Digital Audio Interface Design

'Project'
Module Code: MU70004E

Student: Josh Fairhead (21057665)
Lecturer: Sebastian Lexer

Table of Contents

Introduction

The main goal behind this project was to understand how various forms of digital processing effect audio within the context of a mix. It was chosen because my interest lies in studio work where these processes are ever present and so an understanding of how they function was needed to properly assess side effects.

When mixing I often turn to harmonic saturation because it can alter an instruments timbre in a more useful way than an equaliser and so my patch is an attempt to create this effect for the guitar in a musically related way. The rules of western music theory have their basis in maths beginning with Pythagorean tuning which was an imperfect method due to increasing dissonance in remote keys. Equal temperament was developed but required a compromise of slightly increased dissonance in the home keys in exchange for increased musical flexibility[1]. This slightly dissonant tuning method is what we base our musical expectations on due to continued listening since birth. For this reason I believe that the various distortions we hear (any change from input to output signal), are subjectively liked or disliked according to the same musical rules but on a more granular level.

Being new to MaxMSP and having little math most of my initial theories and attempts to create this musical saturation resulted in failed attempts. The patch created has its flaws but shows an attempt at working with traditional processes (filters, etc.) in a different way.

During the patches period of creation I have dealt with subjects such as synthesis, audio analysis, plug-in design, FFT and effect units. Looking into these areas has been a great lesson in the fundamentals of digital audio and consequentially an enjoyable experience despite the difficulties encountered throughout the process.

The Project

As explained several attempts were made at creating this patch. The methods used were often changed as new knowledge was acquired from research along with learning from the various failed attempts.

Many papers discuss harmonic distortion in terms of an input sine wave compared the output sine waves; these sine waves are harmonics of the fundamental related by octave. Typically the nonlinear distortions in studio equipment colour the sound in some way, these are observable to an extent by running a sine wave through the input and observing the harmonics created on output.

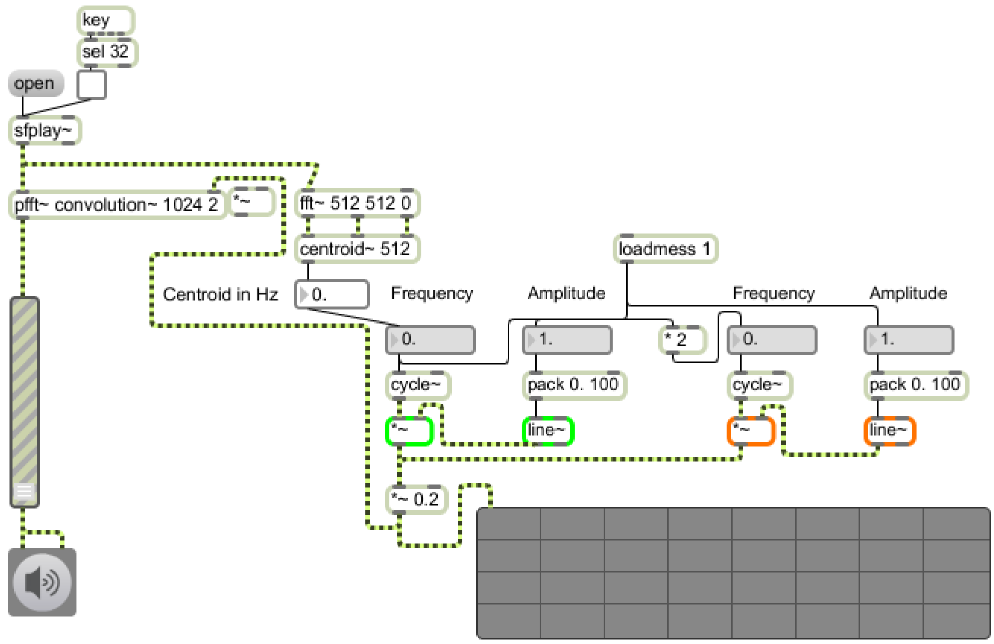
These harmonics and their structure are the focus of several papers based on the subjective sound quality of equipment design particularly Russell O'Hamms paper *Tubes vs. Transistors-Is There an Audible Difference?*[2]. Hamms paper discusses the relevance of different harmonics and how their typically perceived; Its mentioned in this paper that the second harmonic is often perceived as creating the perception of added warmth. This is interesting because the second harmonic is also the closest octave to the fundamental and musically speaking is often played as a strengthener note[3].

To get an idea of how the harmonics sound I wired a few oscillators from Native Instruments FM8s in parallel sounding at f, $2f$, $3f$, $4f$ and $5f$. I then varied the volume of different harmonics and thought about how I perceived the tone to change when certain harmonics were more dominant than others. Like Hamm mentions in his paper adding the second made the tone seem warmer, where as

1 Jourdain, R. (1997) *Music the Brain and Ecstasy* – 1st Edition, Harper Collins, US.
2 Hamm, R. (1973) Tubes Versus Transistors-Is There an Audible Difference – AES Journal, volume 21, Issue 4.
3 This is based on personal observation; many chords on many instruments are built from a close range and so the first octave or $2f$ is usually present in music as a strengthener note to the fundamental/root within a chord.

adding the third added edge and bite.
From reading papers with an emphasis on the importance of the different harmonics generated from a change to the input sine waves shape, I mistakenly oversimplified the process thinking that the same would apply to a complex signal. I later learned that this process is the typical practice of finding the %THD (Total Harmonic Distortion) reading of a device; the harmonically related waves are typically measured in RMS voltage retaliative to the total RMS and expressed as a percent to give a %THD[4].

Wrongly believing the topic of saturation to be this simple I concluded that an oscillator bank could be built, tracked to the fundamental and then convoluted with original audio to introduce a chosen amount of saturation for each harmonic by adjusting the various oscillator levels. I thought that boosting the just the fundamental of a chord and its relative harmonics in this way would result saturation and so I started to build the patch in MSP.



Firstly I found an external for pitch tracking called centroid~ which output numbers in Hz. I wrongly thought centroid~ would track the fundamental of an input but later learned that it actually tracked the weighted mean of frequencies present in a signal. With hind sight this explains why the pitch tracking in this patch always seemed to sound wrong.

I then looked into convolution through FFT and learnt about its various possibilities along with the practical problems such as temporal vs. frequency resolution. This trade off is down to the FFT window size; as an FFT essentially compares a real signal with a self generated imaginary signal the size of the its window must be at least one cycle of the lowest frequency needing detection. However because energy is averaged over the FFT window the larger it is, the more timing is smeared, which can have a pronounced effect on transients. When I eventually convolved the source audio with the tracked oscillator bank I found the result completely unmusical and realised a different approach was needed.

4 Nachbaur, F. (2004) THD Measurement and Conversion – [online: http://www.dogstar.dantimax.dk/tubestuf/thdconv.htm]

The article *Tube Screamers Secret*[5] explains about the importance of phase shifts and its contribution to the presence of a second harmonic. I thought that perhaps manipulating phase in the frequency domain from a cartopol~ object (after a pfft~ object) might produce some musical results but found it not to work either. Also I tried to introducing a phase shift by altering playback speed dependent on amplitude with kink~ object; this worked fine with a sine wave but of course sounded completely unmusical on a complex signal.

After these failures I realised that a complex signal wouldn't work in such a simple way; the main reason was that I hadn't thought about the intermodulation distortion caused by two or more waves interacting. Whats typically referred to as harmonic distortion occurs when the shape of a sine wave is modified generating harmonics related by octave to the fundamental; however when two or more sine waves interact, frequencies are introduced into the signal at the sum and difference. This is intermodulation distortion and not surprisingly it can get incredibly complex when a few frequencies are present at once.[6]

Langevins paper *Intermodulation Distortion in Tape recorders* examines the relationship between harmonious material and the musical significance of the output intermodulation products. It examines the sum and difference of four harmonious intervals; the octave, the fifth, the fourth and the major third. The intermodulation distortion created from these intervals are shown to typically land on musically significant notes or at least within the scale. As an exercise I calculated the products of a few different intervals and found that intermodulation products were mostly harmonious (with a few exceptions) and as stated by Langevin 'It therefore follows that if the music is harmonious there is a good chance IM products will also be harmonious and not too noticeable'[7].

This paper led on to the thought that boosting the fundamental and its harmonics into a clipping stage could create musical byproducts from the intermodulation distortion, however creating this boost was still a problem and soon I found myself looking into digital filters. To gain a practical understanding of the concepts I started to build the *Computer Music Tutorials*[8] various example diagrams in MSP and discovered that creating an audio feedback loop would not work (similar to the Max environments data loops) which seemingly made IIR feedback designs impossible until I discovered the tapin~ and tapout~ objects; these can be used to bypass the problem by storing and receiving data from the same buffer, tapin~ storing a continuous section of incoming signal with tapout~ reading a copy at a specified time interval.

Working through this book soon led me on to reading about comb filters and how they cause frequencies cancel in a harmonic series. Comb filter designs like many filters come in two variations, Finite Impulse Response (FIR) and Infinite Impulse Response (IIR). In FIR filters a copy of the input is delayed causing frequencies to phase cancel, in equalisers this is typically delays of one/two samples cascaded, while in comb filters its usually a single delay that is significantly longer creating the characteristic peaks/notches. In IIR designs some of a delayed output is combined with the input signal, again in equalisers this is generally taps of a sample or two delay with several filters stacked in cascade, in comb filters the delay times are generally longer creating the harmonically spaced peaks and notches.

---

5 Topaktas, B. (2005) Tube Screamers Secret – [online: http://www.bteaudio.com/articles/TSS/TSS.html]
6 Barbati, S. (unknown) *A Perceptual Approach on Clipping and Saturation* - Simulanalog.
7 Langevin, R. (1963) *Intermodulation Distortion in Tape Recorders* – AES Journal, Volume 11, Issue 3.
8 Roads, C. (1996) *The Computer Music Tutorial* – 1st Edition, MIT Press, UK.

A significant bit of information learned from this research into filter design was the effect that nearly all processing has on the phase of a waveform; particularly in equalisation because of its many stacked delays at different amplitudes. The curve of an equaliser is essentially determined from relative delays and volumes numerically represented as filter coefficients.. In particular equalisers with the IIR design are detrimental to phase with ringing caused at high Q values from excessive feedback levels in the filter[9] (Bob Katz also mentions this in his book *Mastering Audio*[10]). In FIR designs we still get phase shifts however it is even across the spectrum and is commonly called a linear phase eq for this reason, these are more processor intensive as many more delay stages must be calculated instead of using a feedback design to do the work[11].

Regardless of filter design its now clear that the timing effects of such phase distortion can be detrimental to musical signals; particularly rhythmic instruments such as the drums. One experiment I tried was to line up hits on a grid and destructively eq them with Audio Suite effects (in Pro Tools and quite radically) so that the timing shifts could be seen by their relative offsets to the grid, which were quite obvious upon inspection.

In the *Computer Music Tutorial* it was also explained that the frequencies affected by the comb filter are proportional to the feedback signals delay time. For this reason I believed that analysis could be used in order to track a fundamental frequency and that this information could then be converted into a corresponding delay time for a comb filter so as to boost (or cut) the tracked frequency and its harmonics. This tracked comb filter could be followed by a bias and clipping stage which I thought could create more musically interesting intermodulation products.

The idea of a bias stage came from reading various different writings on amplification during which I came across an interesting article on Universal Audios website in the 'Ask the doctors' section[12]. Here asymmetry was associated with a more pronounced second harmonic and because of these and other writings such as the Topaktas article[13], it made sense to include a bias stage in my patch to help provide additional tonal coloration, which I placed after the amplification stage.

Before building the patch I found a few pitch tracking externals and tested for the most accurate by playing a recording of full chords with known fundamental and examining the results (e.g. A major should be reported as 110Hz regardless of the other notes present). The most accurate of the externals was the pitch~ object, which is based upon Miller Puckettes fiddle~, and so it was used for pitch tracking within my patch.

The pitch~ object outputs the fundamental frequency of an input and the following equation was used to find the correct sample delay in order to create notches in the comb filter at the same frequency: $f = 1/D \times f s$ where fs is sampling frequency[14] which was obtained through the dspstate~ object. The output result (in samples) was then converted with the sampstoms~ object into a millisecond delay time which was the required format for the comb~ objects second input. This stage now allowed boosting (or cutting if required) of the fundamental and its relative harmonics through control of the filters coefficients. High feedback amounts from the IIR stage causes the most obvious distortions while manipulating the FIR coefficient control is more subtle.

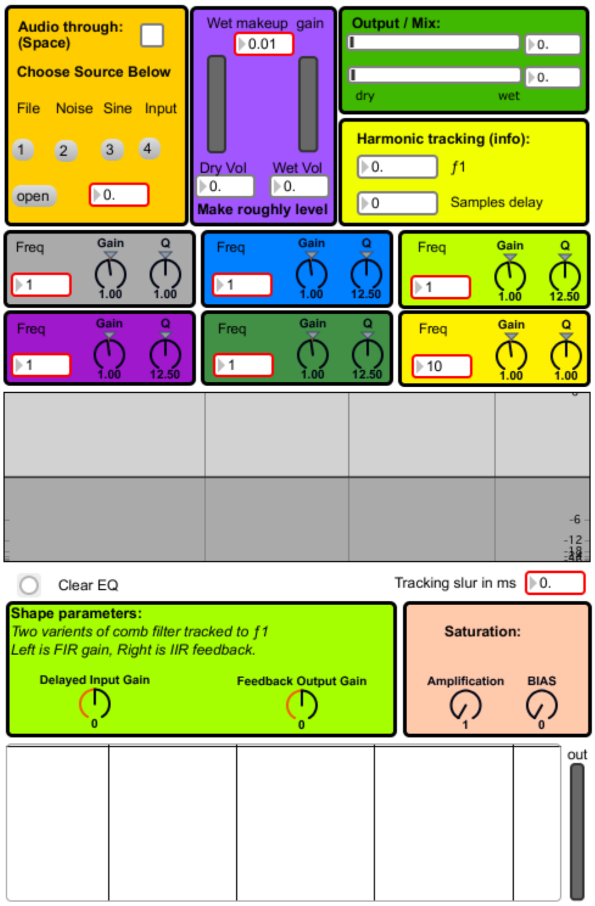
9 Kemp, M (2003) *Sampling Equalisers* – Resolution Magazine, S2 Publications.
10 Katz, B. (2006) - *Mastering Audio; The Art and the Science* - 1st Edition, Focal Press, US.
11 Roads, C. (1996) *The Computer Music Tutorial* – 1st Edition, p410, MIT Press, UK.
12 Berners, D. (2005) *Ask the doctors: Tube vs. Solid-state Harmonics* – [online: http://www.uaudio.com/webzine/2005/october/index2.html]
13 Topaktas, B. (2005) *Tube Screamers Secret* – [online: http://www.bteaudio.com/articles/TSS/TSS.html]
14 Roads, C. (1996) *The Computer Music Tutorial* – 1st Edition, p414, MIT Press, UK.

As the boosted frequencies are harmonious they were followed by an amplification stage followed by a tanh~ object for clipping in order to create the musically related intermodulation products referenced by Langevin[15].

Because I had read about filter design already I decided to put an equaliser before the comb and clipping stage to pre-shape the incoming audio; this was a matter of of connecting a filtergraph~ object to a cascade~. From my normal working process I decided that I wanted to control the equaliser through dials rather than the objects graph; as filtergraph~ only has inputs for controlling one of its available filters a workaround documented in the help files was necessary.

Several filtergraph~ objects were packed and prepened with the cascade message, output to a main filtergraph~ for visual reference and finally the output coefficients were sent to a cascade~ object to processed the audio. I also tried replacing the comb filter with a couple of tracked notch/resonant filters but found the results unusable.

The Interface



The orange section at the top left of the patch controls the audio; clicking the audio through toggle or hitting space turns on the audio converters as well as playing an audio file if its loaded. The signal source may also be chosen in this section which may be a file, noise, a sine wave or a live input. The noise and sine wave options are just for testing purposes.

In the next sections (purple and green panels) the dry volume reports its value and the user should adjust the wet gain to match so as to allow the dry/wet mix to blend smoothly with no volume adjustments. An output gain is also included for necessary adjustment.

The yellow panel reports the frequency being tracked and the equivalent samples delay that is being output to the comb filter.

Below this is an equaliser for pre-shaping the audio, it is placed before the comb filter stage because the signal flows this way. Just below the equalisers graph there is a button to clear the values and a tracking slur box. The tracking slur controls how quickly the comb filter follows the tracked pitch and should be adjusted to avoid jittery audio while still tracking fast enough to boost/cut the notes being played.

Just below this in the green panel are two controls for the comb filter which control the parameters of its shape; each controlling coefficients of the filter (FIR and IIR respectively). Turning either up causes more distortion but makes the filter results less stable. The co efficients can be set to roughly cancel each other by turning them in opposite directions, boosting is a matter of turning either clockwise and cutting by turning either counter clockwise (cutting rarely sounds useable). The amplification and bias dials are used to clip the signal after the comb~ stage.

15 Langevin, R. (1963) *Intermodulation Distortion in Tape Recorders* – AES Journal, Volume 11, Issue 3.

Critical analysis

The final outcome works but has a critical flaw; the tracking can be slightly off at times which creates a jittering sound, this can be fixed by adjusting the tracking slur rate however it is not ideal as it limits the playing style; as a patch it works to an extent, but I would rarely use it if it were a commercial plugin due to these shortcomings.

The positive outcome from the experience was that I developed a better understanding of digital signal processing; as stated before these processes are routinely used in mixing without knowledge of the negative side effects. Although I may not know how to build an algorithm which generates filter co-efficients, the more granular effects of such processing like instance phase distortion and intermodulation distortion,have been made clearer to me.

In the future I would probably approach this subject differently; from reading Micheal Kemps article called *Sampling Equalisers*[16] it seem that filter co-efficients can be sampled from existing hardware and used to build emulations, perhaps the same technique could be applied to emulate the musical qualities of tape or valve technology in a more stable manner. When a single sample is passed through a piece of hardware and rerecorded we can see the phase distortion from examining the input against output. It may be that filter coefficients can be generated from sampling this way, using each sample to set the coefficients of an FIR filter banks relative delays/volumes could perhaps work in emulating hardware, this is speculation however requiring further research.

Word count: 2746

16 Kemp, M (2003) *Sampling Equalisers* – Resolution Magazine, S2 Publications

Bibliography

Barbati, S. (unknown) *A Perceptual Approach on Clipping and Saturation* - Simulanalog.

Berners, D. (2005) *Ask the doctors: Tube vs. Solid-state Harmonics* – [online: http://www.uaudio.com/webzine/2005/october/index2.html]

Hamm, R. (1973) Tubes Versus Transistors-Is There an Audible Difference – AES Journal, volume 21, Issue 4.

Jourdain, R. (1997) *Music the Brain and Ecstasy* – 1st Edition, Harper Collins, US.

Katz, B. (2006) - *Mastering Audio; The Art and the Science* - 1st Edition, Focal Press, US.

Kemp, M (2003) *Sampling Equalisers* – Resolution Magazine, S2 Publications

Langevin, R. (1963) *Intermodulation Distortion in Tape Recorders* – AES Journal, Volume 11, Issue 3.

Nachbaur, F. (2004) THD Measurement and Conversion – [online: http://www.dogstar.dantimax.dk/tubestuf/thdconv.htm]

Roads, C. (1996) *The Computer Music Tutorial* – 1st Edition, p410, MIT Press, UK.

Topaktas, B. (2005) *Tube Screamers Secret* – [online: http://www.bteaudio.com/articles/TSS/TSS.html]

FINAL GRADE

71 / 100

GENERAL COMMENTS

This project has a convincing outcome resulting from detailed research undertaken, which is particularly successful within your writing. The essay gives an comprehensive overview of your journey to realise your initial idea, giving detailed and technical explanations where required, in describing the need for occasional compromises and changes. Although at times your writing is in danger to become a diary, it supports to portray your technical knowledge and familiarity with Max. This submitted patch is fully functional and of considerable success. Throughout the patch is well-designed and all aspects in the audio modification have been considered. The overall use however is slightly awkward, because despite your ability for advanced max programming, you have not dealt with some more basic control structures, which would enhance the value of your project immensely. I.e. chosen elements to route the incoming signal, a present system (in particular the pattr system should be of real interest to you). This omission, to keep an end user in mind, had an influence on your presentation as well, which, although generally very coherent and clear, would have needed a better introduction and overview of your patch and processes implemented, to embed the well-informed detailed descriptions in an even more comprehensive manner.

PAGE 1

PAGE 2

PAGE 3

PAGE 4

1. A possible further developmen might have been to incorporate latency compensation for the dry signal, a technique also found in the convolution examples.

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9