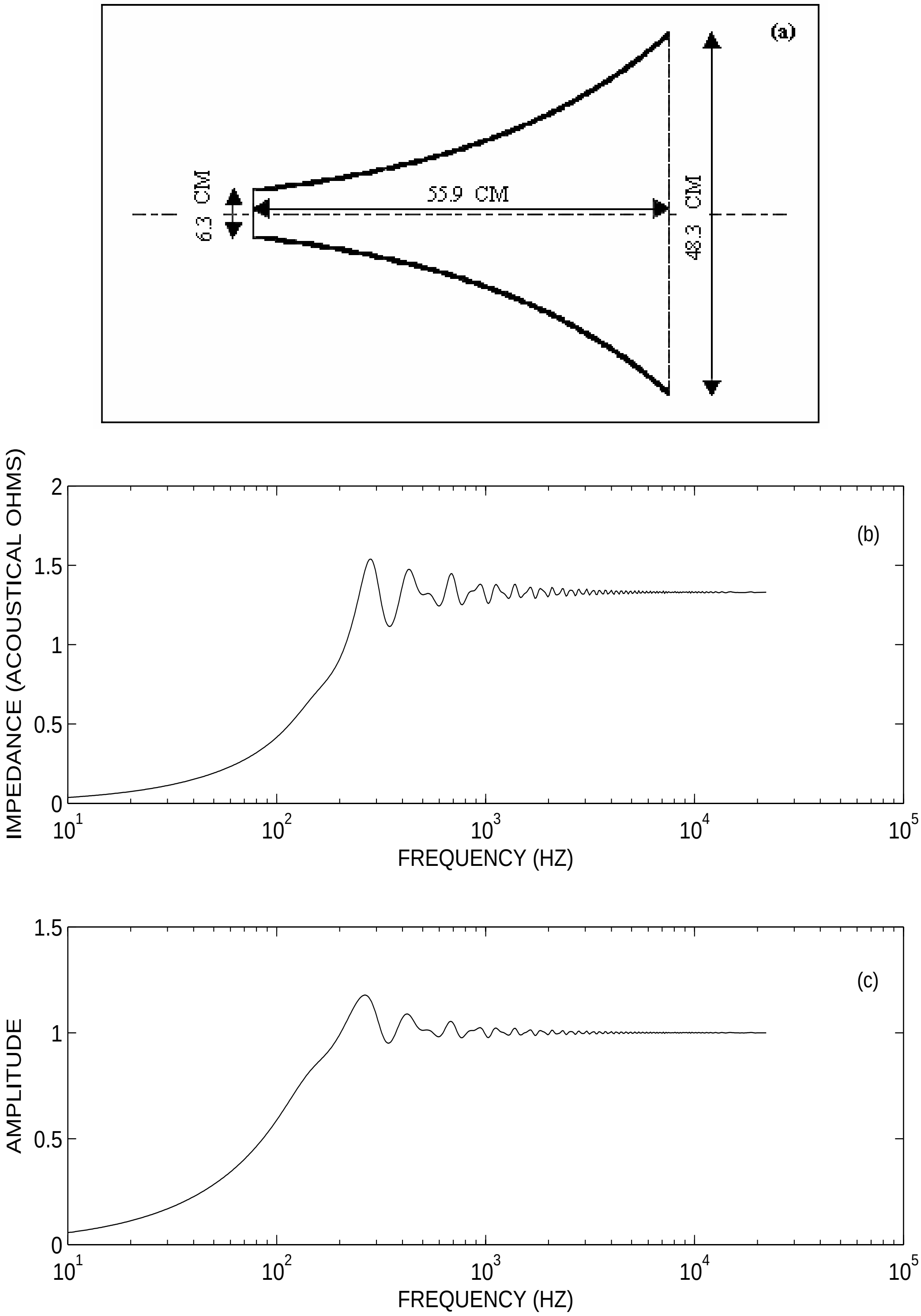


Figure 4.1: (a) Schematic diagram of a gramophone horn. (b) Acoustic impedance of the horn represented in (a). (c) Transmission coefficient of the horn represented in (a).

4.3 Signal Bandwidth

The signal bandwidth of an audio signal recording represents the band of frequencies which a recording technology can store and reproduce. The available bandwidth in the old audio recordings was more limited than the one provided by the CD, so to achieve the desired bandwidth some filters should be used to take away the appropriate frequency components.

To implement this degradation two different Butterworth filters are used. Although the order of this kind of filter is higher than the one obtained for Chebychev or Cauer filters, the smoothness it provides is worthy for the final purpose. Moreover, this application is done offline, so there is no need for saving time in the signal processing (anyway, the time difference would not be so significant).

The first filter is used at the beginning of the process, just after the stereo to mono conversion. It is a low-pass filter or a band-pass filter, depending on the historical recording to simulate. With this filter the audio content in the frequencies out of the desired bandwidth is removed. The second filter, a low-pass filter in all the cases, is used in the final steps, to reduce the high frequency components introduced by the different simulated degradations which could not be possible in the old recordings.

4.4 Dynamic Range

As it was stated previously, the dynamic range indicates the ratio of the maximum signal value to the softest or to the noise level. Nowadays, a compressor is commonly used for reducing the dynamic range of an audio signal. Averaging the power during a short time interval the gain of the signal can be controlled. A compressor attenuates the signal in the loud passages when a threshold in the averaged power is exceeded, reducing thus the dynamic range. A special type of compressor is the limiter, which only saturates the signal when it exceeds a certain level (Välimäki, 2007).

However, it is difficult to imagine when a phonograph cylinder or a gramophone disk was being recorded somebody telling the musicians that they should play softer or louder because of the dynamic range. That is why when implementing this part of the antiquing process the use of this kind of compressors was avoided.

Another way of reducing the dynamic range is to reduce the number of levels in the quantization. It is quite simple to implement and it does not involve any external control when recording. It contributes to the increment of noise and distortion, but taking into account the final purpose of this thesis that it is not really a negative contribution.

There are different options for reducing the number of levels in the quantization. One is to decrease the number of bits per sample, but in this method only some dynamic range values can be reproduced. A way to avoid this problem is explained as follows. The dynamic range can be expressed as:

$$D_R(dB) = 20\log_{10}(2^w) \tag{4.3}$$

where w is the number of bits used to represent a sample and 2^w is the number of levels for the quantization. Once the desired dynamic range has been chosen the corresponding number of levels (N_Q) can be deduced as:

$$N_Q = 10^{D.R./20} \tag{4.4}$$

If M is the maximum value of the original signal, now, having N_Q number of levels, the maximum value for the next step will be M', so that:

$$M' = \frac{N_Q}{2^w} M \tag{4.5}$$

The new signal can be calculated as a quantization of the original signal where the minimum value is 0, the maximum value is M' and N_Q is the number of levels. The relationship between the two signals is represented in figure 4.2 in the case M=15, M'=10 and N_Q=21.

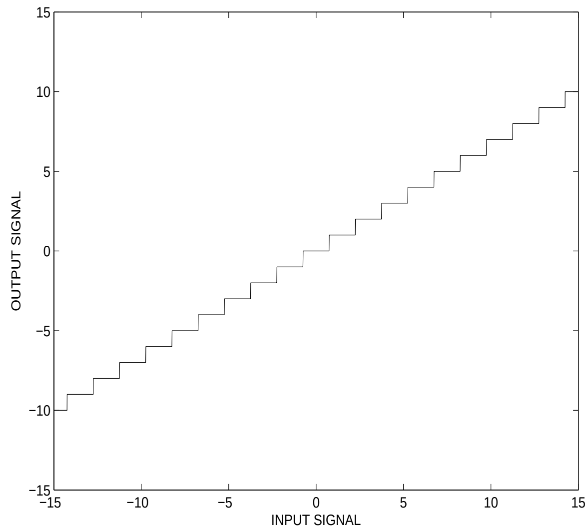


Figure 4.2: Relationship between signal input and output level in Dynamic Range compression.

At this point the audio signal can be normalized to M and the number of levels will remain the same. It should be remembered that the reduction of the loudness is not the point in this step, because at the end all the files should have the same volume to make a good comparison.

4.5 Distortion

Creating this audio waveform deformation is a difficult process since the functions by which the different types of distortion are created in the medium are generally unknown. But what it is clear about this process is the final result: the harmonics, the clipping, etc.

The chosen way to simulate this distortion was to use two different functions: one to create the non-linearity in the loud passages and one for the soft samples.

The hyperbolic tangent is used for the first purpose because of its linearity for the low values of the signal and the shape of its saturation. Introducing a parameter (called equation 4.6 *base*) and normalizing the function, the amount of distortion to introduce can be controlled:

$$y = \frac{\tanh(base \cdot x)}{\tanh(base)} \tag{4.6}$$

For the soft passages equation 4.7 is used. The higher the *base*, the more distortion is introduced (see figure 4.3).

$$y = sign(x) \cdot abs(x)^{base} \tag{4.7}$$

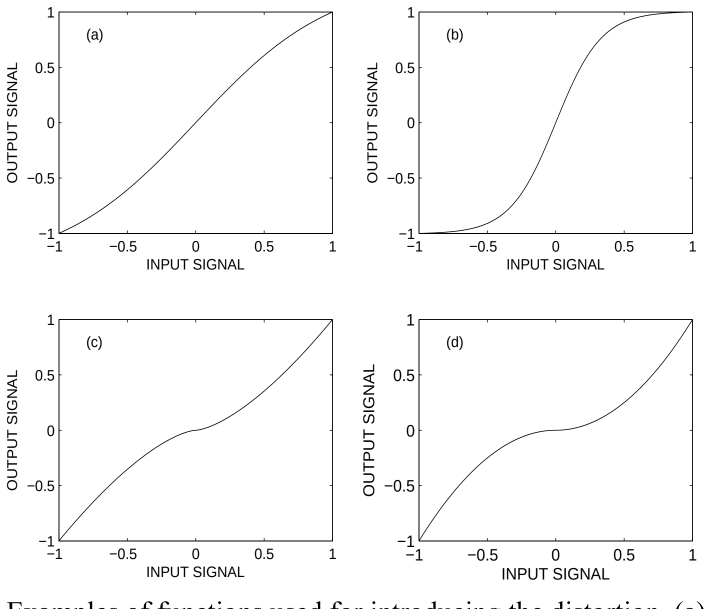


Figure 4.3: Examples of functions used for introducing the distortion. (a) Distortion function with equation 4.6 when *base* = 1. (b) Distortion function with equation 4.6 when *base* = 3. (c) Distortion function with equation 4.7 when *base* = 1.5. (d) Distortion function with equation 4.7 when *base* = 2.

This technique for creating the distortion has acceptable results and the hard studies needed for having reliable data about the distortion sources and functions are avoided. The combination of the two non-linearities and its effect over a sinusoidal signal are represented in figure 4.4.

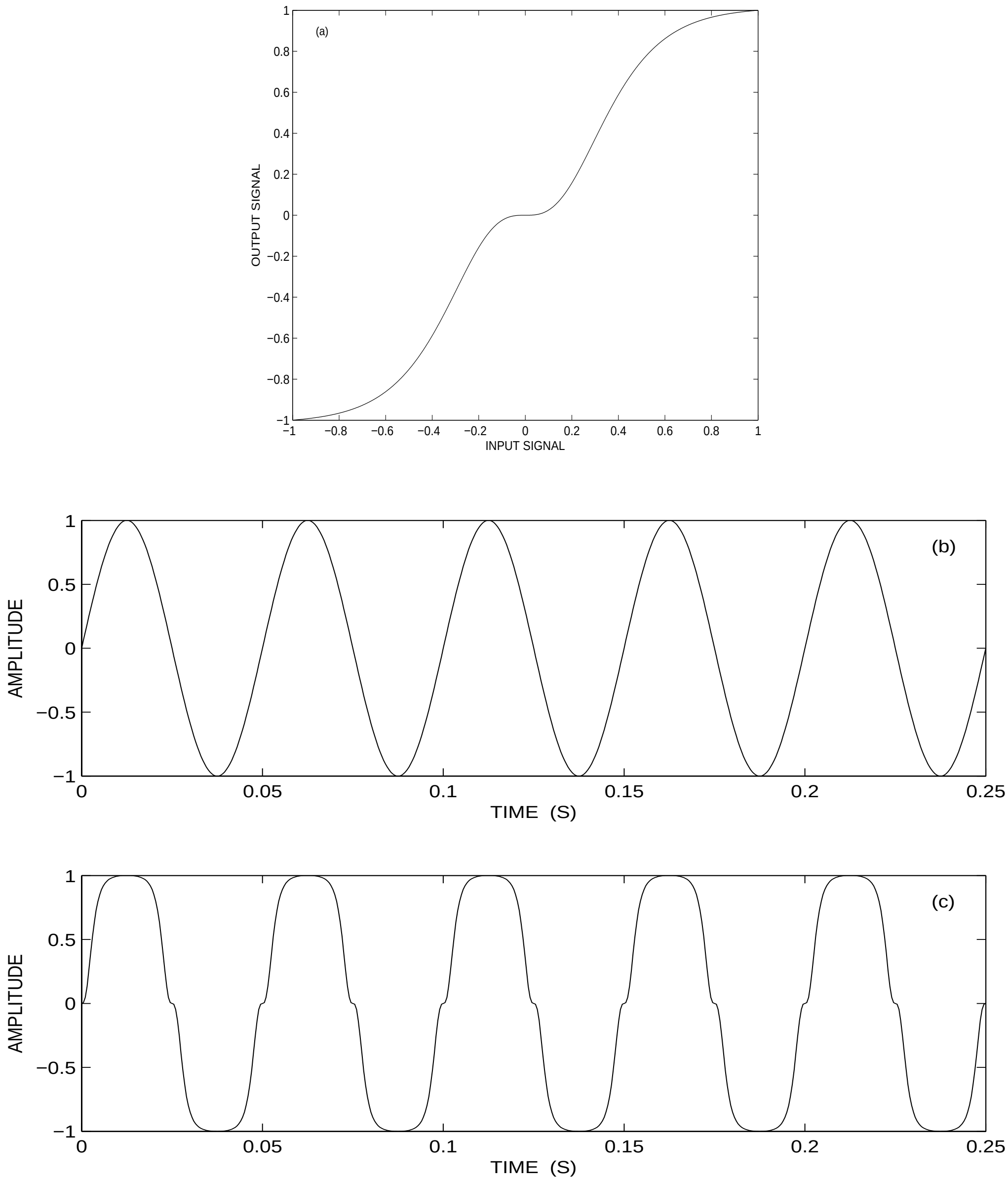


Figure 4.4: (a) Total distortion function (*base* = 2.7 for equation 4.6 and *base* = 2 for equation 4.7). (b) Undistorted sinusoidal signal. (c) Distorted sinusoidal signal.

4.6 Wow and Flutter

The process to introduce the wow or flutter into a music file can be divided into two fundamental parts: creation the time warping function and resampling of the audio file.

The first thing to take into account when creating the pitch variation curve is the nature of this type of degradation in disk and cylinder recordings. In these cases the distortion is usually periodic and has smooth variation with time. Without taking into account the noise, the frequency components of the audio signal can be expressed as (Godsill and Rayner, 1998):

$$F = F_0 P \tag{4.8}$$

where F_0 is the centre frequency, P the pitch variation and F the final frequency.

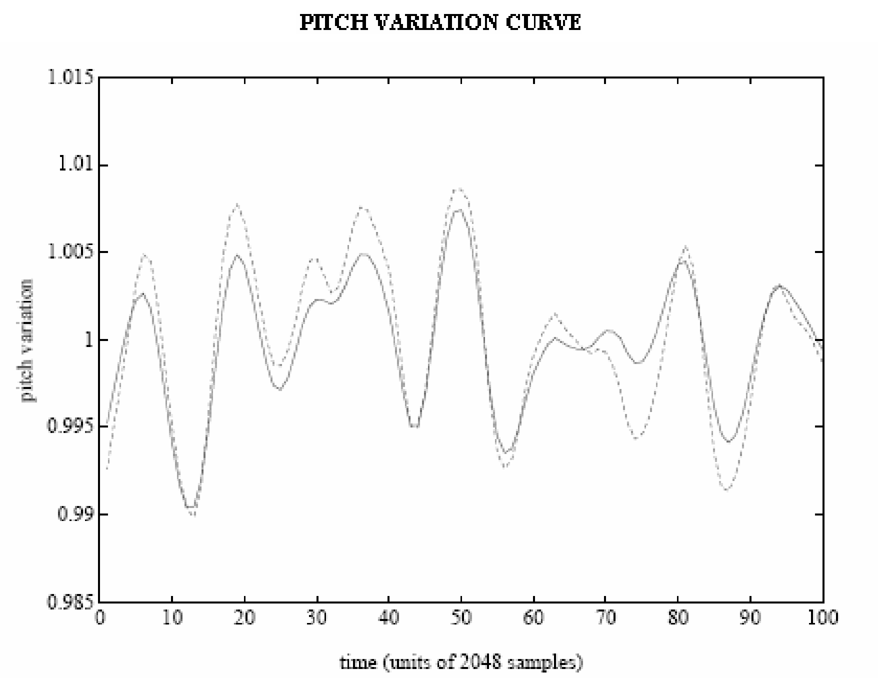


Figure 4.5: Pitch variation curve estimated for an historic recording using two different models.
Adapted from (Godsill and Rayner, 1998)

Knowing the smoothness and periodic nature of this degradation, a sinusoidal function with variable frequency and envelope can simulate a pitch variation curve:

$$P = 1 + A(t) * \sin(2\pi f(t)t) \tag{4.9}$$

This simulation can be split up into the definition of two different parts: frequency variation and envelope variation. The two curves are formed in the same way: having their average value and their deviations (assuming normal probability density functions) a random function is created for each one. Since the changes should be smooth, some time is needed for the transition from one frequency and envelope value to another. That transition was decided to last one fifth of the mean period (a revolution period of the recording in case of wow). To obtain all the needed points (the sampling rate is 44100 Hz), a 'spline' interpolation is used to join all the data in a smooth way. Now that the corresponding frequency and envelope values are known for every sample, the sinusoidal signal (equation 4.9) can be implemented (see figure 4.6).

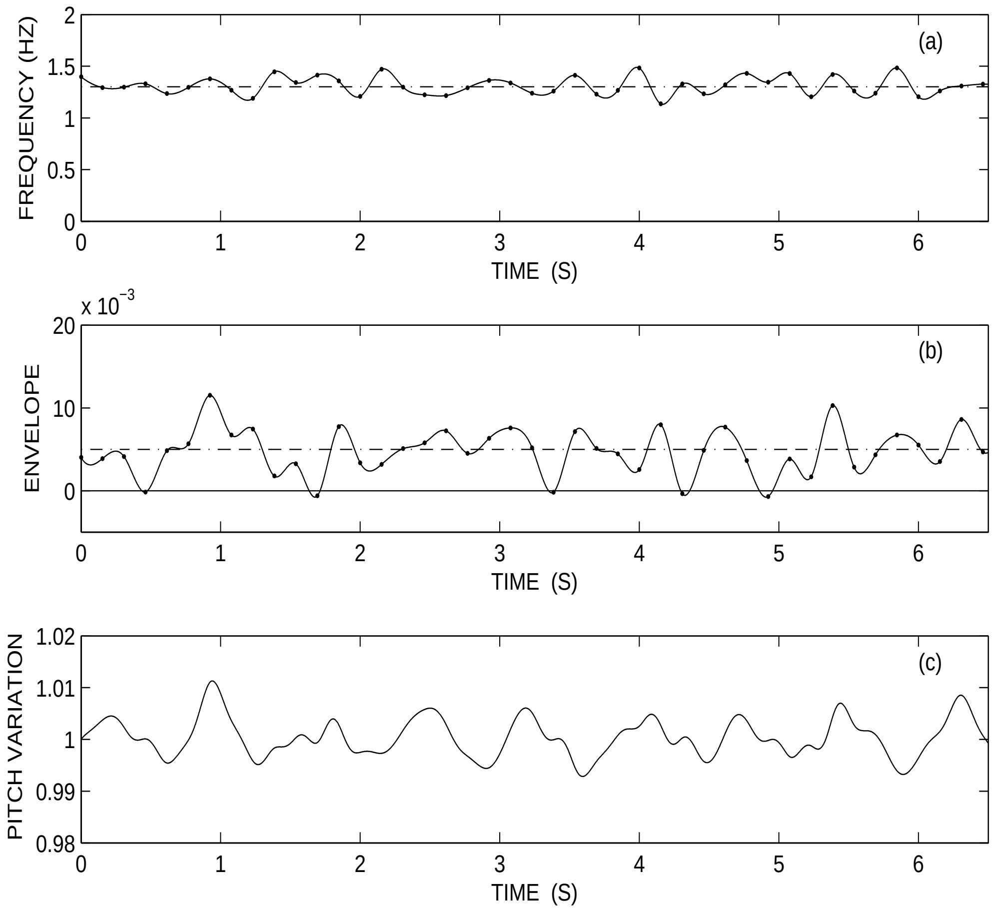


Figure 4.6: (a) Frequency variation curve. (b) Envelope of the pitch variation. (c) Synthetic pitch variation curve.

The resampling for the creation of the defect is done using a 'spline' interpolation. As it can be seen in (Maziewxki, 2005), this technique is between the best ones. At the same time it is easy to implement thanks to the provided Matlab functions. If more accurate results are needed a resampling with a sufficiently long truncated *sinc* function can be used.

4.7 Hiss

In the historical recordings there is usually a part of time where no music is played, so the heard sound is due to the noise the recording system, the storage medium or the playback system has introduced. This noise is mostly formed by hiss, clicks and thumps. While hiss is a kind of global degradation, clicks and thumps are localized degradations, being thus not present during all the time without musical signal. For this reason, if a silent part is present in the audio file, usually at the beginning or at the end, it will be a period while hiss is the most predominant degradation.

An important aspect to explain is that only stationary hiss is considered in the simulation. This assumption is made due to the historical audio sources provided for the realization of the project, where few important changes can be observed looking at the spectral properties of the hiss at different moments. Moreover, it simplifies the recovering data procedure and its performance being the same time quite realistic.

The first step for hiss reproduction is to have the characteristics which represent it in the different historical recordings to imitate. For this, some silent extracts have been obtained from the given audio files and their frequency characteristics have been studied, noticing at which frequencies the hiss has peaks or dips.

Starting from white noise, a filter with a similar frequency response as the corresponding spectral properties of the hiss is used (see figure 5.8 and 5.9). After this process, a noise similar to hiss is achieved. The filters are designed using LPC.

4.8 Tracking Errors

This degradation is produced when a considerably large discontinuity is found by the needle when following the groove. The result is usually a jump to an adjacent groove, repeating or suppressing a part of the audio file. If the original audio recording is available the simulation of this effect is quite simple to do. Knowing the revolution velocity of the playback system, the effect starts with a strong thump and then it goes back or forward a period of the revolution system (for example, if the revolution velocity of an LP is 33 r.p.m., its period is 60/33 s). If it goes back, the effect can be repeated until the needle finally follows the correct groove.

4.9 Clicks

The parameters to define this impulsive disturbance are: duration of the burst of interference, time between them and their amplitude. If the probability distributions of these parameters are known, clicks can be reproduced in a reliable way. Thus, the simulation can be divided into three different parts: harvest of data, modelling of clicks and reproduction.

For the harvest of data, the corrupted and the restored (only restoration of clicks) music files are needed. When these two files are available, the clicks can be extracted by a simple subtraction (remember that the clicks could be modeled by an additive model). The most difficult part is to obtain the restored file and a lot of techniques have been studied to remove the clicks. The restoration of the audio file can be split between the detection of clicks and the interpolation for replacing the distorted samples. At first, it was thought to implement the detection part to can extract some data about the clicks. Several methods to carry out it were found in (Godsill and Rayner, 1998), (Esquef, Biscainho, Diniz and Freeland, 2000), (Kauppinen, 2002) and (Luszick, 2003). A trade-off between the methods explained in (Kauppinen, 2002) and (Luszick, 2003) was done, using the LMS algorithm to predict the next sample of the file and a median threshold curve to the second derivative detection signal based on the residual of the adaptive filtering. Although the results with the artificial clicks were acceptable, there was the problem of setting the parameters in a real corrupted file, because it was not possible to know if the detected clicks were correct without implementing the interpolation. Moreover, just with the detection, information about the clicks amplitude was not collected. Implementing the interpolation would mean a lot of time just for having some data about one of the multiple degradations, so that option was rejected. Looking for an easier solution to solve this problem it was found an option for click/pop removing in Adobe Audition program. By this way, the file can be restored and the final result can be listened, being able to know if the restoration was well-done. It should be taken into account that undistorted samples can suffer from some small changes with this process, so a threshold was used after the subtraction of the corrupted and restored audio file to define above which value it was considered to be a click. It is important too to notice that not only the clicks have been removed, also the pops, but knowing that the duration of the pops is much larger than the duration of the clicks (around 50 ms in comparison with 1 ms), it is easy to filter them. Being able to listen and see all the time the results helps to decide the correct parameters for the harvest of data. The steps followed for the clicks extraction can be seen in figure 4.7.

As it was stated, in the modeling of the clicks three different parameters are important: the duration of the clicks, the duration of the gap between clicks and their amplitude. The goal of this part is to characterize statistically these values to be able to reproduce them in a reliable way. With the available data, histograms can be produced to know their distributions. Then, a known distribution similar to the shape of the histogram is chosen, and the parameters which define it are deduced using some given Matlab function. The probability density functions with which the data is compared are: exponential distribution, gamma distribution, weibull distribution, normal distribution, lognormal distribution, extreme value distribution and poisson distribution. The most adequate is chosen by the highest Probability Plot Correlation Coefficient (PPCC).

Finally, once all the distributions and their parameters are known the clicks can be simulated in the desired audio file.

When implementing this part a problem was found about the credibility of the clicks. They seemed to be very static, while in a real historical recording they are more dynamic. This was due to the frequency component of the clicks. Whereas in an old recording the spectral properties of the clicks are varying from one to another, in the artificial ones the spectrum was very similar in all of them. To solve this problem some Butterworth low-pass filters with variable cut-off frequency are used to make them sound more dynamic (see figure 5.6, 5.17 or 5.29). Before adding the clicks to the audio waveform, the signal with the clicks is framed, and then each frame is filtered with a filter in which the cut-off frequency is decided by a random function.

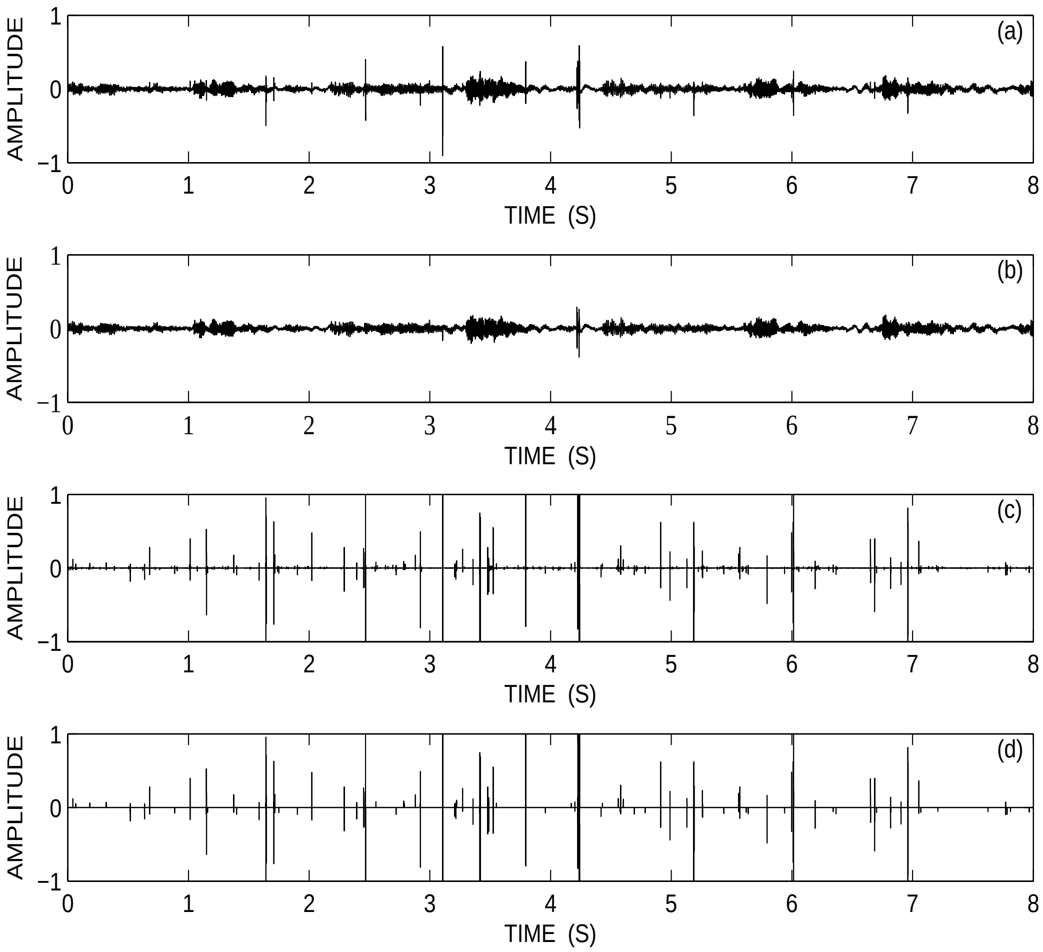


Figure 4.7: (a) Click-degraded signal. (b) Click-restored signal. (c) *(b)* subtracted from *(a)*. (d) Extracted clicks.

4.10 Low Frequency Pulses (Thumps)

For simulating this degradation the structure of some available artificial and real extracted thumps (Esquef, Biscainho and Välimäki, 2003) was studied. It is important to stress the introduced errors in the thump due to the technique used for its extraction. The ideal case of a low frequency pulse is a short discontinuity followed by a long and decaying transient of low-frequency content. When this theoretical thump is introduced in an audio file and then extracted, the errors in the detection algorithm make them seem less smooth than they really are. Figure 4.8 shows the synthetic thump introduced and the one detected.

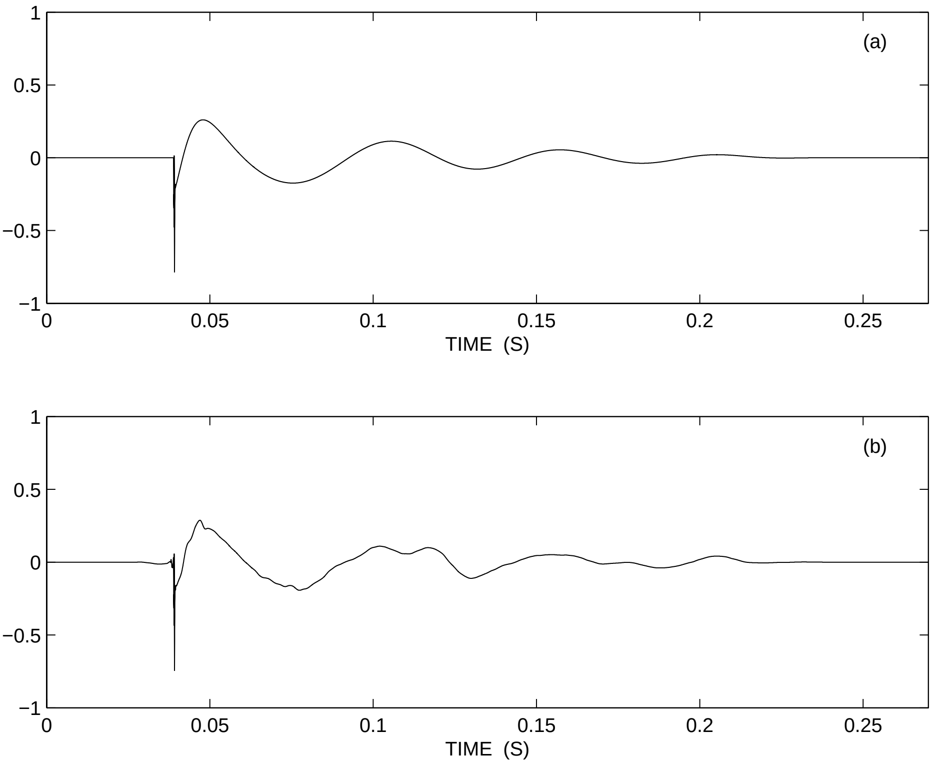


Figure 4.8: (a) Introduced artificial thump. (b) Extracted thump.

Knowing the errors introduced by the algorithm is good to interpret the results obtained when low frequency pulses are extracted from historical recordings. In the same publication, (Esquef, Biscainho and Välimäki, 2003), some extracted pulses from an historical recording are available, so a general study can be done from them. After the analysis, the conclusion which was reached was that the low frequency pulse, as in theory, can be divided into two parts: the initial discontinuity and a long tail. The discontinuity is modeled as a strong click, which is followed by a time where the excited stylus starts an oscillation which decays exponentially until zero.

Having as parameters the lengths, the amplitude and its deviation of the two parts and the frequency and its deviations tail (assuming in all the cases normal pdf), a thump can be simulated (see figure 4.9). In case of the tail, a slightly modified method of the one used for the pitch variation curve is implemented. It should be remembered that the thump is usually caused by deep scratches, so it is repeated during all the revolutions in which the scratch is crossing the groove, with a period equal to the period of the rotational speed of the recording medium. The length of the scratch is another parameter needed for the correct simulation of this degradation.

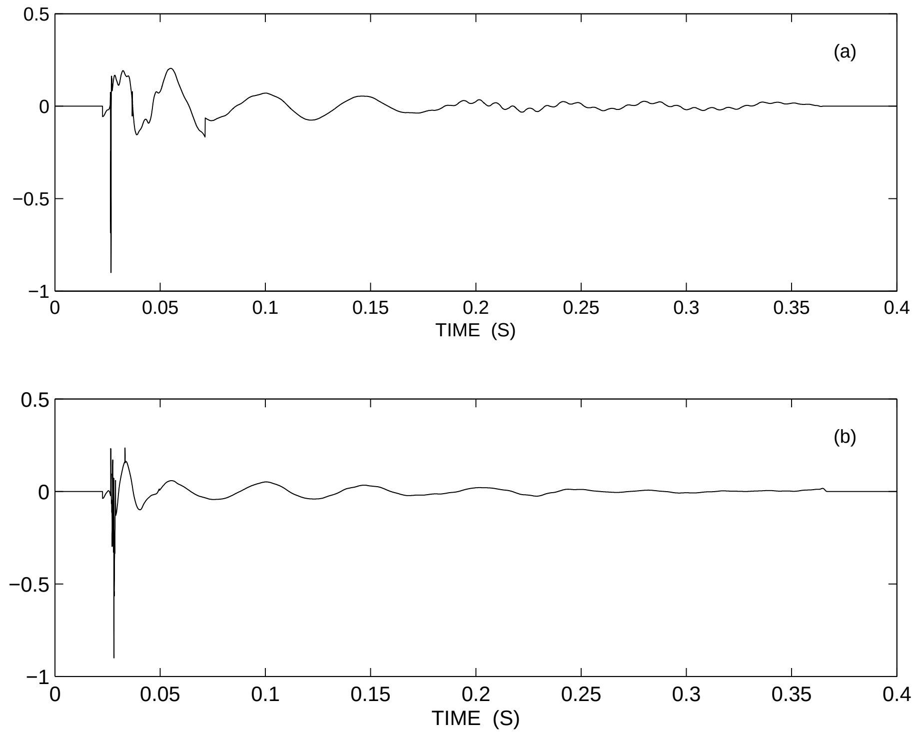


Figure 4.9: (a) and (b) Real extracted thumps.

Chapter 5

Antiquing of Music Files

In this part the steps followed to make a song with CD-quality sound like if it was reproduced by an LP, a gramophone or a phonograph are explained. How to implement the degradations was already explained in Chapter 4, so only specific data for each simulation is showed.

It was decided not to use the dynamic range reduction based on quantization for not being the most appropriate method. Indeed, the addition of different degradations such as hiss or distortion reduces the final dynamic range of the audio file, fulfilling the historical recordings characteristics.

5.1 LP (Vinyl Disk)

The sound quality of an LP depends on many distinct factors, from when it was manufactured to its maintenance. In this section, an early monophonic LP quality is simulated. In figure 5.1 the followed steps for the simulation are shown.

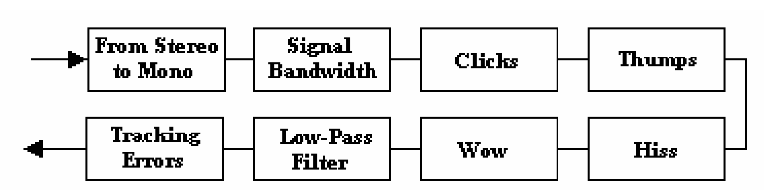


Figure 5.1: Steps followed for LP quality simulation.

5.1.1 From Stereo to Mono

A simple average of the left and right channel is used to convert the stereo file to a mono file.

5.1.2 Signal Bandwidth

In figure 5.2 it can be observed the spectrum of an historical vinyl recording. From there the frequencies to be removed are deduced. In this case, it is only necessary to suppress the high frequencies, so a low-pass Butterworth filter is used. Its parameters are:

Wp (Hz)	Rp (dB)	Ws (Hz)	Rs (dB)
9000	0,45	12000	13

where Wp is the pass-band corner frequency, Ws is the stop-band corner frequency, Rp is the pass-band ripple (the maximum permissible passband loss) and Rs is the stop-band attenuation.

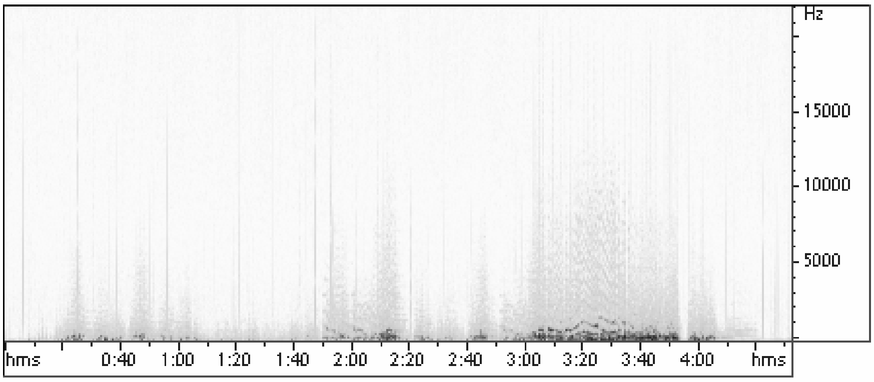


Figure 5.2: About 1955 orchestral music vinyl recording spectrogram.

5.1.3 Clicks

There are three different parameters to model: the duration of the clicks, the duration of the gap between the clicks and their amplitude. In the following figures (5.3, 5.4 and 5.5) their histograms and the parametric distributions which best fits with the data are shown. The selection of the suitable statistic distribution is made by term of the Probability Plot Correlation Coefficient (PPCC) as it was explained before.

For the duration of the clicks (figure 5.3), a Weibull distribution is chosen for the simulation. Its pdf (probability density function) is defined by:

$$f(x/a,b) = ba^{-b}x^{b-1}e^{-(\frac{x}{a})^b} I_{(0,\infty)}(x) \tag{5.1}$$

In the case of the time between the clicks the statistical distribution which best represents it is the Gamma distribution, whose pdf is:

$$f(x/a,b)=\frac{1}{b^{a}\Gamma(a)}x^{a-1}e^{\frac{x}{b}} \tag{5.2}$$

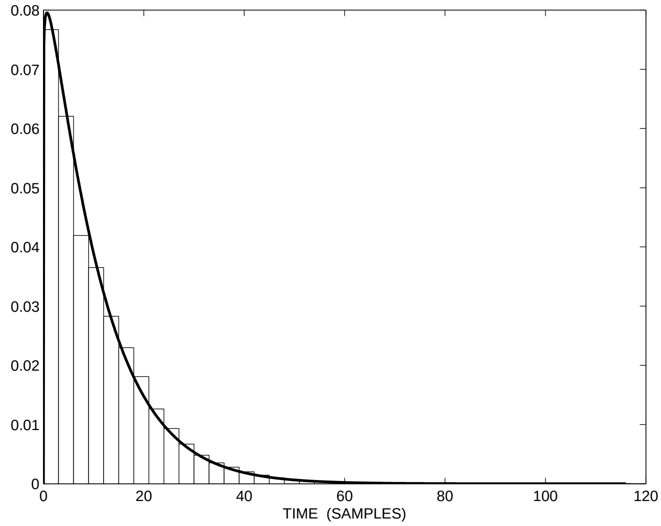


Figure 5.3: Duration of the clicks histogram vs. Weibull distribution (PPCC = 0.99853).
Parameters: a = 10.6907 and b = 1.0606 (see equation 5.1)

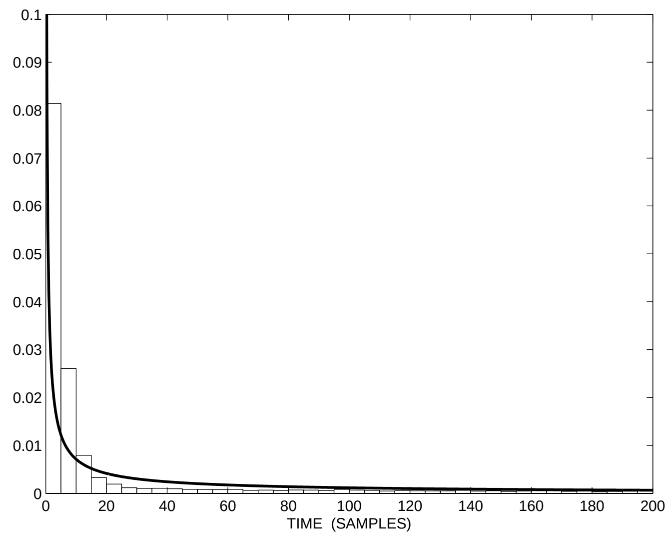


Figure 5.4: Time between the clicks histogram vs. Gamma distribution (PPCC = 0.97029).
Parameters: a = 0.2 and b = 2433.8 (see equation 5.2)

As can be observed, while the distributions of the duration of the clicks and of the gaps between them are discrete probability distributions, the ones with which they are modeled are continuous. For solving this problem, the obtained sample position is rounded to the nearest integer. In case the nearest integer is 0, 1 is chosen instead.

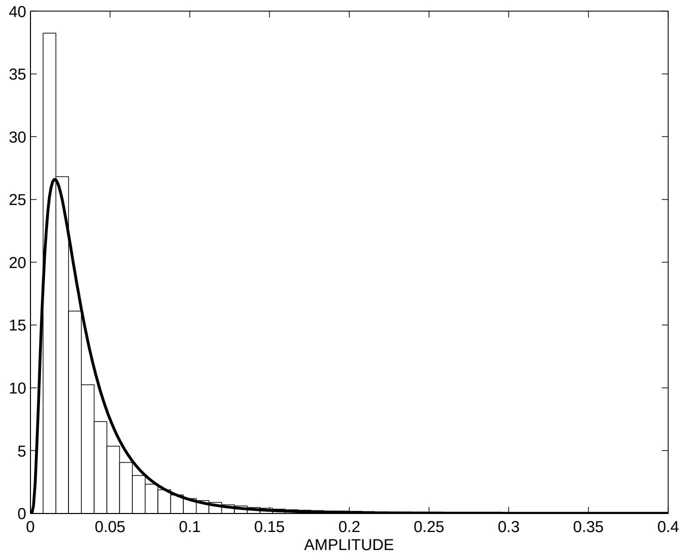


Figure 5.5: Clicks amplitude histogram vs. Lognormal distribution (PPCC = 0.97397). Parameters: μ = -3.6267 and σ = 0.7421 (see equation 5.3)

In case of the amplitude of the clicks, they are simulated with a lognormal distribution. Its pdf is expressed as:

$$f(x/\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{-(\ln x-\mu)^2}{2\sigma^2}} \tag{5.3}$$

The modeling of the amplitude of the clicks was done with the absolute value, so when reproducing it a random function with equally probability for -1 and 1 is used to decide the sign. Moreover, there is a problem with the loudness of the clicks, because the obtained results are dependent on the volume of the file from which there were extracted. Knowing that they are modeled by a Lognormal distribution, the expected value is expressed by:

$$E(X) = e^{\mu+\sigma^2/2} \tag{5.4}$$

If the desired mean is *x*, the obtained samples can be multiplied by *x/E(X)* and the volume of the clicks will increase the wanted quantity. For vinyl recordings the average value is set in this case to 0.2.

The filters used to provide the needed dynamic have a minimum normalized cut-off frequency of 0.1 and a maximum of 0.5. The cut-off frequency is uniformly distributed, so all the values between the minimum and the maximum have the same probabilities. The frame duration is 1000 samples and the order of the Butterworth filter is 3. In figure 5.6 different possible frequency responses of the filter are shown.

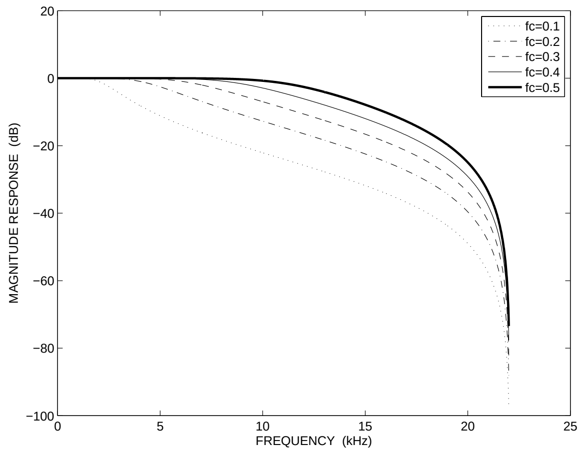


Figure 5.6: Magnitude Response of low-pass filters with different cut-off frequency. In the legend the normalized cut-off frequency for each filter is shown. The sampling rate is 44100Hz.

5.1.4 Low Frequency Pulses (Thumps)

There were supposed to be 8 scratches with different lengths (crossing from 5 to 9 grooves). Distinct thumps were simulated for each scratch. An example of a possible thump is shown in figure 5.7.
The revolution velocity of a vinyl recording is 33 r.p.m. (revolutions per minute), so a revolution period is Td = 60/33 s, time elapsed between thumps created by the same scratch.

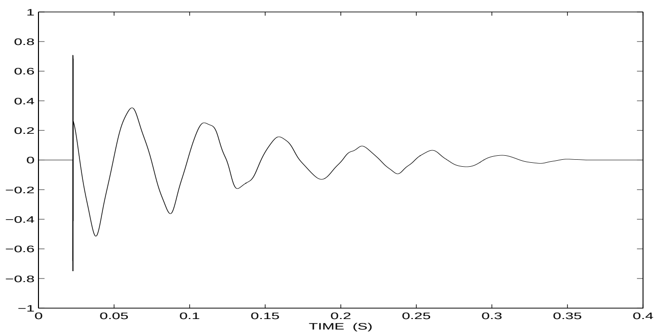


Figure 5.7: Vinyl synthetic thump.

5.1.5 Hiss

The spectrogram of a vinyl silent part can be seen in figure 5.8. The presence of clicks and hiss are easily observed. The filter with which the hiss is produced is estimated with a second-order linear predictor, and its frequency response is shown in figure 5.9.

The signal to noise ratio (SNR = $10\log_{10}(P_{signal}/P_{noise})$) is set approximately to 37 dB.

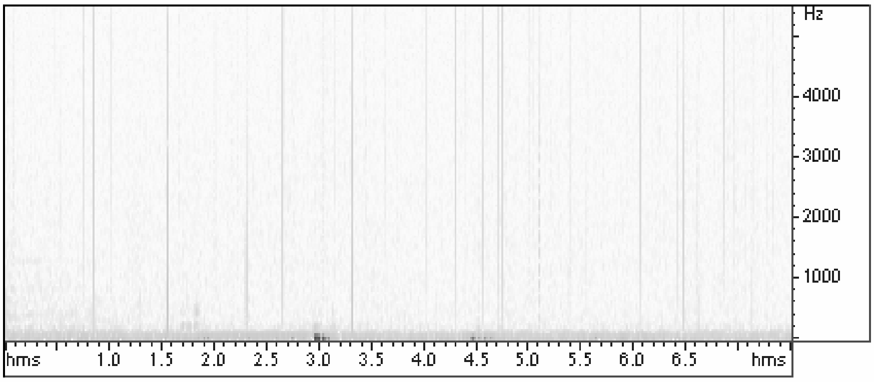


Figure 5.8: Spectrogram of a vinyl silent (no music is played) part.

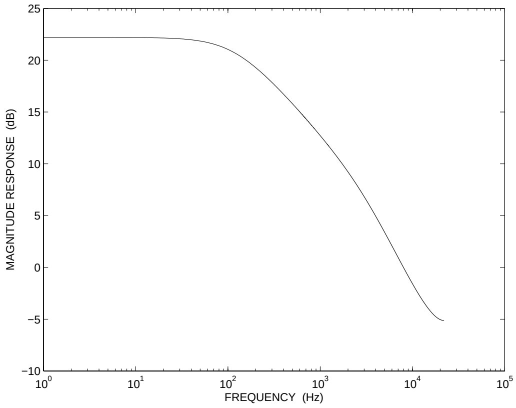


Figure 5.9: Magnitude response of the vinyl hiss filter.

5.1.6 Wow

The wow present in vinyl recordings is mainly due to a wrongly punched disk. That is why the pitch variation curve used for the addition of this degradation (see figure 5.10) is just a sinusoidal with the same period as the revolution period (Td = 60/33 s).

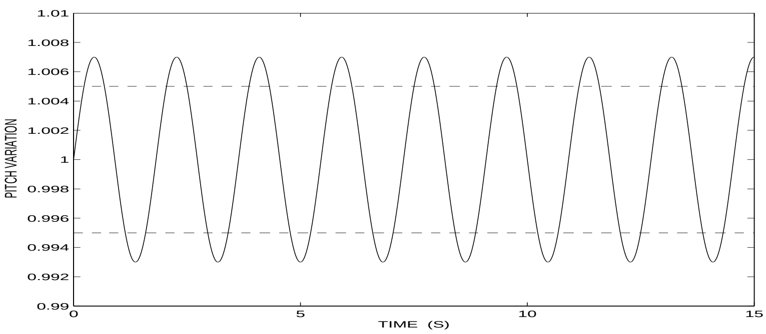


Figure 5.10: Pitch variation curve of a wrongly punched vinyl.

5.1.7 Low-Pass Filter

After the degradations simulation some non-desired high frequencies are present in the music file, so a low-pass filter with a smooth transition is used for slightly attenuating them. The parameters are shown in the next table:

Wp (Hz)	Rp (dB)	Ws (Hz)	Rs (dB)
4000	0.46	18000	10

5.1.9 Tracking Errors

When the vinyl disk has severe damages it is quite common the presence of some tracking errors during the reproduction. At the beginning of the music file some of these errors are introduced, as if the needle jumped to the previous groove reproducing time and time again the same part of the song. Finally, after three times the needle is able to continue in the correct groove.

5.2. Gramophone

To simulate the sound quality of a gramophone more types of degradations than the ones used for the vinyl simulation are needed, like the introduction of distortion in the audio waveform and the use of a horn for the music reproduction (see figure 5.11).

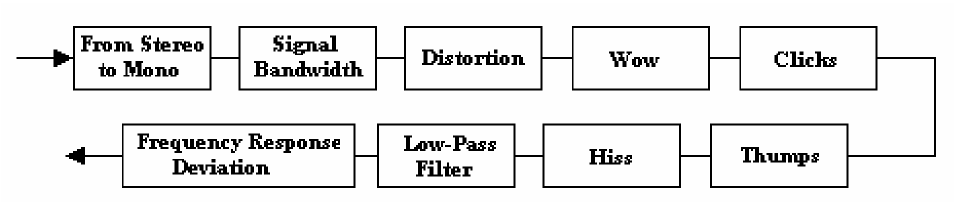


Figure 5.11: Steps followed for gramophone and phonograph quality simulation.

5.2.1 From Stereo to Mono

This step, common to the three simulations, is to calculate the mean of the two channels.

5.2.2 Signal Bandwidth

The figure 5.12 shows frequency components of a gramophone recording. The poor response for high frequencies can be observed. In this case, unlike with the LP simulation, a band-pass filter is used because of the inability of reproducing very low frequencies. The parameters of the band-pass filter are:

Ws1 (Hz)	Rs1 (dB)	Wp1 (Hz)	Rp1 (dB)	Wp2 (Hz)	Rp2 (dB)	Ws2 (Hz)	Rs2 (dB)
100	20	200	0.46	3000	0.46	5000	20

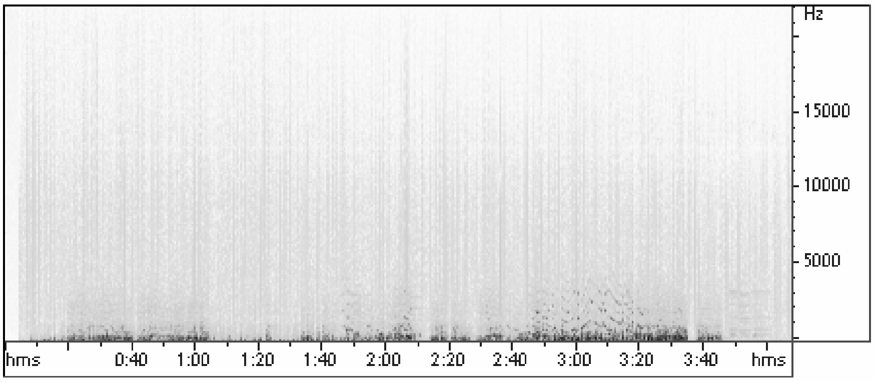


Figure 5.12: A gramophone recording spectrogram.

5.2.3 Distortion

The parameters used for this distortion are *base* = 2.5 for the loud passages and *base* = 1.8 for the soft passages. The final distortion function (the two distortions together) is represented in figure 5.13.

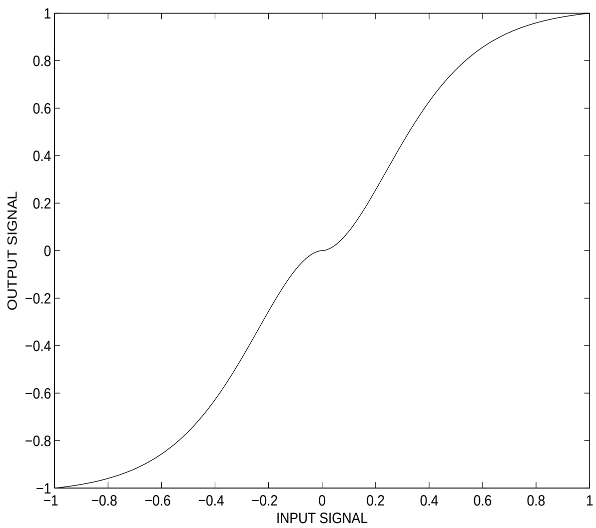


Figure 5.13: Total distortion function for gramophone recording simulation.

5.2.4 Clicks

The procedure is the same as the one that was explained for the LP, so in this case only the specific statistic functions are shown in figure 5.14, 5.15 and 5.16. The mean value used for the amplitude is 0.1.

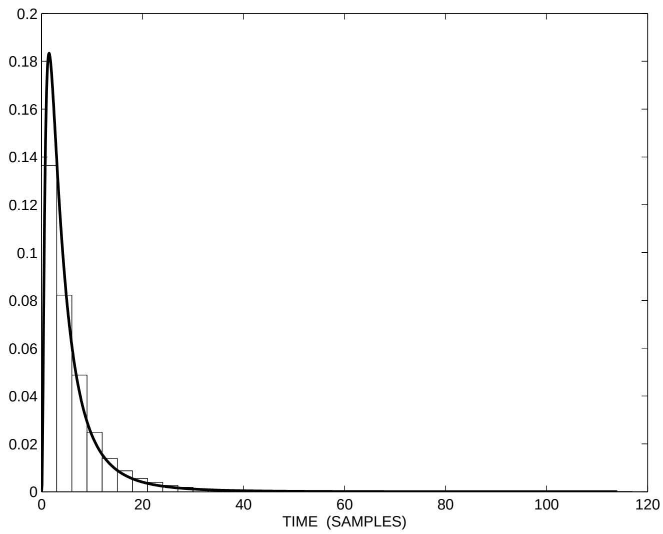


Figure 5.14: Clicks duration histogram vs. Lognormal distribution (PPCC = 0.98428).
Parameters: $\mu = 1.2811$ and $\sigma = 0.9387$ (see equation 5.3).

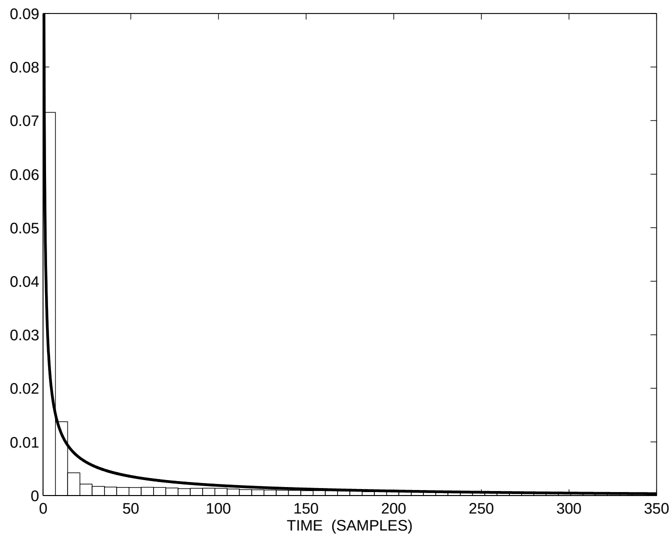


Figure 5.15: Time between clicks histogram vs. Gamma distribution (PPCC = 0.99656).
Parameters: a = 0.3378 and b = 276.6830 (see equation 5.2)

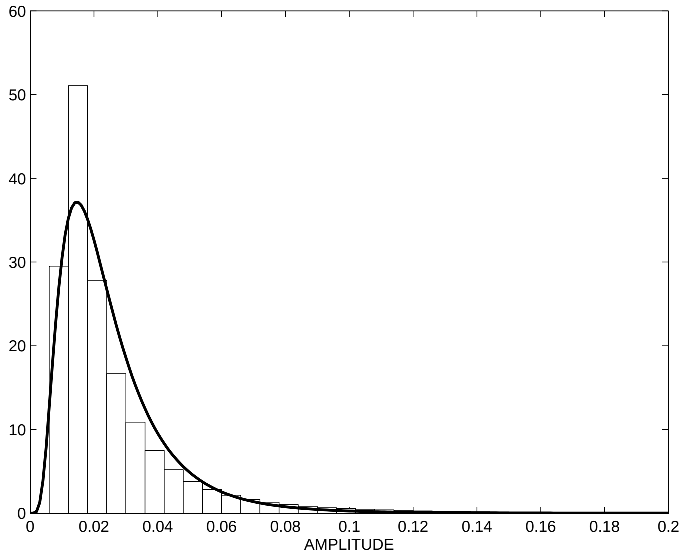


Figure 5.16: Clicks amplitude histogram vs. Lognormal distribution (PPCC = 0.94819). Parameters: μ = -3.8530 and σ = 0.6086 (see equation 5.3)

In the low-pass filter used for providing different spectral components to the different clicks the minimum normalized cut-off frequency is 0.2 while the maximum is 0.4 (see figure 5.17). The order of the Butterworth filters is 3 and the frame duration is 1000 samples.

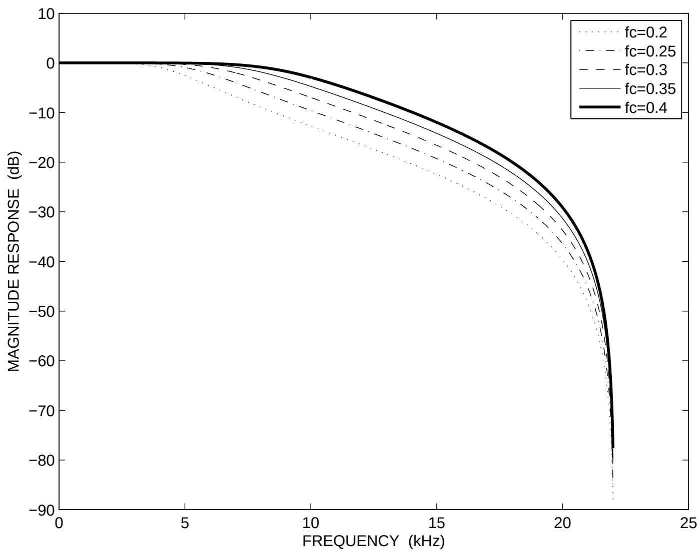


Figure 5.17: Magnitude Response of low-pass filters with different cut-off frequency. In the legend the normalized cut-off frequency for each filter is shown. The sampling rate is 44100Hz.

5.2.5 Low Frequency Pulses (Thumps)

In this case, more scratches on the surface of the disk than in the LP are supposed, 10. Their length varies between 4 and 9 grooves and their amplitude is higher than in the vinyl case (see figure 5.18)

The revolution velocity of a gramophone disk is 78 r.p.m. (revolutions per minute), having a revolution period of Td = 60/78 s, the time between the thumps produce by a scratch.

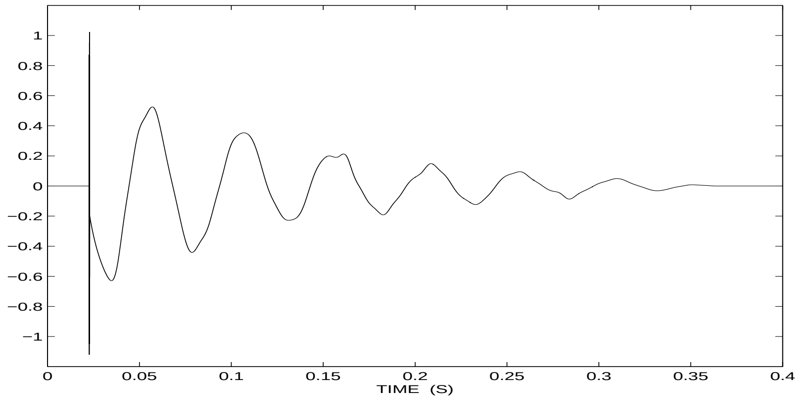


Figure 5.18: Possible gramophone thump.

5.2.6 Hiss

From a extracted part without played music (see figure 5.19) the filter for hiss production (see figure 5.20) is obtained with a linear predictor of order 4.

The signal to noise ratio is set to 30 dB.

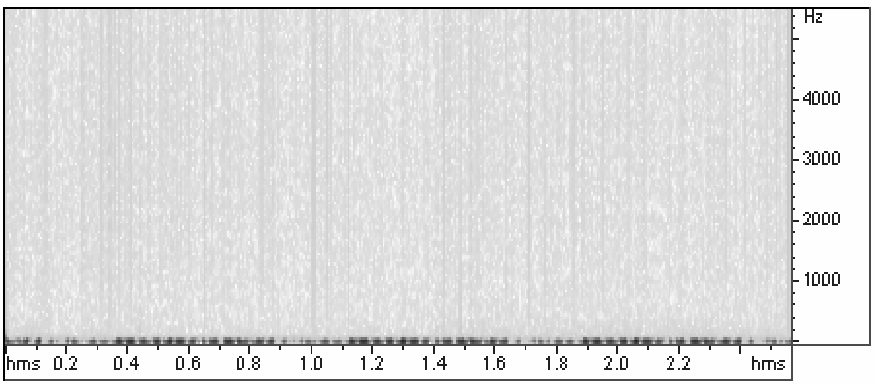


Figure 5.19: Spectrogram of a gramophone silent (no music is played) part.

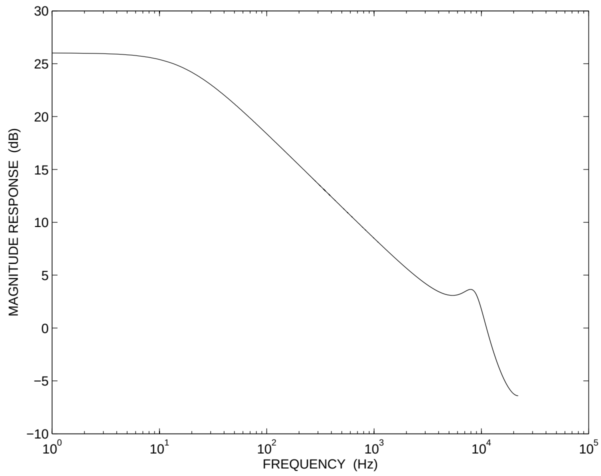


Figure 5.20: Magnitude response of the filter used for gramophone hiss simulation.

5.2.7 Wow

The pitch variation defect can be quite annoying. In this case, the pitch variation curve, although smooth, is changing along the time (see figure 5.21).

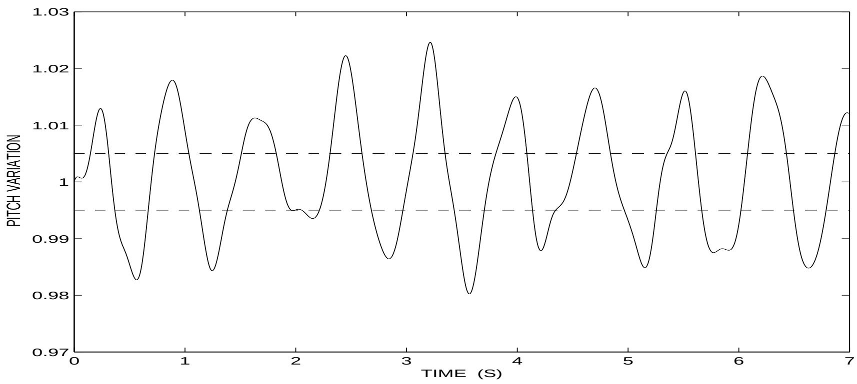


Figure 5.21: Synthetic pitch variation curve for a gramophone disk simulation.

5.2.8 Low-Pass Filter

As in the previous case, a low-pass filter is needed for removing the high frequency components introduced along this process which could not be possible in the gramophone disks.

Wp (Hz)	Rp (dB)	Ws (Hz)	Rs (dB)
3000	0.46	19000	20

5.2.9 Frequency Response Deviation

The horn used in this part was a Victor petalled horn appropriate for Victor II, III, IV or similar machines. It is approximately 56 cm long with 48 cm bell. Although it is made with 8 petals, the simulation is for a hole piece exponential horn due to the limitations of the used program ('Hornresp'). Some pictures of the horn can be observed in figure 5.22 and its frequency response is represented in figure 5.23.

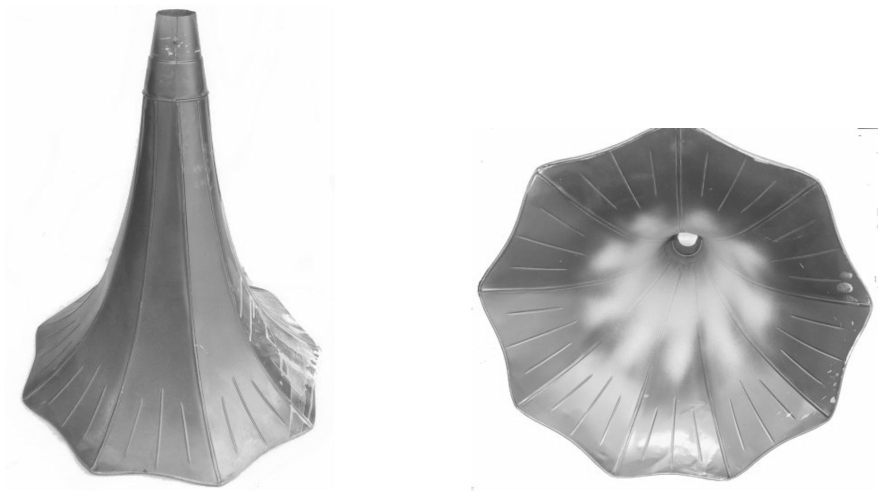


Figure 5.22: Victor petalled horn
(http://www.intertique.com/VictorBlueHorn.html)

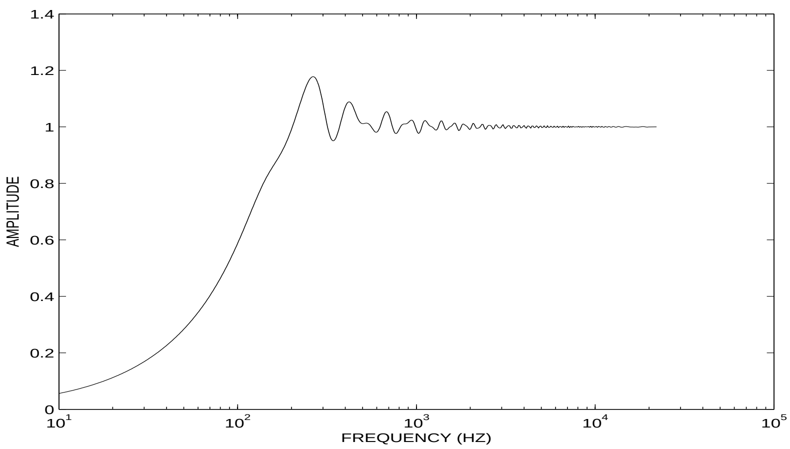


Figure 5.23: Simulated transmission coefficient of a Victor petalled horn.

5.3. Phonograph

This device is the oldest in reproducing and storing sound, and the degradations are in this case more severe as in the previous ones. The steps for achieving the desired quality are the same than the followed in the gramophone, but with more harmful parameters.

5.3.1 From Stereo to Mono

The average of the left and right channel is done.

5.3.2 Signal Bandwidth

The spectrogram of a phonograph recording is represented in figure 5.24. As can be observed, the bandwidth is quite narrow, missing low and high frequencies. The filter used in this case is a band-pass filter and its parameters are:

Ws1 (Hz)	Rs1 (dB)	Wp1 (Hz)	Rp1 (dB)	Wp2 (Hz)	Rp2 (dB)	Ws2 (Hz)	Rs2 (dB)
400	23	1000	0.46	2000	0.46	4000	20

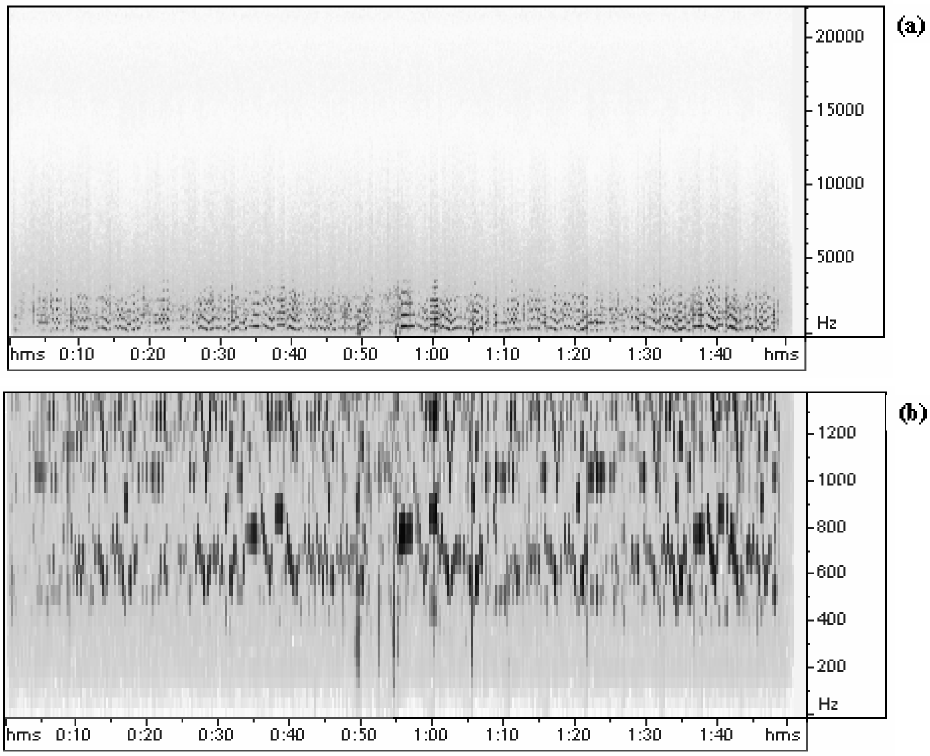


Figure 5.24: Phonograph recording spectrogram. (a) All the frequency components. (b) Low frequencies zoomed.

5.3.3 Distortion

The distortion in phonograph cylinder is an important feature. When listening to a phonograph recording, effects like clipping can be identified. The parameters for the distortion are higher than in the gramophone, with *base* = 3 for loud passages and *base* = 2 for soft passages. In figure 5.25, the final distortion curve is printed.

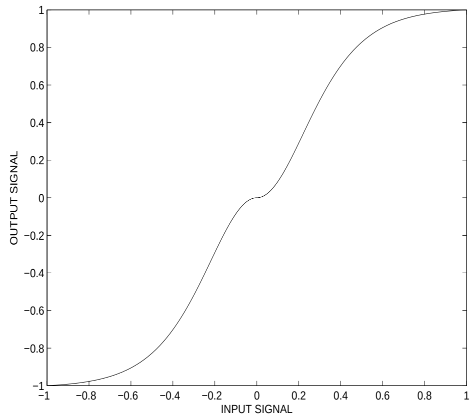


Figure 5.25: Total distortion function for phonograph recording simulation.

5.3.4 Clicks

As in the previous cases, the distinct statistical functions for each parameter are presented in figure 5.26, 5.27 and 5.28. The mean value for the amplitude is set to 0.07, the same than the expected value of the distribution. This time, there is no need for changing the obtained amplitude of the clicks.

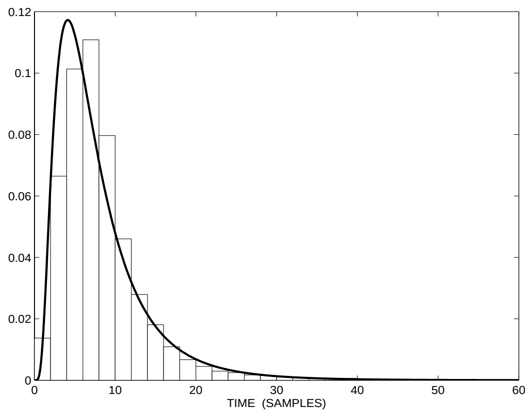


Figure 5.26: Clicks duration histogram vs. Lognormal distribution (PPCC = 0.9939). Parameters: μ = 1.8561 and σ = 0.6617 (see equation 5.3)

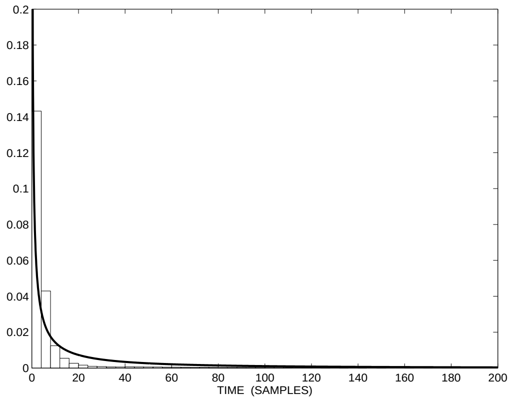


Figure 5.27: Time between clicks histogram vs. Weibull distribution (PPCC = 0.98748). Parameters: a = 17.1571 and b = 0.3975 (see equation 5.1)

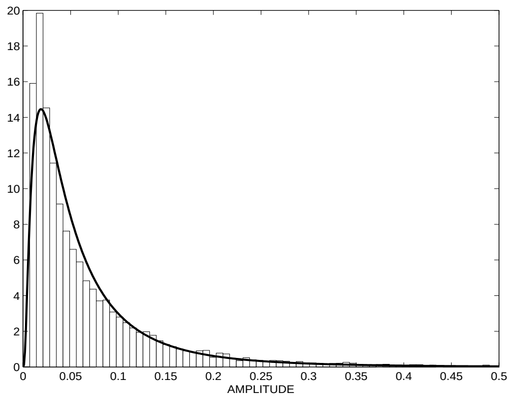


Figure 5.28: Clicks amplitude histogram vs. Lognormal distribution (PPCC = 0.9903). Parameters: μ = -3.0870 and σ = 0.9410 (see equation 5.3)

Some of the filters used for adding dynamic to the clicks are represented in figure 5.29, with a normalized cut-off frequency varying from 0.1 to 0.4.

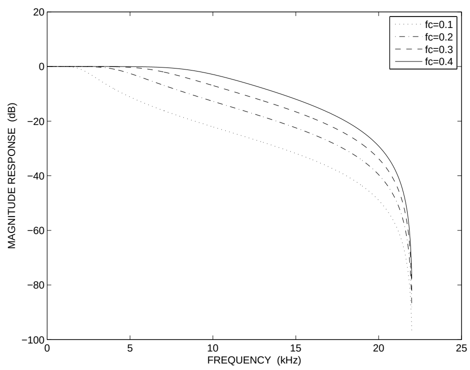


Figure 5.29: Magnitude Response of low-pass filters with different cut-off frequency. In the legend the normalized cut-off frequency for each filter is shown. The sampling rate is 44100Hz.

5.3.5 Low Frequency Pulses (Thumps)

For the phonograph cylinder simulation two different kinds of scratches are supposed: soft and severe. Thirteen scratch of the first type were simulated with a length varying from 4 to 9 grooves and four strong scratches with 10 to 13 grooves length. Figure 5.30 (a) shows a soft thump while figure 5.30 (b) represents a strong one.

The phonograph cylinder is revolving with a velocity of about 120 r.p.m. (revolutions per minute), so the time between thumps is Td = 60/120 s.

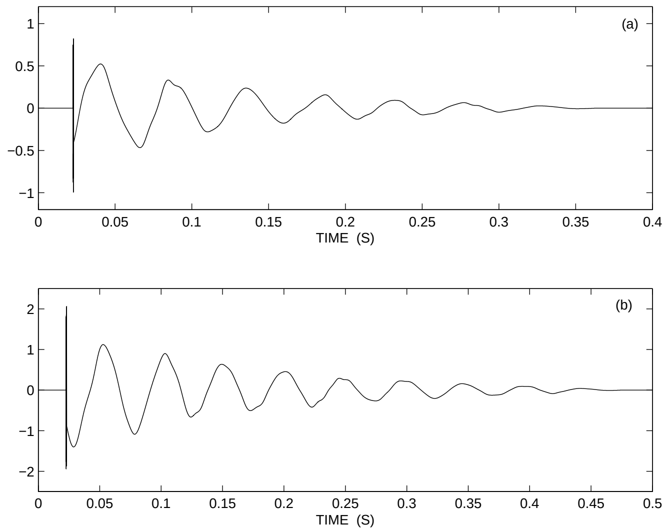


Figure 5.30: (a) Normal synthesized phonograph thump. (b) Strong synthesized phonograph thump.

5.3.6 Hiss

The signal to noise ratio for the hiss is around 23dB, and the filter (see figure 5.31 and 5.32) is obtained with eighth order linear predictor. A resonance at about 1 kHz is observed in the calculated filter.

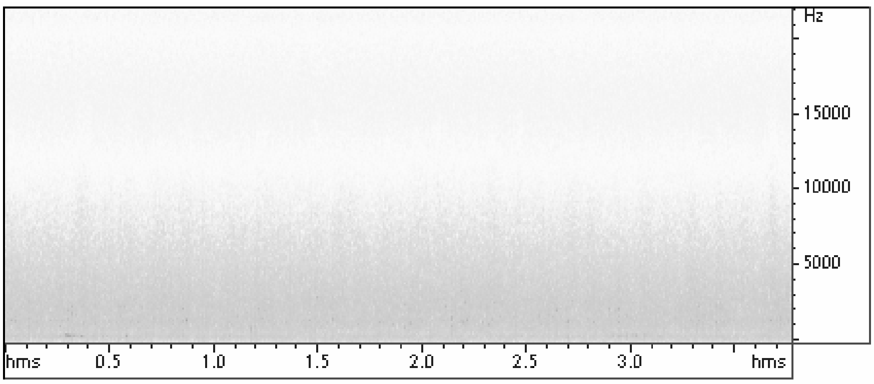


Figure 5.31: Spectrogram of a phonograph silent (no music is played) part.

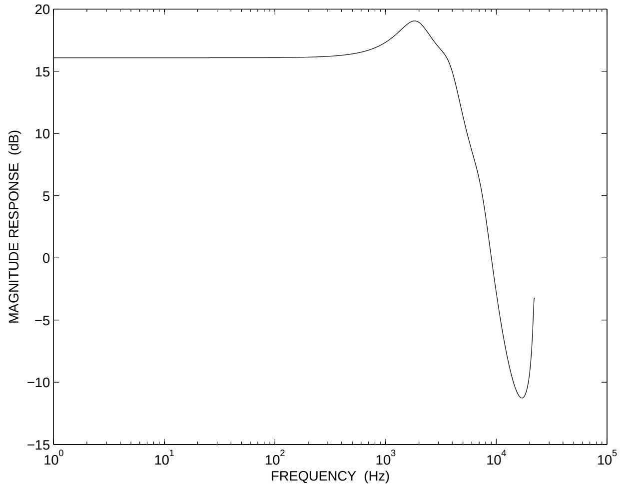


Figure 5.32: Magnitude response of the filter used for phonograph hiss simulation.

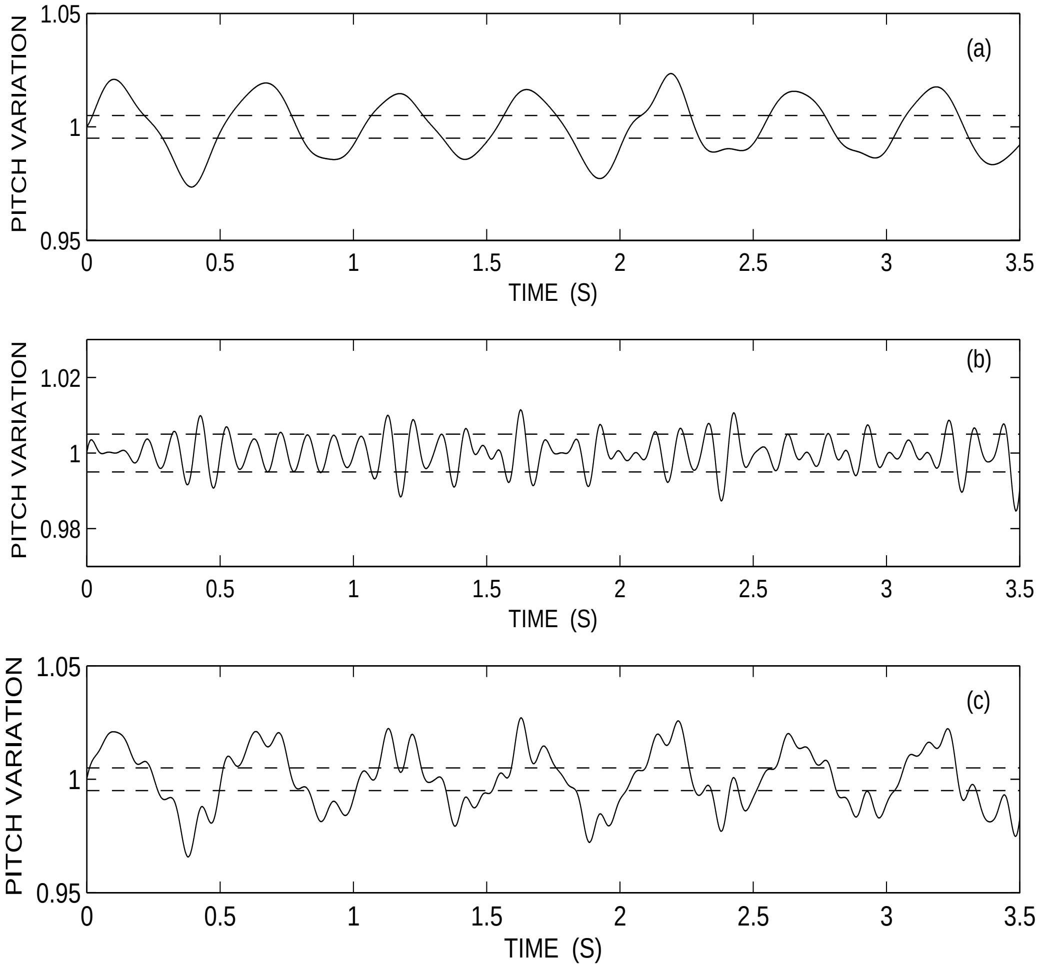
5.3.7 Wow and Flutter

The pitch variation curve in this recording medium is due to wow and flutter. While the mean period of the defect is the same as the revolution period (Td = 60/120 s), the flutter variations are faster, and the chosen value for its mean frequency is 10 Hz. The final pitch variation curve is formed as the combination of these two degradations (wow and flutter). Some examples of these curves are shown in figure 5.33.

5.3.8 Low-Pass Filter

In this case, the stop-band and pass-band corner frequency are much lower than in the vinyl or gramophone simulation due to the narrow bandwidth provided by the phonograph cylinders.

Wp (Hz)	Rp (dB)	Ws (Hz)	Rs (dB)
2000	0,46	7500	20



.Figure 5.33: (a) Simulated wow for phonograph recordings. (b) Simulated flutter for phonograph recordings. (c) Simulated total pitch variation curve for phonograph recordings.

5.3.9 Frequency Response Deviation

An antique horn for Columbia Standard Phonograph (figure 5.34) is used as the base for the simulation. This horn is built with eight scalloped, metal panels. It measurements are approximately 41 cm long by 44.5 cm in diameter. The opening at the tip measures 2.7 cm diameter. As it was explained before, the simulation of the transmission coefficient (figure 5.35) is done for an exponential horn constructed in one piece.

Figure 5.34: Antique horn for Columbia Standard Phonograph.
(http://cgi.ebay.com/Original-Metal-HORN-for-COLUMBIA-STANDARD-PHONOGRAPH_W0QQitemZ230139263128QQihZ013QQcategoryZ38027QQrdZ1QQcmdZViewItem)

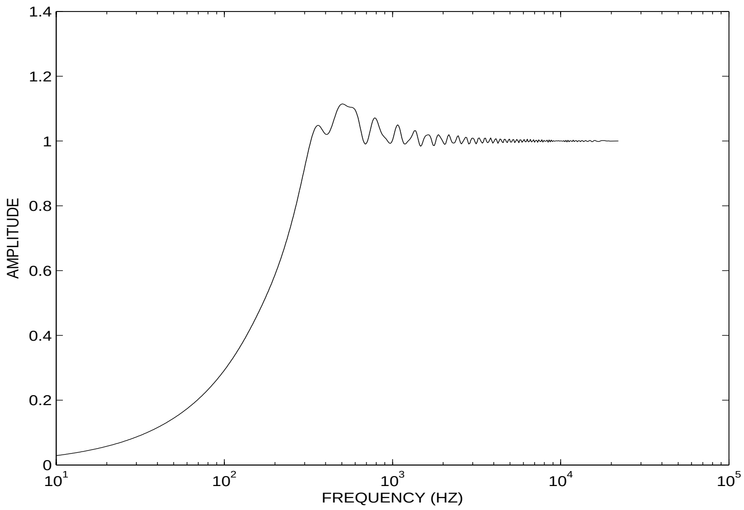


Figure 5.35: Simulated transmission coefficient for a phonograph horn.

Chapter 6

Conclusions

The aim of this project has been to simulate the quality of phonograph, gramophone and LP recordings starting from the quality of a CD. In this work, the different kinds of degradation in these historical recordings were first studied on a theoretical level, analyzing their causes and effects. After that, the best way to simulate each of them was studied. More emphasis was put on the implementation of the more audible ones, like clicks or wow. Finally, suitable parameters for the synthetic degradations were examined for each simulation. When all the data was collected, the needed degradations were added to the initial audio file to achieve the desired quality. In general, the results were reliable, although some manual adjustments could still be made.

Future research could include features like source separation and modification of instrument sounds one at a time (e.g., make the guitar sound older), canceling of effects processing such as de-compression of pop music files, imitation of live recording quality (e.g., a rock concert in a park or a stadium including sounds of the audience), etc. Moreover, the simulation of other recording devices, such as the magnetic tape or C cassette could be implemented.

Bibliography

B. B. Bauer, *Notes on Distortion in Phonograph Reproduction Caused by Needle Wear.* J. Acoust. Soc. Am. vol. 16, issue 4, pp. 246, April 1945

Colección F.B., *Del cilindro al disco.* http://www.coleccionfb.com/del_cilindro_al_disco.htm, 2007a

Colección F.B., *Primeros Registros.* http://www.coleccionfb.com/Primeros%20registros.htm, 2007b

A. Czyzewski , P. Maziewski, M. Dziubinski, A. Kaczmarek and B. Kostek, *Wow detection and compensation employing spectral processing of audio,* 117 Audio Engineering Society Convention, Convention Paper 6212, San Francisco, October 2004.

Encyclopedia Britannica, *Compact Disc: Dynamic Range.* http://www.britannica.com/eb/article-92850/compact-disc, 2007.

P. A. A. Esquef, L. W. P. Biscainho, P. S. R. Diniz and F. P. Freeland, *A double-threshold-based approach to impulsive noise detection in audio signals*. Proc. X European Signal Processing Conf. (EUSIPCO 2000), pp. 2041-2044, Tampere, Finland, September 2000.

P. A. A. Esquef, L. W. P. Biscainho and V. Välimäki, *An Efficient Algorithm for the Restoration of Audio Signals Corrupted with Low-Frequency Pulses*. Journal of the Audio Engineering Society, vol. 51, no. 6, pp. 502-517, June 2003.

J. C. Fesler, *Electrical Reproduction of Acoustically Recorded Cylinders and Disks, Part 2*. Journal of the Audio Engineering Society, vol. 31 number 9 pp. 674-694; September 1983.

N. H. Fletcher and T. D. Rossing, *The physics of musical instrument*. Springer-Verlag, 1991.

U. R. Furst, *Periodic variations of pitch in sound reproduction by phonographs.* Proc. IRE, vol. 34, pp. 887-895, November 1946.

S. J. Godsill and P. J. W. Rayner, *Digital Audio Restoration – A Statistical Model Based Approach*. London, UK: Springer-Verlag, 1998.

IEEE Virtual Museum, *The Phonautograph.* http://ieee-virtual-museum.org/collection/tech.php?taid=&id=2345805&lid=1, 2007.

Izotope, *Izotope Vinyl.* http://www.izotope.com/products/audio/vinyl/, 2007.

I. Kauppinen, *Methods for detecting impulsive noise in speech and audio signals*. Digital Signal Processing, 2002 14th International Conference, vol. 2, pp. 967- 970, July, 2002.

O. Kimmelma, J. Parviainen and V. Välimäki, *Musiikkiäänitteiden digitaalinen vanhentaminen*. "Suomen musiikintutkijoiden symposiumin satoa", URL http://www.jyu.fi/musica/symposium/julkaisu2005/, pp. 55-58, 2005.

J. H. Kogen, *Tracking Ability Specifications for Phonograph Cartridges*. Journal of the Audio Engineering Society volume 12, number 2, pp. 186-190, April 1968.

Laser Turntable, *High Fidelity Pickups*. http://www.smartdev.com/LT/Align.htm, 2006.

C. A. Luszick, *Digital Audio Restoration of Vinyl Records and Implementation on a Digital Signal Processor based System*. Digitale Signalverarbeitung Universitäat Kaiserslautern, Diploma Thesis, November 2003.

P. Maziewski, *Wow defect reduction based on interpolation techniques*. IV Krajowa konferencja elektroninki, 2005.

D. J. McBean, *Horn Loudspeaker Response Analysis Program*. http://www.users.bigpond.com/dmcbean/, 2007.

T. Muraoka, H. Onoye and J. M. Eargle, *On the measurement of phonograph cartridges by the pulse-train method (Part II).* Audio Engineering Society Preprint no. 1088, 1975.

Museum of Sound Recording, *History of Sound.* http://www.museumofsoundrecording.org/history_sound.htm, 2003.

T. D. Rossing, F. R. Moore and P. A. Wheeler, *The Science of Sound, 3rd Edition*. Addison-Wesley, 2002.

E. G. Trendell, *Tracing Distortion Correction*. J. Acoust. Soc. Am. vol. 25, number 5, pp. 273, 1977.

V. Välimäki, *Digital Audio Effects and Reverberation Algorithms.* S-89.3540 Audio Signal Processing, Lecture 6, March, 2007.

Wikipedia, *Compact Disc*. http://en.wikipedia.org/wiki/Compact_Disc, 2007a.

Wikipedia, *Respuesta en frecuencia en audio.* http://es.wikipedia.org/wiki/Respuesta_en_frecuencia, 2007b.