Mathematical Synthesis and Analysis of Sound

Ke Shang and Marcus H. Pendergrass
Department of Mathematics and Computer Science, Hampden-Sydney College, Hampden-Sydney, VA 23943

We present some mathematical and computational means for producing and analyzing novel sound synthesis techniques in this paper. We first develop a new signal analyzing method Instantaneous Frequency Spectrum which decomposes a signal in a totally different way and sometimes it does better job than the Fourier transform for short and fast changing signals. We then give and prove sufficient conditions of the phase modulated network, which is an important approach for producing sounds. At last, we contribute a few examples and ideas for constructing interesting sounds.

Mathematical Synthesis and Analysis of Sound

Ke Shang,
Supervisor: Dr. Pendergrass

March 19, 2012

Abstract

We present some mathematical and computational means for producing and analysing novel sound synthesis techniques in this paper. We first develop a new signal analysing method Instantaneous Frequency Spectrum which decomposes a signal in a totally different way and sometimes it does better job than the Fourier transform for short and fast changing signals. We then give and prove sufficient conditions of the phase modulated network, which is an important approach for producing sounds. At last, we contribute a few examples and ideas for constructing interesting sounds.

1 Introduction

The purpose of this research is to develop and analyze novel sound synthesis techniques by mathematical and computational means. Today, sound synthesis has been applied to many different areas, but most of those techniques are based on either digital sampling technologies, or digital emulation of analog synthesizer technologies, for instance, electronic keyboard, software GarageBand for Mac, iPod and iPad, etc. These techniques certainly provide a rich sound palette, but the increasing computational power of personal computers opens up new possibilities for using more sophisticated mathematical techniques to generate new and interesting sounds at home.

In this paper, we go over some classic methods for signal analysing, like Fourier transform and Hilbert Transform, and invent a new method called instantaneous frequency spectrum for some short and fast changing signals. It involves some concepts and properties in probability theory and statistics. We also study how the phase modulation effects the signal waveforms, and we find sufficient conditions for a unique solution to the phase modulation equations in the special case of a cycle. We eventually provide a few ideas for constructing interesting sounds and present some examples to reach the goal of this project: applying mathematical techniques to study and produce novel sounds.

2 Mathematical Approach

Since the research is to synthesise and analyse sound mathematically, we basically apply following mathematical approaches: Fourier transform, cumulative distribution function (CDF) & probability density function (PDF), Hilbert transform, phase modulation, and dynamical systems. We will make a brief review for each of them next.

Approach 1. Fourier Transform

The Fourier transform is a widely used mathematical tool that decomposes a signal into its constituent frequencies. It is called the time domain representation of the signal if the original signal depends on time, while the Fourier transform depends on frequency and so is called the frequency domain representation of the signal. Let $s(t)$ be a function of time and $S(f)$ be the The Fourier transform of $s(t)$, then

$$S(f) = \int_{-\infty}^{\infty} s(t)e^{-2\pi i f t}\, dt,$$

where i is the imaginary number and $e^{-2\pi i f t} = \cos(2\pi f t) - i\sin(2\pi f t)$. In the result section, we will compute the Frourier transform by given a signal $s(t)$ very often.

Approach 2. CDF & PDF

In probability theory and statistics, the cumulative distribution function (CDF) describes the probability that a real-valued variate X with a given probability distribution occuring at a value less than or equal to x; It is defined as

$$F(x) = P(X \leq x),$$

where the $P(X \leq x)$ is the probability that the variate X takes on a value less than or equal to x.

The CDF can be also defined in terms of its probability density function (PDF) $f(u)$, as

$$F(x) = \int_{-\infty}^{x} f(u)\, du,$$

if f is continuous at x. The PDF describes the relative possibility for that variate X to occur at a given point (less than or equal to x). The PDF and CDF give a complete description of the probability distribution of a random variable. We will use CDF and PDF to invent a new signal analysis method in the result section.

Approach 3. Hilbert Transform

The Hilbert transform is a basic tool in Fourier analysis. The difference between the Hilbert transform and Fourier transform is that the Fourier transform is a frequency domain representation of the signal, whereas the Hilbert transform is a time domain representation of the frequency. It is defined as

$$h(t) = \int_{-\infty}^{\infty} H(f)e^{2\pi i f t}\,\partial f,$$

where $H(f) = -i\mathrm{sgn}(f)S(f)$, $S(f)$ is the Fourier transform and $\mathrm{sgn}(f)$ is defined by

$$\mathrm{sgn}(f) = \begin{cases} -1 & f < 0 \\ 0 & f = 0 \\ 1 & f > 0. \end{cases}$$

We will use the Hilbert transform to help study and analyse sounds. During the testing synthesis and analysis, we sometimes use both Hilbert transform and Fourier transform.

Approach 4. Phase Modulation

Phase modulation is the method we use for generating sounds. We define our signal by

$$s(t) = \cos(2\pi f_0 t + g(t)),$$

where f_0 is the initial frequency and $g(t)$ is a function of time and it helps produce various waveforms. We also design a type of structure called Phase Modulated Network in the result section. The simple case of the network is called Phase modulated Cycle. They connect different notes together and all notes invlove with others by some rules so that the network outputs interesting sounds.

Approach 5. Dynamical Systems

For dynamical systems, we mainly study fixed point and contraction mapping theorem and use them to solve some problems occuring in constrcuting phase modulated network. We will prove some conditions for phase modulated network in the result section by referring fixed point and contraction mapping theorem.

3 Synthesis and Analysis of Sound

Having all those approaches, we execute our ideas and tests with Mathematica, Matlab, and Java. Mathematica is known for symbolic operation. Most of our equations are derived from it. We also plot some graphs with Mathematica because it presents generality very well. Matlab is strong at numerical computing, so we run Matlab for most of complex Fourier transforms and signal analysis. A lof of numerical plot are also generated by Matlab and it helps us get more details of some interesting sounds. We synthesise all sounds by using Java. Designing phase modulated network, composing music with newly generated sounds, collaborating with Matlab for signal analysis, all of them are operated in Java. We also use some other softwares, such as SketchBook Pro, Pages, to help explain some defintions and demonstrate our designs.

4 Results

There are three subsections in the result part. We will present our new signal analysis method Instantaneous Frequency Spectrum first; then we will prove the conditions for the phase modulated network; finally we will show some interesting sounds that we develop.

4.1 Instantaneous Frequency Spectrum

The Fourier transform is a fundamental and powerful tool for signal analysing, however, it in some cases cannot offer people an easily understandable result. The Fourier analysis is a process that decompose the orginal signal by infinite number of sine and cosine functions. Due to the properties of sine and cosine, the Fourier transform shows the probablities of all frequencies that could restore that signal. Mathematically, those frequencies make sense, but in reality either some of them could not be heard by human ears or they are not necessarily generated by the signal directly. The following three graphs of Frourier transforms of three linear chirps can give a better sense.

Let define the linear chirp first. Linear chirp's frequency is proportional to the time. Define F as a function of time, frequency $F(t) = rt + f_0$, where r is the linear chirp rate, in Hertz per second, t is the time, and f_0 is the initial frequency. Then define the signal of linear chirp by $s(t) = \sin(2\pi g(t))$ where $g(t) = \int F(t)\,dt = \frac{1}{2}rt^2 + f_0 t$.

All three linear chirps start at 100 Hz. The first one is $f_1 = 200t + 100$ Hz and we play it for 2 seconds; the second is $f_2 = 2,000t + 100$ Hz and it runs for 0.2 seconds; the third is $f_3 = 20,000t + 100$ Hz for 0.02 second. Technically they have same frequency range, which is from 100 Hz to 500 Hz. All graphs are all based on $\log_{10}$ scale.

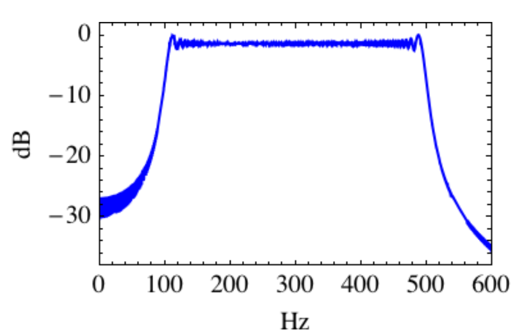

Figure 1. $f_1 = 200t + 100$ Hz for 2 seconds.

The first Fourier transform shows that the frequencies in range 100 Hz to 500 Hz pretty much control the signal. They have very high energy and the frequencies that are under 100 Hz or above 500 Hz can be almost ignored.

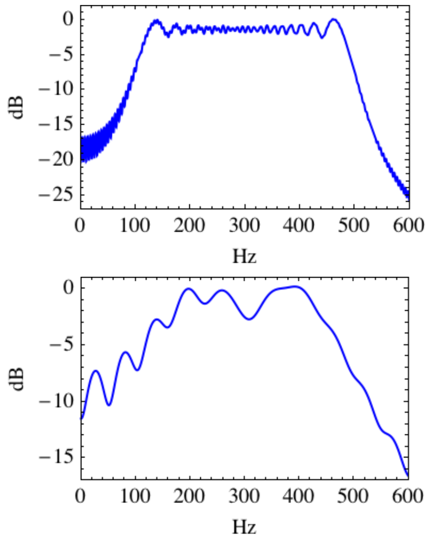

Figure 2. $f_2 = 2,000t + 100$ Hz for 0.2 seconds, $f_3 = 20,000t + 100$ Hz for 0.02 seconds, respectively.

But if we cut down the durtion and increase the rate of the linear chirp, we will observe that more frequencies turn up outside of 100 Hz to 500 Hz. The second and thrid Fourier transform show that the frequency spectrum will become more slippery when the signal runs in a short time or it changes fast.

Although the energies of those very low and high frequencies are relatively small, sometimes people can still get confused while they run Frouier analysis on some instantaneously changing signals. Therefore, we will provide another method for signal analysing: Instantaneous Frequency Spectrum. It effectively solves the former problems and gives a clearer graph about how exactly those frequencies compose the signal. We will derive the general form first and then present three examples (linear chirp, exponential chirp, and sinusoldally modulated chirp).

4.1.1 General Form

Define W as a cumulative distribution function (CDF) of signal energy with respect to the instantaneous frequency f, $W(f) =$ the fraction of signal energy occuring at or below the instantaneous frequency f. Also define F as a function of time t, the instantaneous frequency $f = F(t)$. Assume for simplicity that the instantaneous frequency is strictly increasing with time. So $W(f)$ represents the distribution of the signal energy occurring when the instantaneous frequency $F(t)$ is less than or rqual to frequency f

$$W(f) = \frac{\text{Energy}_{F(t)\le f}}{\text{Total Energy}} \tag{1}$$

If let $s(t)$ represent the signal, where t is the time, then the total energy is given by $c\int_0^T s^2(t)\,dt$, where c is a constant that depends on the units of measurement. Without loss of generality, we assume the units have been chosen so that $c = 1$. T is the duration of the signal. Since T is a constant, the total energy is also a constant. So we can simply mark the total energy as $\|s\|^2$. Similarly, the $\text{Energy}_{F(t)\le f}$ can be written as $\int_0^{t_{F(t)\le f}} s^2(t)\,dt$, where $t_{F(t)\le f}$ is the moment when the instantaneous frequency $F(t) = f$. As figure

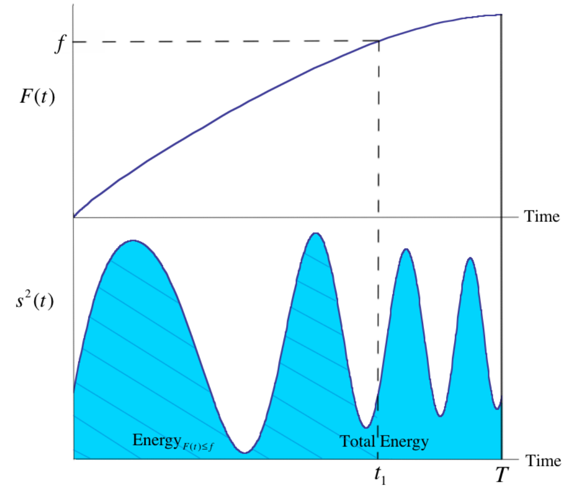

Figure 3. The blue area is the total energy and the blue stripe area is the $\text{Energy}_{F(t)\le f}$. Graph is generated by *Mathematica* and edited by *SketchBookPro* and *Keynote*.

Here the t is a variable and it depends instantaneous frequency f, so we rewrite t as the inverse function of $F(t)$, which is

$$t_{F(t)=f} = F^{-1}_{F(t)=f}(f).$$

Thus, equation(1) becomes

$$W(f) = \frac{\int_0^{t_{F(t)=f}} s^2(t)\,dt}{\int_0^T s^2(t)\,dt} = \frac{\int_0^{F^{-1}_{F(t)=f}(f)} s^2(t)\,dt}{\|s\|^2}. \tag{2}$$

We then need to define P as a probability density function (PDF) of the instantaneous frequency f, $P(f) =$ the density of singal energy with respect to the instantaneous frequency f. According to the defintions in probability theory and statistics, PDF is the derivative of CDF if PDF is continuous at f. Therefore,

$$P(f) = \frac{\partial}{\partial f} W(f) = \frac{1}{\|s\|^2}\frac{\partial}{\partial f}\int_0^{t_{f=f_c}} s^2(t)\,dt.$$

Simplifies it, we get

$$P(f) = \frac{s^2(F^{-1}_{F(t)=f}(f))\frac{\partial}{\partial f}F^{-1}_{F(t)=f}(f)}{\|s\|^2}$$

or in another form (3)

$$P(f) = \frac{s^2(t)}{\|s\|^2|F'(t)|}, \text{where } t = F^{-1}_{F(t)=f}(f)$$

Equation(3) is the instantaneous frequency spectrum method for the simple case when the instantaneous frequency is strictly increasing with time. To expand it into a general form for any instantaneous frequencies, we apply the same logic, as figure

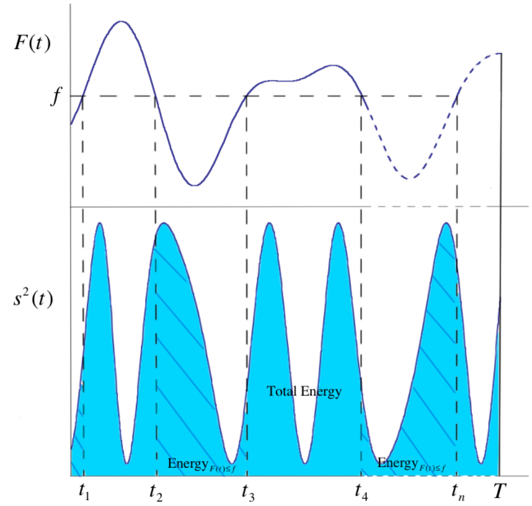

Figure 4. Figure. The blue area is the total energy and the blue stripe area is the Energy$_{F(t)\leq f}$. Graph is generated by *Mathematica* and edited by *SketchBookPro* and *Keynote*.

Now the instantaneous frequency $F(t)$ can be any functions of time, but the core idea is that we still count the signal energy only when it is at or below the instantaneous frequency f. Therefore, we can get a new form of CDF

$$
\begin{aligned}
W(f) &= \frac{\int_0^T s^2(t) I(F(t_i) \leq f)\, dt}{\int_0^T s^2(t)\, dt} \\
&= \frac{1}{\|s\|^2} \int_0^T s^2(t) I(F(t_i) \leq f)\, dt,
\end{aligned}
\tag{4}
$$

where I is function defined by

$$
I(F(t_i) \leq f) = \begin{cases} 1 & \text{if } F(t_i) \leq f \\ 0 & \text{otherwise.} \end{cases} \tag{5}
$$

We then take the derivative of the new CDF and get the general formula of instantaneous frequency spectrum

$$
P(f) = \frac{1}{\|s\|^2} \sum_{t:F(t)=f} \frac{s^2(t)}{|F'(t)|} \tag{6}
$$

Finally, Equation(6) is the our new tool for signal analysing. From the formula itself, we can see that the new method is completely based on the instantaneous frequency. We will apply the instantaneous frequency spectrum method to analyze three examples next.

4.1.2 Three Examples

Example 1. Linear Chirp

Based on same definition, the linear chirp on time interval $[0, T]$ takes the form

$$
s(t) = \begin{cases} \sin(2\pi f_0 t + \pi r^2 t) & \text{if } t \in [0, T] \\ 0 & \text{otherwise.} \end{cases}
$$

Since the linear chirp is strictly increasing with time, we can apply equation(3). Here the $t_{F(t)=f} = F^{-1}_{F(t)=f}(f) = \frac{f-f_0}{r}$. Input $F^{-1}_{F(t)=f}(f) = \frac{f-f_0}{r}$ into $s(t)$ and then input $s(t)$ into equation(3). We get the instantaneous frequency spectrum of linear chirp on $[f_0, f_0 + rT]$

$$
\begin{aligned}
P(f) &= \frac{\sin^2[r\pi(\frac{f-f_0}{r})^2 + 2\pi f_0(\frac{f-f_0}{r})]\frac{\partial}{\partial f}(\frac{f-f_0}{r})}{\|s\|^2} \\
&= \frac{\sin^2[\frac{\pi}{r}(f^2 - f_0^2)]}{r\|s\|^2}
\end{aligned}
$$

and the instantaneous frequency spectrum of linear chirp for all f is

$$
P(f) = \begin{cases} \frac{\sin^2[\frac{\pi}{r}(f^2 - f_0^2)]}{r\|s\|^2} & \text{if } f \in [f_0, f_0 + rT] \\ 0 & \text{otherwise.} \end{cases} \tag{7}
$$

Test equation(7) by the second linear chirp that we mentioned at the beginning of the 4.1 section, $f_2 = 2,000t + 100$ Hz for 0.2 second, which runs in a short time and the frequency changes very fast. We get the following graph

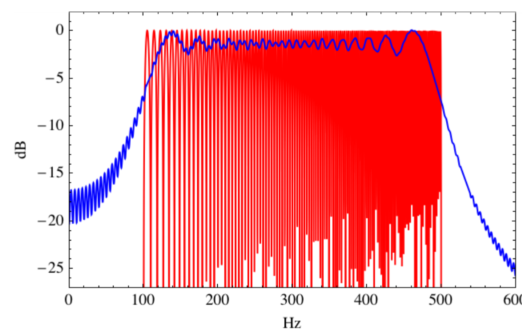

Figure 5. The blue line is the Fourier transform and the red line is the instantaneous frequency spectrum. The range is based on Log scale. Simpling rate is 44,100 Hz. Graph is generated by *Mathematica*.

Through the graph, we can clearly see that there is no frequency occur outside of 100Hz to 500 Hz and it fits what people hear and understand very well. We can also observe how exactly those frequencies distribute into the signal. To show that point, we run both Fourier analysis and instantaneous frequency analysis again to test the third linear chirp $f_3 = 20,000t + 100$ hz for 0.02 second, whose duration is shorter and changing rate is extremely high. Here is what we get

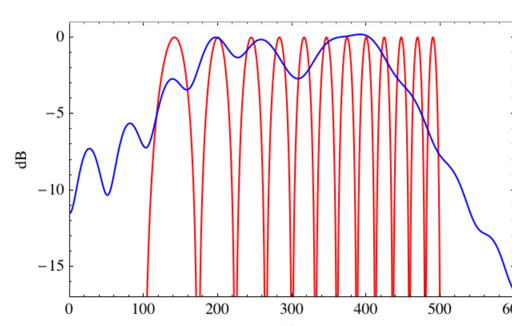

Figure 6. The blue line is the Fourier transform and the red line is the instantaneous frequency spectrum. The range is based on Log scale. Simpling rate is 44,100 Hz. Graph is generated by *Mathematica*.

From this graph, we cannot really tell what the frequency distribution is from the Fourier analysis, but from the instantaneous frequency spectrum it is even more conspicuous. For instance, we see that the energies of some frequencies around 170 $\sim$ 180 Hz, 220 $\sim$ 230 Hz, 260 $\sim$ 270 Hz, 300 Hz, and so on, drop off to zero. That is exactly what happens when we play a sine waveform with respect to a strictly increasing frequency. The moment when the frequency increases to some values, unfortunately the sine function just reach the zero point. So people will hear nothing at that moment and certainly the frequency there has zero energy. Although we can also see the floating of energy distributed on each frequency from the fourier transform graph, it is difficult to read more details. Consequently, the linear chirp test proves that the instantaneous frequency spectrum is good for the short and severely changing signal.

Example 2. Exponential Chirp

Apply the instantaneous frequency spectrum to the exponentical chirp. Let $F(t) = 2^{rt} f_0$, where r is the chirp rate, in octave per second, t is the time, and f_0 is the initial frequency. The signal of the exponentical chirp is defined by $s(t) = \sin(2\pi g(t))$ where $g(t) = \int F(t)\,dt = \frac{f_0}{r \ln 2} 2^{rt}$. The linear chirp on time interval $[0, T]$ takes the form

$$s(t) = \begin{cases} \sin(\frac{\pi f_0}{r \ln 2} 2^{1+rt}) & \text{if } t \in [0, rT] \\ 0 & \text{otherwise.} \end{cases}$$

Same as the linear chirp, we can also use equation(3) for the exponential chirp because it is is rigorously increasing with time. Here we will apply the second form. On the interval $t \in [0, rT]$,

$$t_{F(t)=f} = F^{-1}_{F(t)=f}(f) = \frac{\ln f - \ln f_0}{r \ln 2}$$

and

$$|F'(t)| = |\frac{\partial}{\partial t}(2^{rt} f_0)| = r f_0 2^{rt} \ln 2 = r f \ln 2.$$

Input $t_{F(t)=f}$, $|F'(t)|$ and $s(t)$ to equation(3). We get the instantaneous frequency spectrum of the exponential chirp for all f

$$P(f) = \begin{cases} \frac{\sin^2(\frac{2 f \pi}{r \ln 2})}{r f \ln 2 \|s\|^2} & \text{if } f \in [f_0, 2^{rT} f_0] \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

Now we test equation(8) by two exponential chirps. One is $F(t) = 2^{10t} \times 100$ Hz, whose chirp rate is 10 octaves per second and initial frequency is 100 Hz, and we play it for 0.2 second;

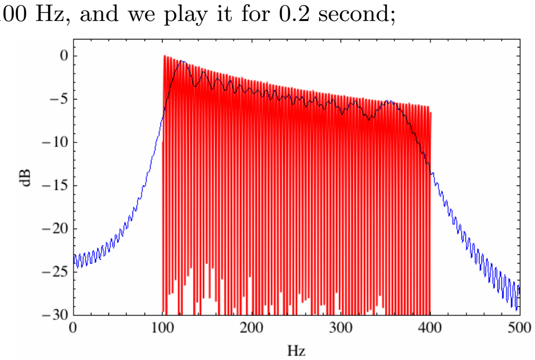

Figure 7. The blue line is the Fourier transform and the red line is the instantaneous frequency spectrum. Simpling rate is 44,100 Hz. The instantaneous frequency spectrum graph is generated by *Mathematica* and the Fourier transform is generated by *Matlab*.

the other is $F(t) = 2^{100t} \times 100$ Hz, whose chirp rate is 100 octaves per second and initial frequency is 100 Hz, and we play it for 0.02 second.

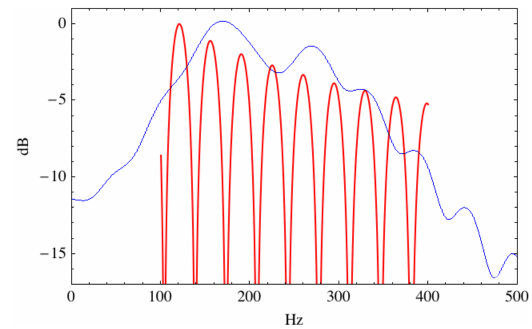

Figure 8. The blue line is the Fourier transform and the red line is the instantaneous frequency spectrum. Simpling rate is 44,100 Hz. The instantaneous frequency spectrum graph is generated by *Mathematica* and the Fourier transform is generated by *Matlab*.

In the first graphs, the fourier transform and the instantaneous frequency spectrum match pretty well. There are some differences on two edges, but in general we can see that the energy of frequency is decreasing with time because the frequency is increasing faster and faster. The second graph shows us a different story. We cannot tell at which frequencies the chirp starts and ends from the fourier transform, but the instantaneous frequency spectrum is still very clear. Therefore, in the competition of exponential chirp, the instantaneous frequency spectrum wins.

Example 3. Sinusoidally Modulated Chirp

The sinusoidally modulated chirp (sounds like a siren sound) is the most interesting and demonstrable example among all three. The frequency of the chirp is defined by

$$F(t) = f_0 + \alpha f_m \cos(2\pi f_m t),$$

where f_0 is the initial frequency, t is the time, α is the modulated weight (is the name right?), and f_m is the modulated frequency. The signal is defined by

$$s(t) = \sin(2\pi g(t)) = \sin(2\pi f_0 t + \alpha \sin(2\pi f_m t)),$$

where $g(t) = \int F(t)\,dt = f_0 t + \frac{\alpha}{2\pi} \sin(2\pi f_m t)$.

The frequency function tells us that the chirp runs at a frequency that is based on a constant value floating with a cosine function of time. Since the range of a cosine function is finite for all t, the range of the siren frequency is also finite no matter how long the signal plays. Therefore, we can observe a significant signal decomposing difference between the fourier transform and instantaneous frequency spectrum. Because the mathematics here is a little bit complicated, we could not derive the symbolic formulas for both frourier transform and instantaneous frequency spectrum. So we will show the numerical result computed by *Matlab*.

Let $f_0 = 100$ Hz, $\alpha = 2$, $f_m = 3$ Hz, and the duration $T = 10$ seconds. The result is

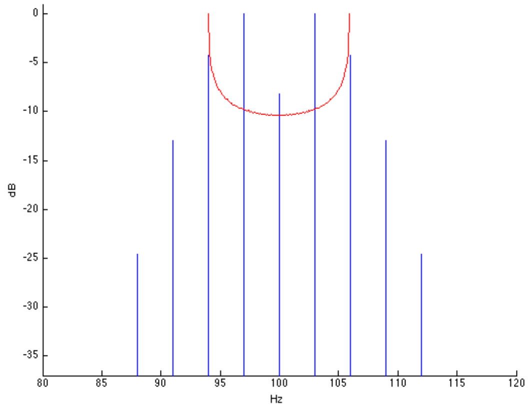

Figure 9. The blue line is the Fourier transform and the red line is the instantaneous frequency spectrum. Simpling rate is 44,100 Hz. Graph is generated by *Matlab*.

The graph displays two totally different views of decomsposing signal. The instantaneous frequency spectrum gives a continuous line from the moment when the chirp starts to run to the time when it stops to describe the distribution and changing of energy on each frequency, but the fourier transform shows all possible frequencies that it needs to restore the original signal. Because $F(t) = f_0 + \alpha f_m \cos(2\pi f_m t)$ and $\alpha = 2$, $f_m = 3$, the frequency range of this chirp should be between 100 Hz $- 2 \times 3$ Hz $= 94$ Hz to 100 Hz$+ 2 \times 3$ Hz $= 106$ Hz. That is exactly what the instantaneous frequency spectrum represents. The fourier transform stands at another side this time. Mathematically, we do can use all frequencies from the fourier transform to reproduce the signal; however, not all of them can be heard or are useful for our purpose. If we extends the interval and range of our graph, we will see that there exist a lot of terribly low and high frequencies at very low energy level, which are not what we need.

As a result after three examples, the instantaneous frequency spectrum gives us a new glasses when we look inside a signal. Normally when we deal with some very complex signals, the Frouier transform is more useful than the instantaneous frequency spectrum because the latter is not good at decomposing signals with multiple changing frequencies, but for a short and simplex signal, the instantaneous frequency spectrum can sometimes do better job.

4.2 Phase Modulated Networks

We will give and prove the conditions for the phase modulated network in this section. We will start at the simple case phase modulated cycle and then step on to the general case phase modulated network.

4.2.1 Phase Modulated Cycles

A phase modulated cycle is a phase modulated network in which the nodes are connected to form a simple cycle. As figure 10

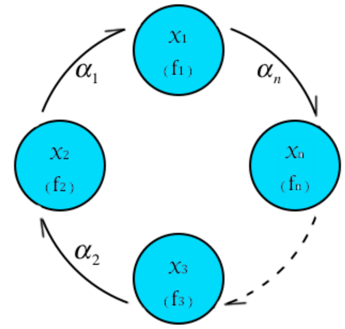

Figure 10. A n-cycle's network.

The network equations are given by

$$
\begin{aligned}
x_1 &= \cos(2\pi f_1 t + \alpha_1 x_2) \\
x_2 &= \cos(2\pi f_2 t + \alpha_2 x_3) \\
x_3 &= \cos(2\pi f_3 t + \alpha_3 x_4) \\
&\vdots \\
x_n &= \cos(2\pi f_n t + \alpha_n x_1) \qquad (9)
\end{aligned}
$$

where f_1, f_2, f_3, ..., f_n and α_1, α_2, α_3, ..., α_n are the frequencies and weights of each note, respectively. Provided these equations have a solution, they define x_1, x_2, x_3, ..., x_n as functions of time.

Phase modulated cycles can generate various sounds based on any frequencies; however, there are some constraints on the connction weights. During the numerical test, some combinations of weights gave a very chaotic kind of sound and frequency spectrum. For example, give a 2-cycle network and let $f_1 = f_2 = 1$ Hz, $\alpha_1 = \alpha_2 = 1.1$, Sample rate fs =44,100 Hz. Playing the sound for 4 seconds, the result shows

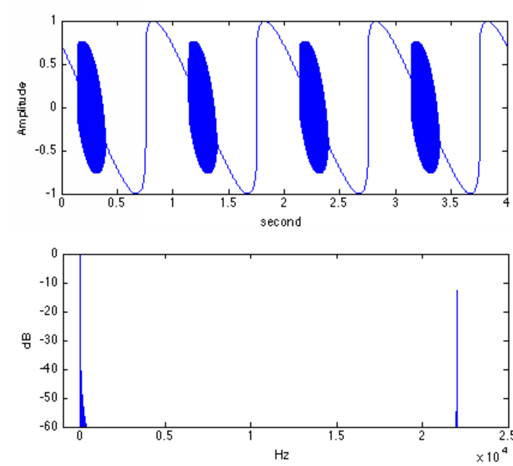

Figure 11. Chaotic 2-cycle sound spectrum frequency spectrum.

The sound spectrum reveals that this 2-cycle network generates extremely high frequencies periodically and the frequency spectrum gives the value of that strange high frequencies, 22050 Hz, which is equal to half of the sample rate. But according to Nyquist-Shannon Sampling Theorem (need citation), 22,050 Hz is the highest frequency that the sampling can get, but not the exact real frequency that the 2-cycle network provides. The graphs simply show that when the weights are beyond some boundaries, the phase modulated cycle will result in chaos.

Therefore, our main result in this section gives sufficient conditions for the system of phase modulated cycle to have a unique solution.

Theorem 1. *In the phase modulated cycles, if the product of the weights is less than 1 in modulus, then the equations* (9) *has a unique solution for all* t*.*

The proof of Throrem 1 depends on the following two propositions, which concern *fixed points*. A point x is called a fixed point of a function f if $f(x) = x$. The following propositions give sufficient conditions in order for a function $f : \mathbb{R} \to \mathbb{R}$ to have a unique fixed point.

Proposition 1. *Let* $I = [a, b]$ *be an interval and let* $f : I \to I$ *be continuous. Then* f *has at least one fixed point in* I*.*

Proposition 2. *Let* $f : I \to I$ *and assume that* $|f'(x)| < 1$ *for all* x *in* I*. Then there exists a unique fixed point in* I*.*

For proofs of proposition 1 and 2, see [2], pages 13 and 14.

Proof. First consider a 1-cycle network. The equation for the waveform produced by the vertex is

$$x = \cos(2\pi f_0 t + \alpha x) \tag{10}$$

where f_0 is a constant frequency and α is the weight of loop at the vertex. We need to show that this equation has a solution for all $t \in \mathbb{R}$, and that the solution is unique. Define a function F by

$$F(x) = \cos(2\pi f_0 t + \alpha x).$$

Any solution of equation (10) satisfies $x = F(x)$. This means that x is a fixed point of F. We need to prove that F has a unique fixed point for each $t \in \mathbb{R}$.

For $x = F(x) = \cos(2\pi f_0 t + \alpha x)$, if $x \in [-1, 1]$, then $F(x) = \cos(2\pi f_0 t + \alpha x) \in [-1, 1]$, so F maps [-1,1] into itself; also F is already continuous. Therefore, by Proposition (1), F has at least one fixed point. To have a unique fiexed point of F by Proposition (2), the derivative of F in absolute value has to be less than one. The derivative of F in absolute value is

$$|F'(x)| = |-\alpha \sin(2\pi f_0 t + \alpha x)| = |\alpha||\sin(2\pi f_0 t + \alpha x)|.$$

Since the maximum value of $|\sin(2\pi f_0 t + \alpha x)|$ is 1 for all t, when $|F'(x)| = |\alpha||\sin(2\pi f_0 t + \alpha x)| < 1$, $|\alpha|$ has to be less than 1. Therefore, if $|\alpha| < 1$, F has a unique fixed point.

The equations for the 2-cycle waveform are produced by

$$\begin{aligned} x_1 &= \cos(2\pi f_1 t + \alpha_1 x_2) \\ x_2 &= \cos(2\pi f_2 t + \alpha_2 x_1). \end{aligned} \tag{11}$$

To solve this system of equations (11), note that it can be written as

$$x_1 = \cos(2\pi f_1 t + \alpha_1 \cos(2\pi f_2 t + \alpha_2 x_1)) \tag{12}$$

Define a function F_1 that $F_1(x) = \cos(2\pi f_1 t + \alpha_1 x)$ and a function F_2 that $F_2(x) = \cos(2\pi f_2 t + \alpha_2 x)$. Then any solutions of equations (11) satisfies $x_1 = F_1(F_2(x_1))$, so x_1 is a fixed point of $F_1(F_2(x_1))$. We want that $F_1(F_2)$ has a unique fixed point for all $t \in \mathbb{R}$ as well. Because $F_1(F_2)$ is cosine function, obviously, $F_1(F_2)$ is continuous and if $x \in [-1, 1]$, $F_1(F_2)$ maps [-1,1] into itself. By Proposition (1), therefore, $F_1(F_2)$ has at least one fixed point. The absolute value of the derivative of $F_1 \circ F_2$ is

$$\begin{aligned} &|\frac{\partial}{\partial x_1}[F_1(F_2(x_1))]| = \\ &\quad |\alpha_1\alpha_2||\sin(2\pi f_2 t + \alpha_2 x_1)\sin(2\pi f_1 t + \alpha_1 F_2(x_1))| \end{aligned}$$

Since the absolute value of sine function is always less or equal to one, Proposition (2) implicts that the fixed point of $F_1 \circ F_2$ is unique wherever $|\alpha_1\alpha_2|$ is less than 1.

Finally, apply the same idea to the n-cycle case. The n-cycle waveform is generated by equation (9). Solve it by rewriting x_1 by

$$\begin{aligned} x_1 = \cos(2\pi f_1 t + \alpha_1 &\cos(2\pi f_2 t \\ &+ \alpha_2 \cos(2\pi f_3 t \\ &+ \alpha_3(\cdots F_n(x_1))))) \end{aligned} \tag{13}$$

Similar to the 2-cycle's proof, define functions $F_1(x) = \cos(2\pi f_1 t + \alpha_1 x)$, $F_2(x) = \cos(2\pi f_2 t + \alpha_2 x)$, ..., and $F_n(x) = \cos(2\pi f_n t + \alpha_n x)$. Then any solutions of the system of equations (13) satisfies $x_1 = F_1(F_2(...(F_n(x_1))))$. Therefore x_1 is a fixed point of $F_1(F_2(...(F_n)))$. Since $F_1(F_2(...(F_n)))$ is still a cosine function, it is continuous and maps [-1,1] into itself. By Proposition (1), $F_1(F_2(...(F_n)))$ has at least one fixed point.

The derivative of $F_1(F_2(...(F_n)))$ in modulus is the product of all weights in modulus times the product of n times of sine functions in modulus, like

$$|\frac{\partial}{\partial x_1}[F_1(F_2(...(F_n)))]| = |\alpha_1\alpha_2\alpha_3\cdots\alpha_n||O(x_1)|$$

where $|O(x_1)| = |\sin(2\pi f_n t + \alpha_n x_1)\sin(2\pi f_{n-1} t + \alpha_{n-1}\cos(2\pi f_{n-2} t + \alpha_{n-2} x_1))\sin(\cdots)\cdots\sin(\cdots)|$. There is no doubt that the absolute value of the product of those sine functions cannot be greater than one. By Proposition (2), therefore, for n-cycle phase modulated cycles, if the product of all weights in modulus $|\alpha_1\alpha_2\alpha_3\cdots\alpha_n| < 1$, then the waveform equations have a unique fixed point, which means the phase modulated cycle network has a unique solution for all $t \in \mathbb{R}$. $\square$

As a comparison of the chaotic graphs, we change that $\alpha_1 = \alpha_2 = 1.1$ to $\alpha_1 = \alpha_2 = 0.99$ to test our theorem numerically. The result gives

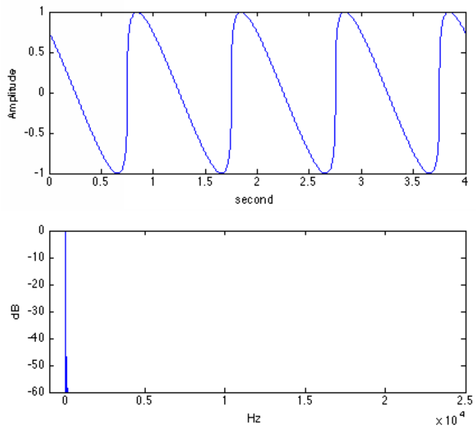

Figure 12. Figure. Normal 2-cycle sound spectrum frequency spectrum.

which is very reasonable.

4.2.2 Phase Modulated Networks

For the general n-note phase modelated network, as figure 13

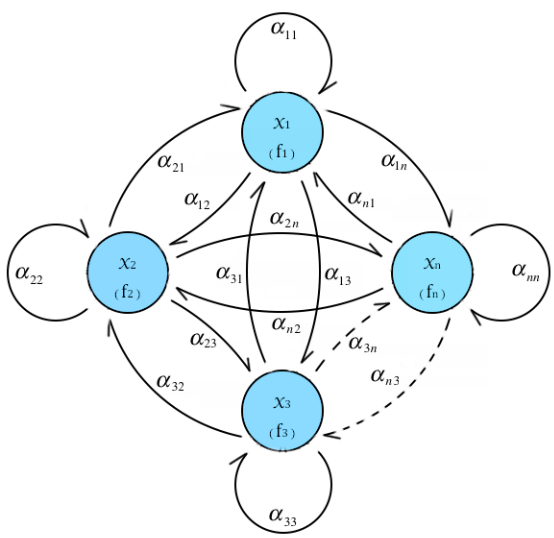

Figure 13. PMNetwork.

each note connects itself and all others. The network equations are given by

$$
\begin{aligned}
x_1 &= \cos(2\pi f_1 t + \alpha_{11}x_1 + \alpha_{21}x_2 + \cdots + \alpha_{n1}x_n)\\
x_2 &= \cos(2\pi f_2 t + \alpha_{12}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{n2}x_n)\\
x_3 &= \cos(2\pi f_3 t + \alpha_{13}x_1 + \alpha_{23}x_2 + \cdots + \alpha_{n3}x_n)\\
&\vdots\\
x_n &= \cos(2\pi f_n t + \alpha_{1n}x_1 + \alpha_{2n}x_2 + \cdots + \alpha_{nn}x_n).
\end{aligned} \tag{14}
$$

Same as the phase modulated cycle, we need some sufficient conditions to make sure that the system of equations have a unique solution so that the network will not show chaos. The approach is still related to the fixed point. Since the weights of the general n-note phase modulate network is a $n \times n$ matrix,

$$
\begin{bmatrix}
\alpha_{11} & \alpha_{21} & \dots & \alpha_{n1}\\
\alpha_{12} & \alpha_{22} & \dots & \alpha_{n2}\\
\alpha_{13} & \alpha_{23} & \dots & \alpha_{n3}\\
\vdots & \vdots & \ddots & \vdots\\
\alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn}
\end{bmatrix}
$$

we will apply the Banach fixed point theorem (contraction mapping theorem).

Theorem 2. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuous and $x_1, x_2 \in \mathbb{R}^n$. If*

$$\|F(x_1) - F(x_2)\| \leq K\|x_1 - x_2\|$$

where $0 \leq K < 1$, then F has a unique fixed point.

Define a function F by

$$
F\begin{pmatrix} x_1\\ x_2\\ x_3\\ \vdots\\ x_n \end{pmatrix} =
\begin{bmatrix}
\cos(2\pi f_1 t + \alpha_{11}x_1 + \alpha_{21}x_2 + \cdots + \alpha_{n1}x_n)\\
\cos(2\pi f_2 t + \alpha_{12}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{n2}x_n)\\
\cos(2\pi f_3 t + \alpha_{13}x_1 + \alpha_{23}x_2 + \cdots + \alpha_{n3}x_n)\\
\vdots\\
\cos(2\pi f_n t + \alpha_{1n}x_1 + \alpha_{2n}x_2 + \cdots + \alpha_{nn}x_n)
\end{bmatrix} \tag{15}
$$

Banach fixed point theorem guarantees the existence and uniqueness of fixed point if we can find the conditions on F so that there exists a $K \in [0, 1)$ for all $x \in \mathbb{R}^n$.

To discover the conditions, we simplify F first. We know the Jacobian matrix of a function describes the orientation of a tangent plane to the function at a given point. If p is a point in $\mathbb{R}^n$ and F is differentiable at p, then derivative of F is given by $J_F(p)$. In this way, the linear map described by $J_F(p)$ is the best linear approximation of F near the point p, which means

$$F(x) \approx F(p) + J_F(p)(x - p).$$

Therefore, if $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and $x_1, x_2, p \in \mathbb{R}^n$, we can write

$$
\begin{aligned}
&F(x_1) - F(x_2) \approx\\
&\quad J_F(p)(x_1 - p) + F(p) - J_F(p)(x_2 - p) + F(p)\\
&\quad = J_F(p)(x_1 - p) - J_F(p)(x_2 - p)\\
&\quad = J_F(p)(x_1 - x_2)
\end{aligned} \tag{16}
$$

Based on Banach fixed point theorem

$$\|F(x_1) - F(x_2)\| \approx \|J_F(p)(x_1 - x_2)\| \leq K\|x_1 - x_2\|,$$

we have a conjecture.

Conjecture 1. *If all singular values of the Jacobian matrix of the n-note phase modulate network are less than 1, $\sigma_{J_F} \in [0, 1)$, for all $x_1, x_2, x_3, \cdots, x_n \in [-1.1]$, then the system of the n-note phase modulated network will have a unique solution for all $t \in \mathbb{R}$.*

Since the Jacobian matrix of $n \times n$ system is huge, its symbolic eigenvalue is difficult to compute. During the research period we had only a conjecture, but no proof. However, during the paper writting, Dr. Brain Lins found a theorem in a math paper published in 1978 that could prove our conjecture is correct.

Theorem 3. *Let (X, d) be a complete metric space and $g : (X, d) \to (X, d)$ a local radical contraction. Suppose for some $x_0 \in X$ the points x_0 and $g(x_0)$ are joined by a path of finite length. Then g has a unique fixed point in X.*

For the proof of theorem 3, see [3].

The proof of our conjecture requests some background in real analysis and matrix analysis. The basic idea is to show that if all the singular values of the Jacobian matrix are less than 1 for each $x \in X^n$, where $X = [-1, 1]$, then the function F is a local radical contraction inside X^n. Since F involves with only sine functions, it is obvious that x and $F(x)$ is joined by a path of finite length. Therefore, our conjecture can be proved by the theorem. I will finish the proof in the future.

4.3 Interesting Sound

As we talked about the phase modulated cycle and network, in this section we will apply them to generate some interesting sounds. Different combinations of frequencies can produce various sound effects. We will present three examples. The first one is a simple phase modulated cycle, but it sounds like an Oboe; the second is a phase modulated network which teems with changing tones; the last one is a composite of cycle and network in which results a euphonious major triad.

Example 1. Phase Modulated Cycle

We use 440 Hz as the basic frequency, 800 Hz, 1200 Hz, 1600 Hz, and 2800 Hz as harmonizing frequencies. the connection and weights are as figure

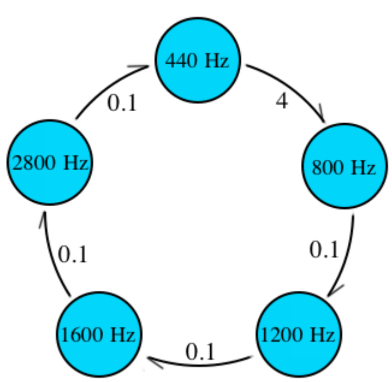

Figure 14. Structure of "Oboe" sound

and intensity of each frequency that we play is 0.7, 0.6, 0.6, 0.25, and 0.05 (maximum intensity is 1), respectively. Because high frequency is just for harmomizing and 440 Hz is what we want to hear, we play 2800 Hz weaklier and 440 Hz firmlier. We run the sound for 6 seconds and plot its waveform versus time and the Fourier transform to see the details of the sound.

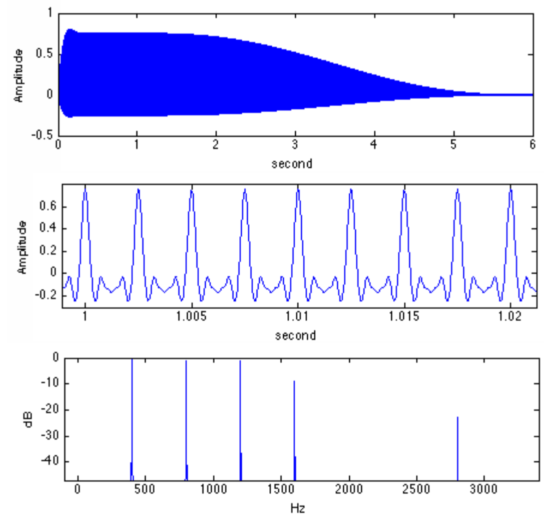

Figure 15. The first graph is the whole time domain waveform. We add an envelope to make it sound naturelier. The second graph is the locally magnified waveform from 1 second to 1.02 second. The third is the frouier transform. Graph is generated by Matlab.

The Fourier transform shows that we can hear clear frequencies 440 Hz, 800 Hz, 1200 Hz, weak 1600 Hz and very weak 2800 Hz. It is not easy to say exactly how each frequency effect others at every single second, but since all those harmomizing frequencies are overtones, the final output has a clear, smooth, and penetrating voice. Because it sounds simliar to an oboe, we name this phase modulated cycle "KeOboe".

Example 2. Phase Modulated Network

In the first example, we know that some lucky overtones can generate interesting sounds. We will play with some "little" frequencies in the phase modulated network example. We still set 440 Hz as our basic frequency, but we add four "little" frequencies around the basic one, as figure

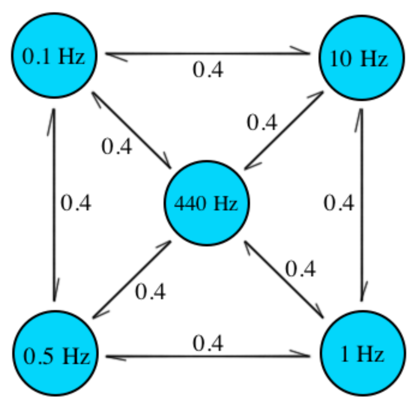

Figure 16. Structure of Phase Modulated Network's sound

The "little" frequencies are 0.1 Hz, 0.5 Hz, 1 Hz, and 10 Hz. All their connection weights are 0.4. We turn on only 440 Hz and turn off all four modulated frequencies because they are too low to be heard by human ears anyway and we want to concentrate on how those low frequencies change the sound. Here are the wavefrom and Fourier transform graphs.

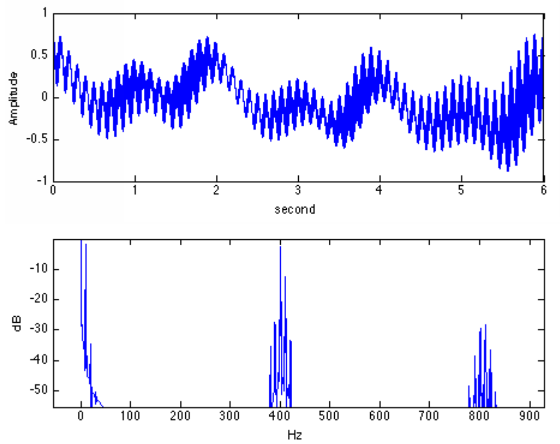

Figure 17. Graph is generated by Matlab.

First of all, the signal gives a very funny looking waveform. We can see that the volume and frequency is changing all the time. We play only 440 Hz, but the Frourier transform shows three different frequency groups. Although the first one is so low that we can not hear, it has very high energy. The energy of the third group is not very high, but we can still hear a little bit high frequency sound. We also find it is interesting even for the 440 Hz group if we zoom in it, as figure

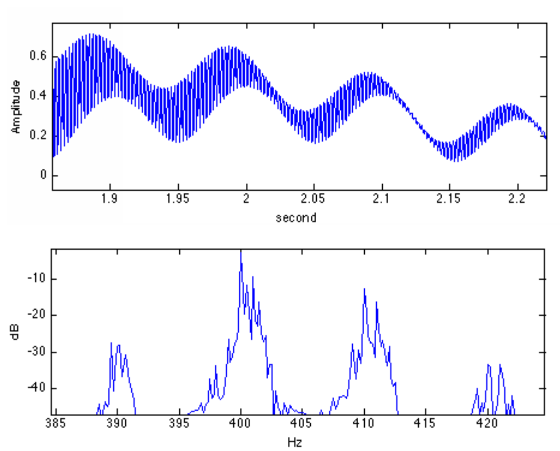

Figure 18. The first graph is the locally magnified waveform from 1.9 second to 2.2 second. The second is the magnified 440 Hz group from frouier transform. Graph is generated by Matlab.

We see that 390 Hz, 400 Hz, 410 Hz, 420 Hz exist in the 440 Hz group, but the 440 Hz itself is not there (too weak). We also zoom in the waveform and it is still changing fancy.

This phase modulated network does not give an agreeable voice, but it does generate a bizarre sound by only four irrelevant low frequencies. Therefore, this example offers us a way to decorate any normal sounds.

Example 3. Composition

The last example will combine ideas from former two: harmony and decoration. We call it "KeMajorTriadC4". As figure

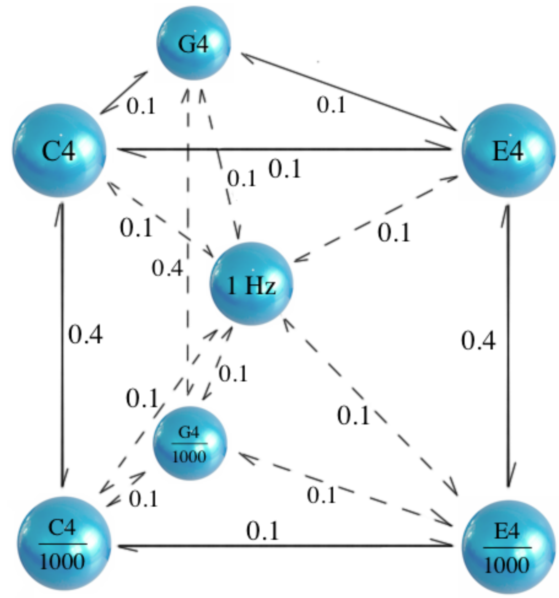

Figure 19. Structure of KeMajorTriadC4

The top is a phase modulated cycle (triangle) with the major chord C4, E4, and G4 (216.63 Hz, 329.63 Hz, and 392 Hz, based on equal-tempered scale); the bottom is same structure as the top, but their frequencies are 1000 times smaller the the top. The bottom plays the role of "little" frequencies. We then connect the bottom and the top to make a solid. Eventually put an 1 Hz note at the center of the solid and connect with every other notes. The waveform and Frourier transform shows

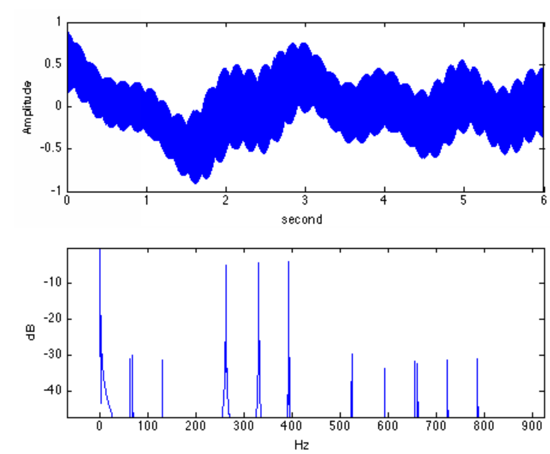

Figure 20. Graph is generated by Matlab.

The waveform shows that it has a fantastically variable sound. The Fourier transform gives three frequency groups and all of them are harmonic. We can also see how different frequencies contribute to the sound: the low frequency group ($0 \sim 1$ Hz) mainly decorates the whole signal and makes it be full of change; the major chord group (C4, E4, and G4) lest the voice be musical; the high frequency group ($500 \sim 800$ Hz, actually they are double higher than the major chord) has very low energy, but it is still important because it adds more colorful elements to the sound.

As a result, applying those constructing methods, we can produce interesting sounds as many as we want. Sometimes, it is pleasurable to get both pretty graphs and dulcet sounds.

5 Discussion and Future Work

Due to the lack of enough mathematical skills and knowledge, I was not able to finish some computing and proofs during the research period. Some work requested a lot of computer skills as well, but I did not have enough ability to make programs to get exactly what I want, so I got lots of helps from Dr. Pendergrass for coding, especially for Java.

Generally, I finished my project as expected. I got three main results and all of them are derived or proved successfully. The contraction proof that I have mentioned at the 4.2 section will be kept working. We will also keep working and thinking some new designs of constructing sounds so that we can get more pretty graphs and sounds.

6 Conclusion

During this ten weeks' project, we utilized the computers and software that we have, and presented some interesting ideas and sounds with our computers. Besides, we provided our new signal analysis method Instantaneous Frequency Spectrum, and we proved the sufficient conditions for the phase modulated network. I got the taste of mathematical research and writing and I also learnt many stuffs that I have never seen. It was a fun but conscientious process. This project is about math and sound, but I also read and researched a lot of things about music, and realize how strong the connections are between math and sound, math and music. As I said, I will keep working and thinking about this research since math and music are two of the most magical objects in the world.

7 Acknowledgments

We thank Dr. B. Lins and Dr. H. Hulsizer for their kindly help.

References

[1] The Spectra of Frequency Modulated Waveforms

[2] An Introduction to Chaotic Dynamical Systems

[3] Local Contractions in Metric Spaces

[4] Digital Signal Processing, Oppenheim and Schafer

[5] The Math Behind The Music, L. Harkleroad, Cambridge University Press, 2006