Analysis of Effects on Sound Using the Discrete Fourier Transform

A Thesis

Presented in Partial Fulfillment of the Requirements for the Degree
Master of Science in the Graduate School of The Ohio State University

By

Timothy Tussing, B.A., B.S.

Graduate Program in Department of Mathematics

The Ohio State University

2012

Master's Examination Committee:

Bart Snapp, Advisor

Herb Clemens

James Cogdell

Abstract

The purpose of this study was to show how mathematics can be used to analyze effects on sound. Our hope is that this may inspire student interest in mathematics.

We analyzed five common industry standard effects. Research data was gathered using *Mathematica* and *GarageBand* software. Three versions of each effect were used to alter pure tone sound waves of ten different frequencies using *GarageBand*. Then using *Mathematica*'s `Fourier` command, the frequency spectrum of each altered sound wave was generated. Through observation of each set of 30 frequency spectra, the most prominent and common pure tone components were determined. For each effect, *Mathematica*'s `Fit` command was used to determine a best fit model of the magnitude of each component as a function of frequency.

Our models provide descriptions of the effects that are consistent with the traditional descriptions of the industry standard effects in our study. If similar research is to be conducted, our recommendation is that more versions of each effect, a wider range of input frequencies, and a higher sampling rate would produce function models that are even more consistent with traditionally accepted effect descriptions. Furthermore, an understanding of the hardware and software design used to build effects on sound is highly recommended.

This thesis is dedicated to my aunt, JoAnne Blackmore, who passed away at the time of this writing, and to Mom and Dad. You are the three kindest, most generous people I have ever known. The love and support of each of you will always be remembered and appreciated!

Acknowledgments

This work would not have been possible without the creative guidance of Bart Snapp. Thank you, Bart! Also, thank you Steve Fields, Steve Hoehner, Charity Turner, Jeff Dunworth, Rodica Costin, Herb Clemens, James Cogdell, my friends and colleagues in the Ohio State math department, and last but not least my brother, Tom Tussing, for your sincere interest and support.

Vita

January 13, 1969	Born - Columbus, Ohio
1995	B.A. Music Technology, The Evergreen State College
2004	B.S. Applied Mathematics, California State University - Chico
2010-present	Graduate Teaching Associate, The Ohio State University.

Fields of Study

Major Field: Department of Mathematics

Table of Contents

List of Figures

Chapter 1: Introduction

Our goal is to explain how to use the discrete Fourier transform to analyze sound waves and effects on sound waves. We will show that the discrete Fourier transform produces frequency spectra of sound waves and we will analyze the frequency spectra of sound waves that are altered by one of five standard effects. Then, using least squares approximation, we will mathematically describe each effect.

1.1 Sound waves

Sound is the oscillation of compression waves in the atmosphere. When these compression waves oscillate in equal intervals of time, we hear these waves as *tones*. The number of cycles of a wave that oscillate per unit of time is called the *frequency*, often measured in cycles per second (Hertz). The duration of each cycle is called the *period*, often measured in seconds per cycle. Hence the frequency f and the period T of a wave are related by the formula $f = 1/T$. The *amplitude* of a sound wave is the difference between the equilibrium pressure (no pressure) and the maximum pressure (maximum displacement of air molecules) caused by the wave. Humans hear the frequency of a sound wave as *pitch* and the amplitude as *volume*.

1.2 Modeling with trigonometric functions

Sound waves are longitudinal. The vibration of the sound wave (and the displacement of the air molecules) is parallel to the direction in which the sound wave propagates. When the frequency and amplitude of a sine (or cosine) wave are identical to those of a sound wave, the sine wave can be used to represent those relevant aspects of sound waves. Hence although sine and cosine waves are not longitudinal, they can be used to model sound waves. The simplest of all sound waves is a *pure tone*.

Definition. A **pure tone** is a sound wave that is modeled by a function of the form

$$p(t) = A\sin(2\pi ft)$$

where A is the amplitude and f is the frequency of the pure tone.

The sound waves that we hear are rarely pure tones. Usually what we hear is the composition or superposition of pure tones, each with its own frequency and amplitude. The displacement of air molecules caused by a sound wave is equal to the sum of the displacements of the individual sound waves. This is called the *superposition principle*.

Definition. A **tone** is a sound wave that is modeled by a function of the form

$$s(t) = \sum_{m=0}^{\infty} A_m \sin(2\pi f_m t)$$

where A_m is the amplitude of the pure tone of frequency f_m of the tone.

Chapter 2: The Frequency Spectrum

One way to analyze a tone is to determine the amplitudes of its pure tone components. The *frequency spectrum* of a tone is a description of its pure tone composition in terms of the amplitudes that correspond to the frequencies of its pure tone components.

2.1 The Fourier series

Under certain restrictions, a periodic function that models a tone has a representation as a *Fourier series* that describes the tone as a sum, or superposition, of pure tones. The frequencies of these pure tones are integer multiples of the *fundamental frequency* of the tone. Since the fundamental frequency $\frac{1}{T}$ is the frequency of the pure tone component whose wave oscillates exactly once on the interval $[0, T]$, the nth pure tone in the sum has n periods on the interval $[0, T]$ and frequency $\frac{n}{T}$. Hence the Fourier series of a periodic function that models a tone decomposes the tone into a sum of its pure tone components of frequency $\frac{n}{T}$, providing a complete description of the tone's frequency spectrum.

It is important to note that for the tones of our study, and throughout the discussions and proofs of the remainder of this thesis, we will assume that $T = 1$. Notice then that the nth pure tone of the tones of our study have frequency n and that the fundamental frequency is 1. For more general discussions and proofs, see [2].

Definition. Given a periodic function f with some restrictions, including piecewise continuity on a closed and bounded interval $[0, 1]$, the **complex form of the Fourier series** of f is

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i2\pi nx}$$

where the *Fourier coefficients* c_n are given by

$$c_n = \int_0^1 f(x) e^{-i2\pi nx} dx.$$

2.1.1 Fourier series convergence

In general, when a series converges to a function f, the series is said to be equal to $f(x)$ for all values of x. What conditions are sufficient for a Fourier series to converge to a function? To answer this question is beyond the scope of this thesis. Instead we show that the Fourier series of the piecewise continuous functions of our study converge to these functions. To show this, we need the following definition.

Definition. A function f is *Lipschitz at a point* x_L if there exists a positive constant A such that for all x near x_L where $x \neq x_L$

$$|f(x) - f(x_L)| \leqslant A\,|x - x_L|\,.$$

Proposition 2.1.1. *Let f be Lipschitz at x_L. The Fourier series of f converges to f at x_L.*

Proof. Suppose a function s is sampled at N equally spaced points x_k for $k = 0, \ldots, N-1$. Then the function f that models the samples is a step function of the form

$$f(x) = \begin{cases} s(x_0) & : x_0 \leqslant x < x_1 \\ s(x_1) & : x_1 \leqslant x < x_2 \\ \ldots & \\ s(x_k) & : x_k \leqslant x < x_{k+1} \\ s(x_{k+1}) & : x_{k+1} \leqslant x < x_{k+2} \\ \ldots & \\ s(x_{N-2}) & : x_{N-2} \leqslant x < x_{N-1} \\ s(x_{N-1}) & : x_{N-1} \leqslant x < x_N. \end{cases}$$

By definition, with the possible exception of the endpoints, f is Lipschitz at each point x_L within each kth interval since for each of these points, $f(x) = s(x_k) = f(x_L)$. We want to show that the Fourier series of f converges to f at each point x_L. Equivalently, we want to show that for each x_L

$$\lim_{|n|\to\infty} S_N(x_L) = \lim_{|n|\to\infty} \sum_{n=-N}^{N} c_n e^{i2\pi n x_L} = f(x_L).$$

Let the function g be defined by

$$g(x) = \frac{f(x) - f(x_L)}{e^{i2\pi x} - e^{i2\pi x_L}}.$$

Since f is Lipschitz at x_L,

$$\frac{|f(x) - f(x_L)|}{|e^{i2\pi x} - e^{i2\pi x_L}|} \leqslant A \frac{|x - x_L|}{|e^{i2\pi x} - e^{i2\pi x_L}|}.$$

By L'Hospital's Rule,

$$\lim_{x\to x_L} \frac{|x - x_L|}{|e^{i2\pi x} - e^{i2\pi x_L}|} = \lim_{x\to x_L} \frac{1}{2\pi\,|ie^{i2\pi x}|} = \frac{1}{2\pi}.$$

Therefore since $\lim_{x\to x_L} \frac{|x-x_L|}{|e^{i2\pi x}-e^{i2\pi x_L}|} < \infty$, $g(x) = \frac{f(x)-f(x_L)}{e^{i2\pi x}-e^{i2\pi x_L}}$ is bounded in magnitude for x near x_L. Hence $\int_0^1 |g(x)|\,dx$ is finite.

Now let d_n denote the nth coefficient of the Fourier series for $g(x)$ so that

$$d_n = \int_0^1 \frac{f(x) - f(x_L)}{e^{i2\pi x} - e^{i2\pi x_L}} e^{-i2\pi n x} dx.$$

Then

$$\begin{aligned}
d_{n-1} - d_n e^{i2\pi x_L} &= \int_0^1 \frac{(f(x) - f(x_L))e^{-i2\pi nx}}{e^{i2\pi x} - e^{i2\pi x_L}} (e^{i2\pi x/T} - e^{i2\pi x_L}) dx \\
&= \frac{1}{T} \int_0^1 f(x)e^{-i2\pi nx} - f(x_L)e^{-i2\pi nx} dx \\
&= c_n - \int_0^1 f(x_L)e^{-i2\pi nx} dx \\
&= c_n - f(x_L) \int_0^1 e^{-i2\pi nx} dx.
\end{aligned}$$

When $n = 0$, $e^{-i2\pi nx} = 1$ and $\int_0^1 e^{-i2\pi nx} dx = 1$.

When $n \neq 0$, $\int_0^1 e^{-i2\pi nx} dx = \frac{-f(x_L)}{i2\pi n}(e^{-i2\pi n} - e^{-i2\pi n(0)}) = 0$.

Hence

$$c_n = \begin{cases} d_{n-1} - d_n e^{i2\pi x_L} + f(x_L) & : n = 0 \\ d_{n-1} - d_n e^{i2\pi x_L} & : n \neq 0. \end{cases}$$

The partial sum $S_N(x_L) = \sum_{n=-N}^{N} c_n e^{i2\pi n x_L}$ telescopes:

$$\begin{aligned}
\sum_{n=-N}^{N} c_n e^{i2\pi n x_L} &= c_{-N} e^{-i2\pi N x_L} + \dots \\
&+ c_{-1} e^{-i2\pi x_L} + c_0 e^0 + c_1 e^{i2\pi x_L} + \dots \\
&+ c_N e^{i2\pi N x_L} \\
&= d_{-N-1} e^{-i2\pi N x_L} + \dots + \left(d_{-2} - d_{-1} e^{i2\pi x_L}\right) e^{-i2\pi x_L} \\
&+ d_{-1} - d_0 e^{i2\pi x_L} + f(x_L) \\
&+ \left(d_0 - d_1 e^{i2\pi x_L}\right) e^{i2\pi x_L} + \dots + d_N e^{i2\pi(N+1)x_L} \\
&= d_{-N-1} e^{-i2\pi N x_L} + \dots + f(x_L) + \dots + d_N e^{i2\pi(N+1)x_L}.
\end{aligned}$$

Now consider the inequality

$$
\begin{aligned}
0 &\leqslant \int_0^1 \left| g(x) - \sum_{n=-N}^{N} d_n e^{i2\pi nx} \right|^2 dx \\
&= \int_0^1 \left(g(x) - \sum_{m=-N}^{N} d_m e^{i2\pi mx} \right) \left(\overline{g(x)} - \sum_{n=-N}^{N} \overline{d_n} e^{-i2\pi nx} \right) dx \\
&= \int_0^1 |g(x)|^2 \, dx - \sum_{n=-N}^{N} |d_n|^2
\end{aligned}
$$

by orthogonality.[1] Therefore

$$\sum_{n=-N}^{N} |d_n|^2 \leqslant \int_0^1 |g(x)|^2 \, dx.$$

Since this inequality holds for all N, we have[2]

$$\sum_{n=-\infty}^{\infty} |d_n|^2 \leqslant \int_0^1 |g(x)|^2 \, dx.$$

Recall that $\int_0^T |g(x)| \, dx$ is finite. It follows that $\lim_{|n|\to\infty} d_n = 0$.

Then since for each x_L

$$S_N(x_L) = \sum_{n=-N}^{N} c_n e^{i2\pi n x_L} = d_{-N-1} e^{-i2\pi N x_L} + \cdots + f(x_L) + \cdots + d_N e^{i2\pi(N+1)x_L},$$

we have

$$
\begin{aligned}
\lim_{|n|\to\infty} S_N(x_L) &= \lim_{|n|\to\infty} \sum_{n=-N}^{N} c_n e^{i2\pi n x_L} \\
&= \lim_{|n|\to\infty} d_{-N-1} e^{-i2\pi N x_L} + \cdots + f(x_L) + \cdots + d_N e^{i2\pi(N+1)x_L} \\
&= f(x_L)
\end{aligned}
$$

for each x_L. $\square$

Hence the piecewise continuous step functions of our study are Lipschitz at each point x_L, and the Fourier series of these functions converge to these functions at each of these points x_L.

[1] We will define and discuss orthogonality later in this chapter.

[2] This is known as Bessel's Inequality.

2.2 The Fourier transform

For the past two centuries the Fourier transform has been the primary mathematical tool used to determine the frequency spectrum of periodic waves such as heat, light, electricity, and sound, in a wide range of fields such as geology, astronomy, acoustics, and quantum physics. It is important to notice in the following definition that while the domain of f is time or space, the domain of the Fourier transform $\widehat{f}$ is frequency. By transforming an otherwise difficult problem into the frequency domain, the Fourier transform is often the key to successful problem solving.

Definition. Let f be an absolutely integrable function over $\mathbb{R}$. The **Fourier transform** of f is denoted by $\widehat{f}$ and is uniquely defined as a function of frequency ξ by

$$\widehat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi\xi x} dx.$$

The integral defining the Fourier transform is closely related to the integral that determines the coefficients of the Fourier series of f. Indeed, if $f(x)$ is a function defined only on an interval of length 1, then it can be extended to a periodic function of period 1. In this case the Fourier transform *equals* the integral that computes the Fourier series coefficients.

2.3 The continuous to discrete Fourier transform

While the Fourier transform has been used for centuries, the *discrete Fourier transform* was birthed in the digital age of computers as an approximation of the Fourier transform. In a sense, the discrete Fourier transform was born out of the necessity to process the vast amount of information that is now in discrete, sampled form. It is the most appropriately chosen mathematical tool for any study of discrete sound wave data samples.

As an approximation of the Fourier transform of a periodic function f, the discrete Fourier transform does not exist unless f is zero outside of some finite interval. Since each tone of our study has a 1-second duration so that $T = 1$, $f(x) = 0$ when $x < 0$ or $x > 1$. Therefore for the functions f our study, the Fourier transform of f is

$$\widehat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi\xi x}dx = \int_{0}^{1} f(x)e^{-i2\pi\xi x}dx.$$

The key to the approximation of the Fourier transform of a function f that is zero outside of $[0, 1]$ is in the approximation of its integral by a *Riemann sum*:

$$\int_{0}^{1} f(x)dx \approx \sum_{k=0}^{N-1} f(x)\Delta x.$$

Using this Riemann sum approximation, the Fourier transform is given by

$$\widehat{f}(\xi) = \int_{0}^{1} f(x)e^{-i2\pi\xi x}dx \approx \sum_{k=0}^{N-1} f(x)e^{-i2\pi\xi x}\Delta x$$

where $\Delta x = \frac{1}{N}$. Now notice that when a periodic function[3] f is sampled on the interval $[0, 1]$ at N equally spaced points $x_k = k\Delta x = \frac{k}{N}$,

$$\int_{0}^{1} f(x)dx = \sum_{k=0}^{N-1} f(x_k)\Delta x.$$

[3] f might not be explicitly defined as a function. It could be given as a discrete set of data values. We often assume that f is defined at equally spaced values x_k of some variable x when, in fact, f might be known *only* at these points.

Hence for the functions f of our study, the Fourier transform of f is given by

$$\begin{aligned}\widehat{f}(\xi) &= \int_{-\infty}^{\infty} f(x)e^{-i2\pi\xi x}dx \\ &= \int_{0}^{1} f(x)e^{-i2\pi\xi x}dx \\ &= \sum_{k=0}^{N-1} f(x_k)e^{-i2\pi\xi x_k}\Delta x \\ &= \frac{1}{N}\sum_{k=0}^{N-1} f(x_k)e^{-i2\pi\xi x_k}.\end{aligned}$$

As we shall see, this sum will lead to the discrete Fourier transform.

2.4 Inner product

The inner product of two complex vectors is the primary "mathematical mechanism" of the discrete Fourier transform.

Definition. Given complex vectors $\mathbf{x} = (x_0, \ldots, x_{N-1}) \in \mathbb{C}^N$ and $\mathbf{y} = (y_0, \ldots, y_{N-1}) \in \mathbb{C}^N$, an **inner product** on $\mathbb{C}^N$ is a function $\langle ., . \rangle : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}$ that associates to $\mathbf{x}$ and $\mathbf{y}$ a complex number $\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{C}$ that satisfies the following properties:

1. Positive-definiteness: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

2. Linearity in the second argument: $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$ and $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$, for all complex numbers, α.

3. Conjugate symmetry: $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$.

Proposition 2.4.1. *Given two complex vectors $\mathbf{x} = (x_0, \ldots, x_{N-1}) \in \mathbb{C}^N$ and $\mathbf{y} = (y_0, \ldots, y_{N-1}) \in \mathbb{C}^N$,*

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=0}^{N-1} x_k \overline{y_k}$$

defines an inner product on $\mathbb{C}^N$.

Proof. We will first show that $\langle \mathbf{x}, \mathbf{y} \rangle$ is positive-definite. Let $x_k = a+bi$ where $a, b \in \mathbb{R}$. Now

$$x_k \overline{x_k} = (a+bi)(a-bi) = a^2 + b^2.$$

Since for all $a, b \in \mathbb{R}$, $a^2 \geq 0$ and $b^2 \geq 0$, we have $x_k \overline{x_k} \geq 0$. Hence $\langle \mathbf{x}, \mathbf{x} \rangle$ is the sum of nonnegative real numbers, and is also nonnegative. Furthermore, since $\langle \mathbf{x}, \mathbf{x} \rangle$ is the sum of nonnegative real numbers $x_k \overline{x_k}$,we have $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if for all k, $x_k = 0$. Hence $\langle \mathbf{x}, \mathbf{y} \rangle$ is positive-definite.

To show that $\langle \mathbf{x}, \mathbf{y} \rangle$ is linear in the second-argument, we will first show that $\langle \mathbf{x}, \mathbf{y}+\mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$. Let $y_k = c+di$ where $c, d \in \mathbb{R}$, and let $z_k = f+gi$ where $f, g \in \mathbb{R}$, so

$$(y+z)_k = y_k + z_k = c + di + f + gi = c + f + i(d+g).$$

Then

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y}+\mathbf{z} \rangle &= \sum_{k=0}^{N-1} x_k \overline{(y+z)_k} \\
&= \sum_{k=0}^{N-1} (a+bi)(c+f-i(d+g)) \\
&= \sum_{k=0}^{N-1} (a+bi)(c-di+f-gi) \\
&= \sum_{k=0}^{N-1} x_k (\overline{y_k} + \overline{z_k}) \\
&= \sum_{k=0}^{N-1} x_k \overline{y_k} + \sum_{k=0}^{N-1} x_k \overline{z_k} \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle.
\end{aligned}$$

Hence $\langle \mathbf{x}, \mathbf{y}+\mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$.

Now we want to show that $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$, for all complex numbers, α.

$$\begin{aligned}
\langle \mathbf{x}, \alpha\mathbf{y} \rangle &= \sum_{k=0}^{N-1} x_k \overline{\alpha y_k} \\
&= \sum_{k=0}^{N-1} (a+bi)(\overline{\alpha(c+di)}) \\
&= \sum_{k=0}^{N-1} (a+bi)(\overline{\alpha c + \alpha di}) \\
&= \sum_{k=0}^{N-1} (a+bi)(\alpha c - \alpha di) \\
&= \sum_{k=0}^{N-1} \alpha(a+bi)(c-di) \\
&= \alpha \sum_{k=0}^{N-1} x_k \overline{y_k} \\
&= \alpha\langle \mathbf{x}, \mathbf{y} \rangle.
\end{aligned}$$

Hence $\langle \mathbf{x}, \mathbf{y} \rangle$ is linear in the second argument.

To show that $\langle \mathbf{x}, \mathbf{y} \rangle$ is conjugate symmetric, recall that the conjugate of a sum of complex numbers is equal to the sum of their conjugates. So

$$\overline{\langle \mathbf{y}, \mathbf{x} \rangle} = \overline{\sum_{k=0}^{N-1} y_k \overline{x_k}} = \sum_{k=0}^{N-1} \overline{y_k \overline{x_k}} = \sum_{k=0}^{N-1} \overline{y_k} x_k = \sum_{k=0}^{N-1} x_k \overline{y_k} = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Hence $\langle \mathbf{x}, \mathbf{y} \rangle$ is conjugate symmetric.

Hence $\langle \mathbf{x}, \mathbf{y} \rangle$ is positive-definite, linear in the second-argument, and conjugate symmetric, and therefore defines an inner product on $\mathbb{C}^N$. □

2.5 Orthogonality

Heuristically speaking, an inner product of two vectors gives information that shows how "aligned" two vectors are. This leads us to our next definition.

Definition. Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^N$ are **orthogonal** with respect to an inner product on $\mathbb{C}^N$ if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Proposition 2.5.1. *For $k = 0, \ldots, N-1$, let*

$$\mathbf{s_m} = (e^{i2\pi m(0)/N}, \ldots, e^{i2\pi mk/N}, \ldots, e^{i2\pi m(N-1)/N}) \in \mathbb{C}^N.$$

If $m, n \in \mathbb{Z}$ and $m \neq n$, then $\mathbf{s_m}$ and $\mathbf{s_n}$ are orthogonal.

Proof. We want to show $\sum_{k=0}^{N-1} e^{i2\pi mk/N}\overline{e^{i2\pi nk/N}} = 0$. Notice that for integers m and n, $\left(e^{i2\pi(m-n)/N}\right)^N = 1$. Hence $e^{i2\pi(m-n)/N}$ is a root of the polynomial

$$z^N - 1 = (z-1)(z^{N-1} + z^{N-2} + \cdots + z + 1) = (z-1)\sum_{k=0}^{N-1} z^k.$$

Since $m \neq n$, $e^{i2\pi(m-n)/N} \neq 1$. Hence the sum is zero, and we have

$$\begin{aligned}
\sum_{k=0}^{N-1} z^k &= \sum_{k=0}^{N-1} (e^{i2\pi(m-n)/N})^k \\
&= \sum_{k=0}^{N-1} e^{i2\pi(m-n)k/N} \\
&= \sum_{k=0}^{N-1} e^{i2\pi mk/N} e^{-i2\pi nk/N} \\
&= \sum_{k=0}^{N-1} e^{i2\pi mk/N}\overline{e^{i2\pi nk/N}} = 0.
\end{aligned}$$

Hence if $m, n \in \mathbb{Z}$ and $m \neq n$, then $\mathbf{s_m}$ and $\mathbf{s_n}$ are orthogonal. $\square$

Proposition 2.5.2. *Let $\mathbf{s_m} = (e^{i2\pi m(0)/N}, \ldots, e^{i2\pi mk/N}, \ldots, e^{i2\pi m(N-1)/N}) \in \mathbb{C}^N$. Then $\langle \mathbf{s_m}, \mathbf{s_m} \rangle = N$.*

Proof. $\langle \mathbf{s_m}, \mathbf{s_m} \rangle = \sum_{k=0}^{N-1} e^{i2\pi mk/N}\overline{e^{i2\pi mk/N}} = \sum_{k=0}^{N-1} e^0 = N$. $\square$

2.6 The discrete Fourier transform and the frequency spectrum

The discrete Fourier transform is used to determine the frequency spectrum of tones. This is analagous to using a prism to decompose white light while computing the magnitude of each of its component colors. Orthogonality is the key to this process.

In order to determine if a pure tone of frequency n is present in a tone x, we take N samples of x, generating the vector $\mathbf{x} = (s_0, \ldots, x_{N-1}) \in \mathbb{C}^N$, and compute $\langle \mathbf{x}, \mathbf{s_n} \rangle$ where

$$\mathbf{s_n} = (e^{i2\pi n(0)/N}, \ldots, e^{i2\pi nk/N}, \ldots, e^{i2\pi n(N-1)/N}) \in \mathbb{C}^N.$$

In this computation, orthogonality filters out the amplitude A_m of any pure tone of frequency m unless $m = n$. In other words, $\langle \mathbf{x}, \mathbf{s_n} \rangle = 0$ unless there exists a pure tone of frequency n in tone x.

The discrete Fourier transform has several commonly used forms. We'll adopt this one:

Definition. Given a vector $\mathbf{x} = (x_0, \ldots, x_k, \ldots, x_{N-1}) \in \mathbb{C}^N$ and a vector

$$\mathbf{s_n} = (e^{i2\pi n(0)/N}, \ldots, e^{i2\pi nk/N}, \ldots, e^{i2\pi n(N-1)/N}) \in \mathbb{C}^N,$$

the **discrete Fourier transform** of $\mathbf{x}$ is the vector $\mathbf{X} = (X_0, \ldots, X_n, \ldots, X_{N-1}) \in \mathbb{C}^N$ where

$$X_n = \langle \mathbf{x}, \mathbf{s_n} \rangle = \sum_{k=0}^{N-1} x_k \overline{s_n(k)} = \sum_{k=0}^{N-1} x_k e^{-i2\pi nk/N}.$$

Recall from section 2.3 that for the functions f of our study the Fourier transform of f is given by

$$\widehat{f}(\xi) = \frac{1}{N} \sum_{k=0}^{N-1} f(x_k) e^{-i2\pi \xi x_k}.$$

Hence $\mathbf{F} = (F_0, \ldots, F_\xi, \ldots, F_{N-1})$ is the discrete Fourier transform of

$$\mathbf{f} = (f(x_0), \ldots, f(x_k), \ldots, f(x_{N-1}))$$

where

$$F_\xi = \langle \mathbf{f}, \mathbf{s}_\xi \rangle = \sum_{k=0}^{N-1} f(x_k) e^{-i2\pi\xi x_k} = N\widehat{f}(\xi)$$

and $\widehat{f}(\xi)$ is the Fourier transform of f.

Proposition 2.6.1. *Suppose that N samples of a tone*

$$f(x) = \sum_{m=0}^{\infty} A_m \sin(2\pi m x)$$

are taken, generating a vector $\mathbf{f} = (f(x_0), \ldots, f(x_k), \ldots, f(x_{N-1}))$ where

$$f(x_k) = \sum_{m=0}^{\infty} A_m \sin(2\pi m x_k) = \sum_{m=0}^{\infty} A_m \sin(2\pi mk/N),$$

A_m is the amplitude of the pure tone of frequency m, and $\mathbf{F} = (F_0, \ldots, F_n, \ldots, F_{N-1})$ is the discrete Fourier transform of $\mathbf{f}$. Then for all nonnegative integers n, $S_n = |F_n|$ is directly proportional to A_n.

Proof. The discrete Fourier transform of $\mathbf{f}$ is the vector $\mathbf{F}$ where

$$\begin{aligned}
F_n = \langle \mathbf{f}, \mathbf{s_n} \rangle &= \sum_{k=0}^{N-1} f(x_k) \overline{e^{i2\pi nk/N}} \\
&= \sum_{k=0}^{N-1} \left(\sum_{m=0}^{\infty} (A_m \sin(2\pi mk/N)) e^{-i2\pi nk/N} \right) \\
&= \sum_{k=0}^{N-1} \left(\sum_{m=0}^{\infty} (A_m \frac{e^{i2\pi mk/N} - e^{-i2\pi mk/N}}{2i}) e^{-i2\pi nk/N} \right) \\
&= \frac{1}{2i} \sum_{m=0}^{\infty} A_m \sum_{k=0}^{N-1} \left(e^{i2\pi mk/N} e^{-i2\pi nk/N} - e^{-i2\pi mk/N} e^{-i2\pi nk/N} \right) \\
&= \frac{1}{2i} \sum_{m=0}^{\infty} A_m (\langle \mathbf{s_m}, \mathbf{s_n} \rangle - \langle \mathbf{s_{-m}}, \mathbf{s_n} \rangle).
\end{aligned}$$

Suppose $n = 0$. Then[4]

$$\begin{aligned}
F_n = F_0 &= \frac{1}{2i}\sum_{m=0}^{\infty} A_m(\langle \mathbf{s_m}, \mathbf{s_0}\rangle - \langle \mathbf{s_{-m}}, \mathbf{s_0}\rangle) \\
&= \frac{1}{2i} A_0(\langle \mathbf{s_0}, \mathbf{s_0}\rangle - \langle \mathbf{s_0}, \mathbf{s_0}\rangle) \text{ (by orthogonality)} \\
&= 0.
\end{aligned}$$

Therefore for $n = 0$, $S_n = |F_n| = 0$. Since A_0 is the amplitude of the pure tone of frequency 0, $A_0 = 0$. Hence for $n = 0$, S_n is directly proportional to A_n.

Now suppose $n \in \mathbb{Z}^+$. Then

$$\begin{aligned}
F_n &= \frac{1}{2i}\sum_{m=0}^{\infty} A_m(\langle \mathbf{s_m}, \mathbf{s_n}\rangle - \langle \mathbf{s_{-m}}, \mathbf{s_n}\rangle) \\
&= \frac{1}{2i} A_n(\langle \mathbf{s_n}, \mathbf{s_n}\rangle - \langle \mathbf{s_{-n}}, \mathbf{s_n}\rangle) \text{ (by orthogonality)} \\
&= \frac{1}{2i} A_n(\sum_{k=0}^{N-1} 1 - \sum_{k=0}^{N-1} 0) \\
&= \frac{NA_n}{2i}.
\end{aligned}$$

Therefore for $n \in \mathbb{Z}^+$, $S_n = |F_n| = \left|\frac{NA_n}{2i}\right| = \frac{NA_n}{2}$.

Hence for all nonnegative integers n, S_n is directly proportional to A_n. $\square$

We conclude this chapter by defining the frequency spectrum of a tone s in terms of the discrete Fourier transform.

Definition. The **frequency spectrum** of a tone

$$s(x) = \sum_{m=0}^{\infty} A_m \sin(2\pi m x)$$

is the vector $\mathbf{S} = (S_0, \ldots, S_{N-1})$ where $S_n = |F_n|$, $\mathbf{F} = (F_0, \ldots, F_{N-1})$ is the discrete Fourier transform of $\mathbf{s} = (s(x_0), \ldots, s(x_{N-1}))$, and $s(x_k)$ is the kth sample of s.

[4]Since $F_0 = \langle \mathbf{f}, \mathbf{s_0}\rangle = \sum_{k=0}^{N-1} f(x_k)\overline{e^{i2\pi(0)k/N}} = \sum_{k=0}^{N-1} f(x_k)$, F_0 is equal to the sum of the elements of the input data.

Chapter 3: Analysis of Effects

Our goal is to analyze and model five standard effects on sound waves, namely *vibrato*, *chorus*, *flanger*, *phaser*, and *overdrive*. In order to widen the scope of our analysis, we will use pure tones of ten selected frequencies ranging from 220 Hz to 1320 Hz: 220, 440, 495, 528, $586\frac{2}{3}$, 660, $733\frac{1}{3}$, 792, 880, and 1320 Hz. These frequencies represent the A Dorian Minor scale, with the addition of the 220 and 1320 Hz frequencies.

Each pure tone will be separately altered by each of three versions of each of the five effects, generating 150 tones. Using *Mathematica*'s `Fourier` command, we will generate the frequency spectrum of each of these tones. Then based on an analysis of 30 frequency spectra for each effect, we will use *Mathematica*'s `Fit` command to estimate a 'best fit' function that models how the magnitude of each prominent component varies with input frequency for each of the five effects. Using these models, as well as our general frequency spectra observations, we will define how the effect alters pure tones of varying frequencies. The definitions for each effect of our analysis are below. We outline the analysis procedures as follows.

3.1 Analysis procedures

3.1.1 Analysis definitions

We will classify each effect in our study as an *effect*, a *pure tone effect*, or a *tone effect*. Since a tone usually consists of more than one pure tone, we need to define what we mean when we refer to the frequency of a tone.

Definition. The **frequency of a tone**

$$s(t) = \sum_{m=0}^{\infty} A_m \sin(2\pi f_m t)$$

is the frequency f_n corresponding to the maximum amplitude A_n, when there is a maximum amplitude. A maximum amplitude exists when a pure tone has a greater amplitude than the amplitude of all of the other pure tones of the tone.

We'll denote the set of all pure tones by $\mathcal{P}$ and the set of all pure tones of frequency n by $\mathcal{P}_n$. Likewise, we'll denote the set of all tones by $\mathcal{T}$ and the set of all tones of frequency n by $\mathcal{T}_n$. Notice that, by definition, every pure tone is a tone. Hence, $\mathcal{P}$ is a proper subset of $\mathcal{T}$. Also notice that, by defintion, two or more distinct tones can have the same frequency, but two or more pure tones of the same frequency are not unique. Hence $\mathcal{P}$ is the disjoint union of the $\mathcal{P}_n$'s where n runs from zero to infinity, and $\mathcal{T}$ is a proper super-set of the joint union of the $\mathcal{T}_n$'s where n runs from zero to infinity.

Definition. An **effect** is a function $\mu : \mathcal{T} \to \mathcal{T}$.

Definition. A **pure tone effect** is a function $\gamma : \mathcal{P} \to \mathcal{T}$ such that for all values of f, $\gamma : \mathcal{P}_f \to \mathcal{T}_f$.

Definition. A **tone effect** is a function $\Gamma : \mathcal{T} \to \mathcal{T}$ such that for all values of f, $\Gamma : \mathcal{T}_f \to \mathcal{T}_f$.

By definition, the input of a pure tone effect is a pure tone, and the output is a tone whose frequency is equal to the frequency of the pure tone. Likewise, the input of a tone effect is a tone, and the output is a tone whose frequency is equal to the frequency of the input tone. An effect has no frequency restrictions. Its input and output are tones of any frequency.

Note: After we determined that the effect that overdrive has on tones does not have a linear relation to its effect on pure tones, we decided that describing effects based on our definition of *tone effect* would be an overwhelmingly time-consuming endeavor. Based on this observation we decided to focus on describing how the five effects alter pure tones. Although it was not a primary focus of our analysis, we determined which effects satsify our definition of *pure tone effect*.

3.1.2 Step-by-step procedures

The first step was to generate 10 1-second pure tones using *Mathematica*'s `Sin` command. Then for each pure tone, we

1. used *Mathematica*'s `Play` command to take data samples of the tone at 11,025 samples per second.

2. used *Mathematica*'s `Export` command to convert the data into a .aif file.

3. used *GarageBand* to alter the sound wave data with one of 15 effects.

4. used *Mathematica*'s `Import`, `Fullform`, and `Table` commands to convert the new `.aif` file into 11,025 data samples.

5. used *Mathematica*'s `Fourier` and `ListLinePlot` commands to generate and plot the frequency spectrum of the altered sound wave data.

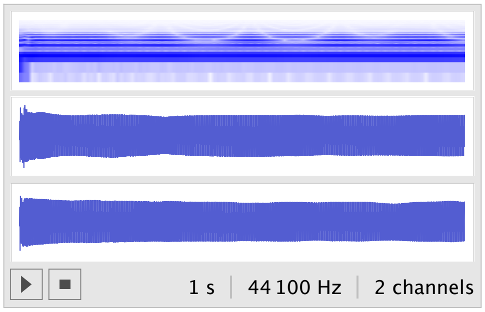


Figure 3.1: Using *Mathematica*'s `Import` command, a tone is imported into *Mathematica* as the output of the *GarageBand* effect, *Choral Stack*, whose input was a 528 Hz pure tone.

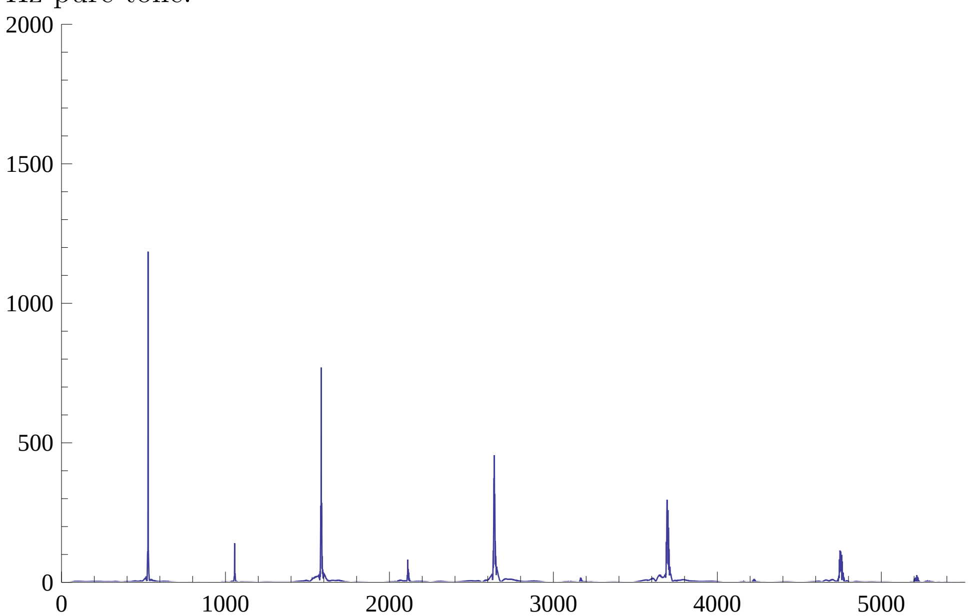


Figure 3.2: Frequency spectrum of the output tone of *Choral Stack* whose input was a 528 Hz pure tone

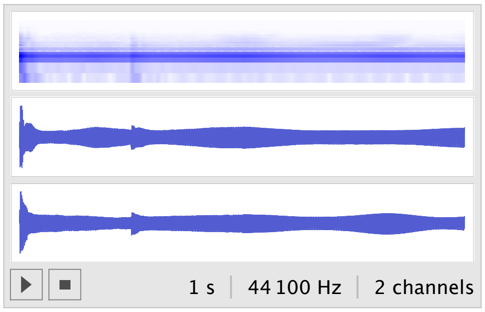


Figure 3.3: Using *Mathematica*'s `Import` command, a tone is imported into *Mathematica* as the output of the *GarageBand* effect, *Chorus Shimmer*, whose input was a 528 Hz pure tone. Notice the the tone's 'thin' appearance correspnding to its 'thin' frequency spectrum below.

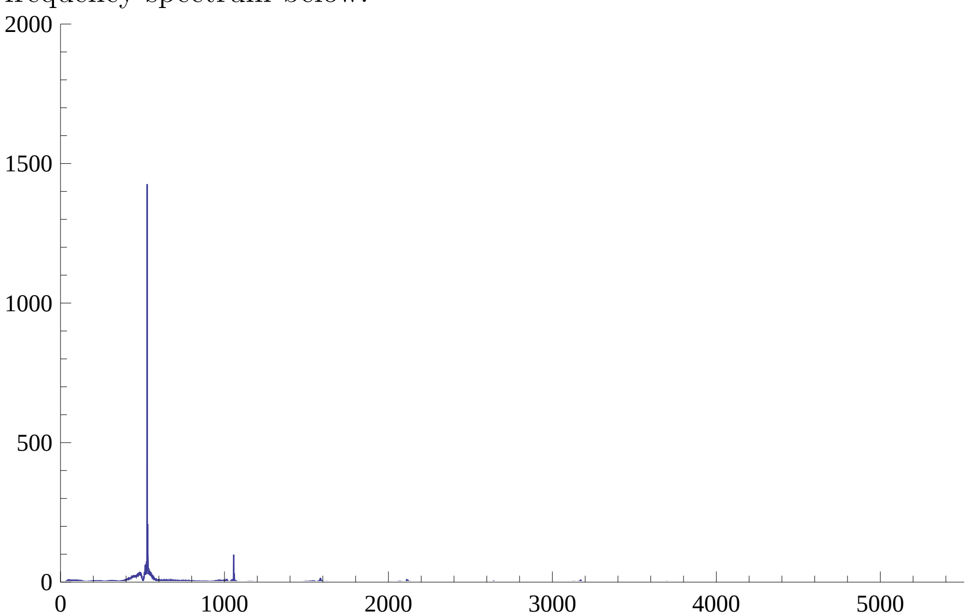


Figure 3.4: Frequency spectrum of the output tone of *Chorus Shimmer* whose input was a 528 Hz pure tone

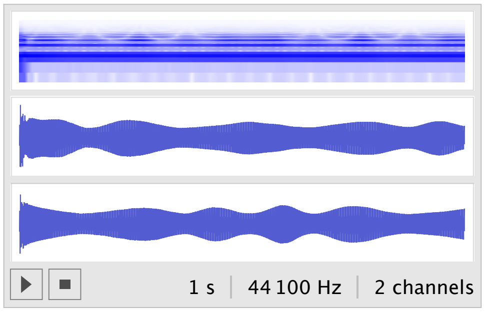


Figure 3.5: Using *Mathematica*'s `Import` command, a tone is imported into *Mathematica* as the output of the *GarageBand* effect, *Chunky Chorus*, whose input was a 528 Hz pure tone.

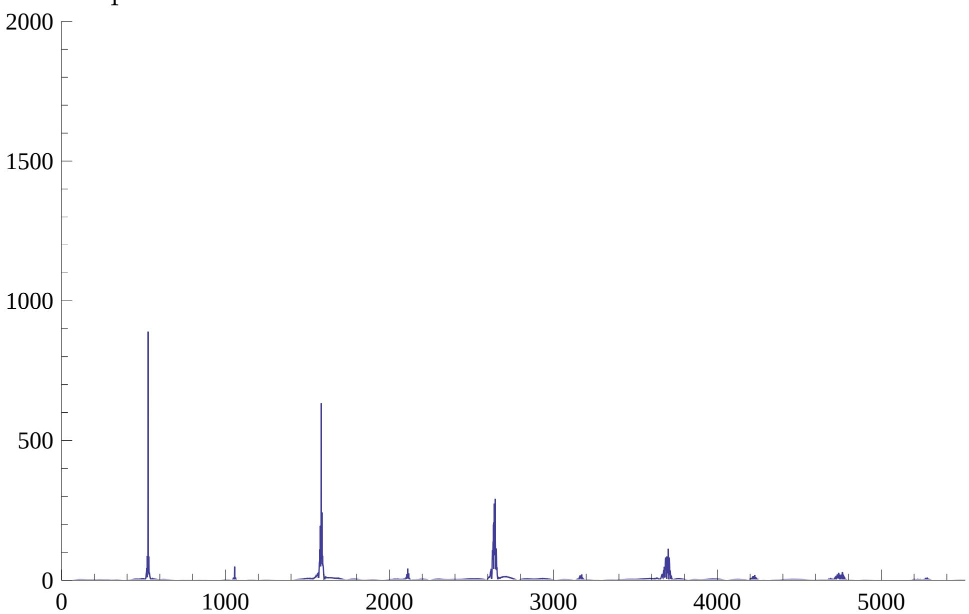


Figure 3.6: Frequency spectrum of the output tone of *Chunky Chorus* whose input was a 528 Hz pure tone

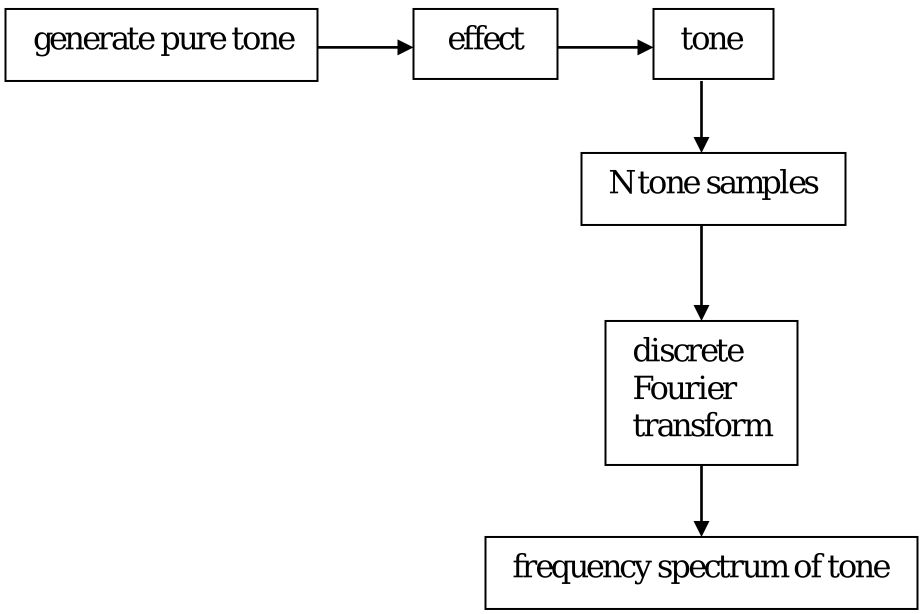


Figure 3.7: Flowchart of analysis procedures for determining the frequency spectrum of an output tone of an effect on sound

Execution of these steps generated 150 frequency spectra, 30 for each of the five standard effects of our study. Our next step was to determine the four greatest S_n values of each of the 30 frequency spectra for each effect.[5] Through careful obseravtion of each set of 120 data points, we determined the most significant and common pure tone components of the output tones of each effect. Then recognizing that the magnitude of each prominent component varies with input frequency, we decided that the best way to model each of the five effects was to use *Mathematica*'s `Fit` command to find a 'best fit' function model that describes how the prominence of each component varies with the frequency of the input pure tone of frequency f. Therefore the next step was to create a list of (f, S_n) data points for each component of frequency n.

For example, for overdrive we determined that the most significant components are of frequencies $n = f$, $n = 3f$, $n = 4f$, $n = 5f$, and $n = 7f$. For each of these five components, 30 (f, S_n) data points were placed in a list, one for each of 10 frequencies and three chosen examples of the overdrive effect. For example, the list for component $n = 5f$ contains 30 (f, S_{5f}) data points. Over the domain of 10 frequencies from 220 Hz to 1320 Hz, the data points in this list quantify how much the $5f$ component is present in the frequency spectra of the three overdrive effects.

We found in our effect analysis that third degree polynomials best model the behavior of effects on sound. Hence, to model how the magnitude of each component varies with frequency, we used cubic functions[6] of the form $ax^3 + bx^2 + cx + d$. In conjunction with *Mathematica*'s `Fit` command, we determined a best fit model for

[5]This provided us with 600 total data points to consider, 40 for each of 15 effects, grouped in sets of 120 per five standard effects.

[6]This choice was made based primarily on our frequency spectrum observations. It is certainly possible that for some effects a 2nd or 4th degree model is a better choice.

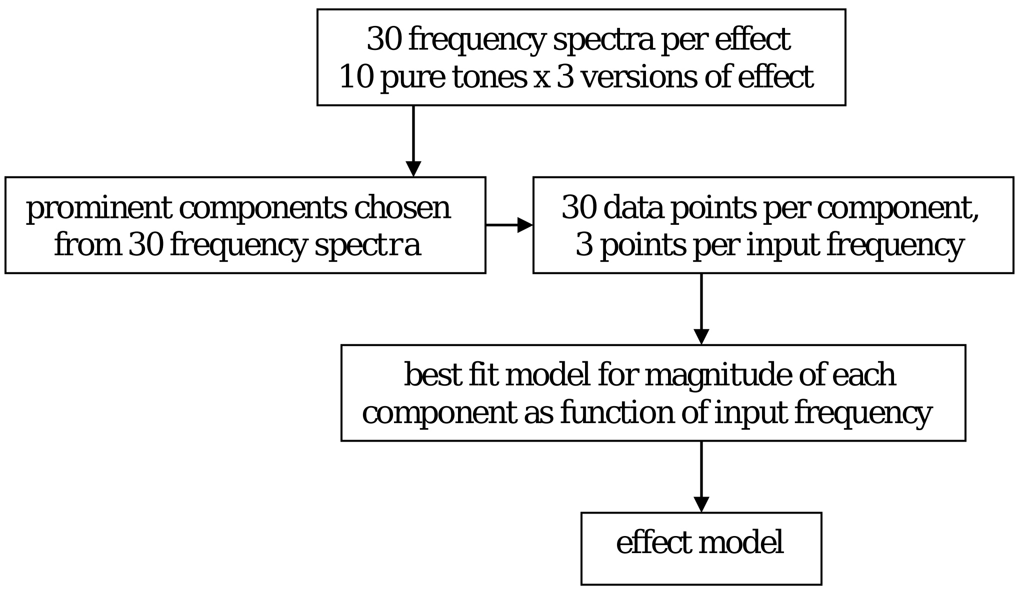


Figure 3.8: Flowchart of analysis procedures for modeling an effect on sound

each significant component of each effect. Using these models we were able to define each effect in terms of the behavior of these components over the 220 Hz to 1320 Hz frequency domain. The final step of our analysis was to compare these definitons with their traditionally accepted definitions. The results are decribed below.

Now that we have outlined our step-by-step procedures, we describe the method that we used to estimate the best fit function models for the effects, *the method of least squares*.

3.1.3 The method of least squares

When a set of data points do not "fit" along a line or curve, the *method of least squares* is often preferred to determine a "least squares approximation" function that models the data. *Mathematica*'s `Fit` command uses this method.

The method of least squares determines the best fit function model such that the sum of the squares of the distances between the data points and the points of the model is minimzed. Simply put, this method minimizes the *discrete least squares error*, a real-valued, nonnegative function.

The method of least squares is based on the variance σ_x^2 of a set of data $\{x_0, \ldots, x_{N-1}\}$, a tool commonly used by statisticians that quantifies how much a set of data varies from its mean $\overline{x}$. The variance is given by

$$\sigma_x^2 = \frac{1}{N}\sum_{i=0}^{N-1}(x_i - \overline{x})^2.$$

When the method of least squares is used, larger errors are weighted more because each error is squared in the sum. One might choose to remove such weights by using $\frac{1}{N}\sum_{i=0}^{N-1}(x_i - \overline{x})$, but then not only could terms of the sum be negative, the sum would always be zero. One might also consider using absolute values as in $\frac{1}{N}\sum_{i=0}^{N-1}|(\mathbf{x}_i - \overline{x})|$.

But this is also not a good choice since the absolute value function is not differentiable. Hence although the use of the sum of the squares causes a disproportionate effect on the fit, it enables us to work with a continuous differentiable quantity.

The discrete least squares error that is minimized by the method of least squares is simply $N\sigma_x^2$ where σ_x^2 is the variance of the data set of size N. Given a set of N data points

$$\{(f_0, A_{n_0}) \ldots, (f_{N-1}, A_{n_{N-1}})\},$$

the discrete least squares error is the function

$$\begin{aligned} E(a,b,c,d) &= \sum_{k=0}^{N-1} (A_{n_k} - g(f_k))^2 \\ &= \sum_{k=0}^{N-1} \left(\left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right)^2 \right). \end{aligned}$$

To find the best fit model, we need to determine the values of the a, b, c, and d coefficients such that this discrete least squares error function is minimized. This equates to finding the values of these coefficients such that $\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = \frac{\partial E}{\partial c} = \frac{\partial E}{\partial d} = 0$. Notice that the $E(a,b,c,d)$ values increase with the absolute values of the coefficients. Therefore we do not need to check the boundary points for minimum values.

It suffices to compute the partial derivatives and set them equal to zero:

$$\begin{aligned} \frac{\partial E}{\partial a} &= \sum_{k=0}^{N-1} 2\left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right)(-f_k^3) = 0 \\ \frac{\partial E}{\partial b} &= \sum_{k=0}^{N-1} 2\left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right)(-f_k^2) = 0 \\ \frac{\partial E}{\partial c} &= \sum_{k=0}^{N-1} 2\left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right)(-f_k) = 0 \\ \frac{\partial E}{\partial d} &= \sum_{k=0}^{N-1} 2\left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right) = 0. \end{aligned}$$

Then dividing by 2 we have

$$\sum_{k=0}^{N-1} \left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right)(-f_k^3) = 0$$
$$\sum_{k=0}^{N-1} \left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right)(-f_k^2) = 0$$
$$\sum_{k=0}^{N-1} \left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right)(-f_k) = 0$$
$$\sum_{k=0}^{N-1} \left(A_{n_k} - (af_k^3 + bf_k^2 + cf_k + d)\right) = 0.$$

Rewriting these equations in terms of linear combinations of the sums, we have

$$\left(\sum_{k=0}^{N-1} f_k^6\right) a + \left(\sum_{k=0}^{N-1} f_k^5\right) b + \left(\sum_{k=0}^{N-1} f_k^4\right) c + \left(\sum_{k=0}^{N-1} f_k^3\right) d = \sum_{k=0}^{N-1} f_k^3 A_{n_k}$$
$$\left(\sum_{k=0}^{N-1} f_k^5\right) a + \left(\sum_{k=0}^{N-1} f_k^4\right) b + \left(\sum_{k=0}^{N-1} f_k^3\right) c + \left(\sum_{k=0}^{N-1} f_k^2\right) d = \sum_{k=0}^{N-1} f_k^2 A_{n_k}$$
$$\left(\sum_{k=0}^{N-1} f_k^4\right) a + \left(\sum_{k=0}^{N-1} f_k^3\right) b + \left(\sum_{k=0}^{N-1} f_k^2\right) c + \left(\sum_{k=0}^{N-1} f_k\right) d = \sum_{k=0}^{N-1} f_k A_{n_k}$$
$$\left(\sum_{k=0}^{N-1} f_k^3\right) a + \left(\sum_{k=0}^{N-1} f_k^2\right) b + \left(\sum_{k=0}^{N-1} f_k\right) c + \left(\sum_{k=0}^{N-1} 1\right) d = \sum_{k=0}^{N-1} A_{n_k}.$$

Hence the values of the a, b, c, and d coefficients that minimize the discrete least squares error function satisfy the matrix equation

$$\begin{pmatrix} \sum_{k=0}^{N-1} f_k^6 & \sum_{k=0}^{N-1} f_k^5 & \sum_{k=0}^{N-1} f_k^4 & \sum_{k=0}^{N-1} f_k^3 \\ \sum_{k=0}^{N-1} f_k^5 & \sum_{k=0}^{N-1} f_k^4 & \sum_{k=0}^{N-1} f_k^3 & \sum_{k=0}^{N-1} f_k^2 \\ \sum_{k=0}^{N-1} f_k^4 & \sum_{k=0}^{N-1} f_k^3 & \sum_{k=0}^{N-1} f_k^2 & \sum_{k=0}^{N-1} f_k \\ \sum_{k=0}^{N-1} f_k^3 & \sum_{k=0}^{N-1} f_k^2 & \sum_{k=0}^{N-1} f_k & \sum_{k=0}^{N-1} 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} \sum_{k=0}^{N-1} f_k^3 A_{n_k} \\ \sum_{k=0}^{N-1} f_k^2 A_{n_k} \\ \sum_{k=0}^{N-1} f_k A_{n_k} \\ \sum_{k=0}^{N-1} A_{n_k} \end{pmatrix}.$$

Solutions to this equation exist if and only if the matrix

$$M = \begin{pmatrix} \sum_{k=0}^{N-1} f_k^6 & \sum_{k=0}^{N-1} f_k^5 & \sum_{k=0}^{N-1} f_k^4 & \sum_{k=0}^{N-1} f_k^3 \\ \sum_{k=0}^{N-1} f_k^5 & \sum_{k=0}^{N-1} f_k^4 & \sum_{k=0}^{N-1} f_k^3 & \sum_{k=0}^{N-1} f_k^2 \\ \sum_{k=0}^{N-1} f_k^4 & \sum_{k=0}^{N-1} f_k^3 & \sum_{k=0}^{N-1} f_k^2 & \sum_{k=0}^{N-1} f_k \\ \sum_{k=0}^{N-1} f_k^3 & \sum_{k=0}^{N-1} f_k^2 & \sum_{k=0}^{N-1} f_k & \sum_{k=0}^{N-1} 1 \end{pmatrix}$$

is invertible if and only if the determinant of M is nonzero.

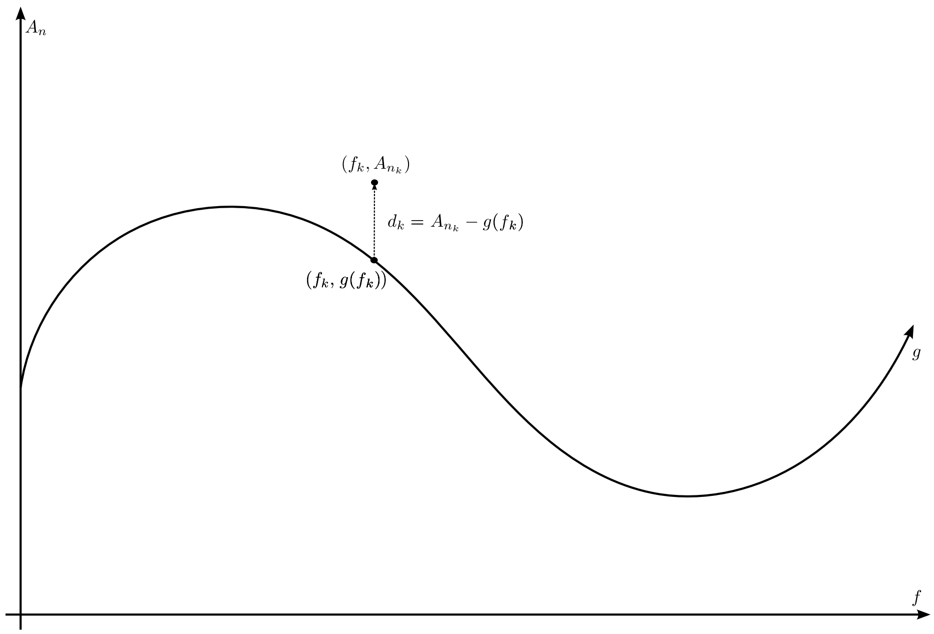


Figure 3.9: *Mathematica*'s `Fit` command uses vertical fitting. The squared distances are vertical, not perpendicular, to the $g(f)$ "best fit" model.

The determinant of M is nonzero if the f_k values are not all equal. Hence a least squares solution exists unless the f_k values are all equal. For our study, this means that a least squares solution exists unless the input of each effect is a tone of the same frequency.

3.2 Effect modeling and analysis

With the exception of overdrive, the effects that we have chosen to analyze are often categorized as *modulation effects*. Modulation effects change in real time according to user specifications. These changes occur when the parameters of the effect's processor are varied or modulated. The most common parameters are called "delay time" and "pitch shift." These parameters can be manually or automatically modulated. For automatic modulation, most often a *low-frequency oscillator* is used to control the value of the selected parameter over time. A low-frequency oscillator is a sinusoidal signal that is commonly used to modulate parameters of other signals at a frequency that is usually below 20 Hz.

3.2.1 Vibrato

Vibrato is a modulation effect characterized by a regular modulation of pitch. It is often confused with *tremelo* which is characterized by the modulation of amplitude. It has two fundamental parameters, the amount of pitch variation and the frequency of pitch variation. Both of these parameters are typically controlled by a low-frequency oscillator. The amplitude of the low frequency oscillator (or other modulation source) determines how much the pitch modulates, and the frequency of the low frequency oscillator determines how often the pitch modulates. Both the frequency and the amplitude of the modulation are very low. The frequency of modulation (how *often* the pitch varies) is typically between 4 and 7 Hz. The amplitude of the modulation (*amount* of pitch variation) is typically between one-tenth of a semi-tone (10 cents) and a whole tone (200 cents).

Vibrato can be used to give sound a natural, “warm” quality or to add expression within a musical composition. It can be produced in a variety of ways by musicians. It naturally results from the free oscillation of vocal cords, especially for high register vocalists, and it can be produced by modulating the air flow into wind instruments or by bending the strings of a guitar.

After observing the frequency spectra of 30 output tones of the three vibrato effects that we used in our analysis, we determined that by definition, since for all m and j, $\alpha(m,j) > \beta(m,j)$, vibrato is a pure tone effect.

Definition. Vibrato is a pure tone effect $\xi : \mathcal{P} \to \mathcal{S}$ such that

$$\begin{aligned}\omega : \sin(2\pi m t) \mapsto \sum_{j=1}^{\infty} \alpha(m,j)\sin(2\pi j m t) &+ \sum_{j=1}^{m-1} \beta(m,j)\sin(2\pi(m-j)t) \\ &+ \sum_{j=1}^{\infty} \beta(m,j)\sin(2\pi(m+j)t)\end{aligned}$$

where m is the frequency of the input pure tone,

$$\alpha(m,j) = \begin{cases} f(m,j) & : f(m,j) \geq 0 \\ 0 & : otherwise, \end{cases}$$

$$\beta(m,j) = \begin{cases} g(m,j) & : g(m,j) \geq 0 \\ 0 & : otherwise, \end{cases}$$

and

$f(m,1) \approx 0.05 + 0.0016m - 3.0 \times 10^{-6}m^2 + 1.5 \times 10^{-9}m^3,$

$f(m,3) \approx 0.3 - 0.0012m + 1.7 \times 10^{-6}m^2 - 7.0 \times 10^{-10}m^3,$

$f(m,5) \approx -0.2 + 0.0016m - 3.0 \times 10^{-6}m^2 + 1.75 \times 10^{-9}m^3,$

$f(m,j) \approx 0, \forall j \notin \{1,3,5\},$

$g(m,4) \approx g(m,6) \approx g(m,8) \approx -0.05 + 0.0004m - 4.0 \times 10^{-7}m^2 + 1.2 \times 10^{-10}m^3,$

$g(m,j) \approx 0, \forall j \notin \{1,2,3\}.$

Our traditional and mathematical definitions are not contradictory. According to our traditional description, vibrato is characterized by the periodic variation of the pitch of a sound wave. In our mathematical description, vibrato receives a pure tone of frequency m as input and outputs a tone with a most prominent pure tone component of frequency m and less prominent components of frequency $m \pm j$ where $m \in \{1, 3, 5\}$. Clearly the variations of pitch of these latter components are characteristic results of the effect of vibrato on the original pure tone.

Also according to our traditional description, vibrato is characterized by pitch modulation that typically varies between one-tenth of a semi-tone and a whole tone. According to our mathematical description, vibrato is characterized by pitch modulation of 1, 3 and 5 Hz. We should note that these values may vary with each vibrato effect. It should not be assumed that *all* vibrato effects are characterized by these particular pitch shift values. 1, 3 and 5 Hz were the most prominent pitch shift values found in our analysis, so we defined our model accordingly.

We analyzed 30 tones, each the output of one of 3 vibrato effects whose input was a pure tone of frequency f in Hz where $220 \leqslant f \leqslant 1320$. One semi-tone above 220 Hz is approximately 233 Hz. Then one-tenth of a semi-tone above 220 Hz corresponds to a pitch variation of approximately 1.3 Hz, and one whole tone above 220 Hz corresponds to a pitch variation of approximately 26 Hz. So according to our traditional definition, the pitch modulation for a 220 Hz pure tone *typically* varies between 1.3 Hz and 26 Hz. According to our model, the pitch modulation values are 1, 3 and 5 Hz. Since our 1 Hz minimum variation is only slightly below the *typical* minimum, we claim that our definitions do not contradict each other.

We suspect that for the Apple/Garageband designers of the vibrato effects that we analyzed, 1 Hz was a simpler, easier choice than a decimal value slightly above 1 Hz for the minimum pitch variation.

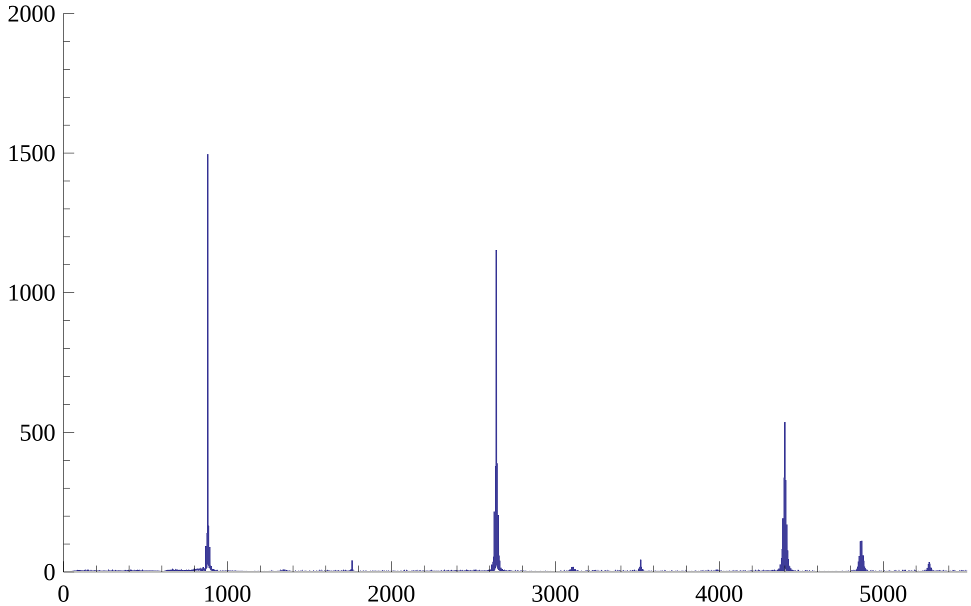


Figure 3.10: Frequency spectrum plot of the *Vibrato Dri* output tone whose input was an 880 Hz pure tone

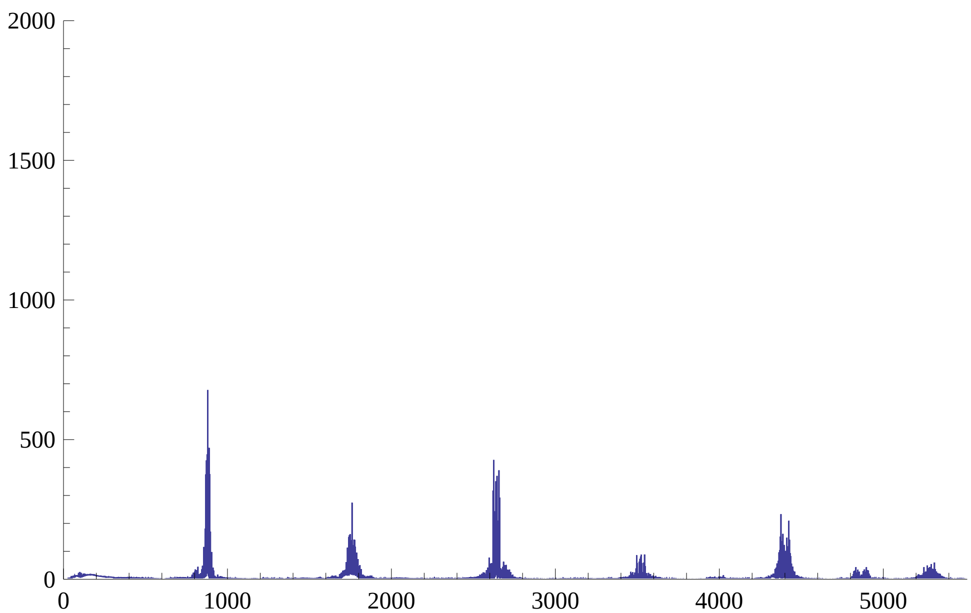


Figure 3.11: Frequency spectrum plot of the *Dark Vibrato* output tone whose input was an 880 Hz pure tone

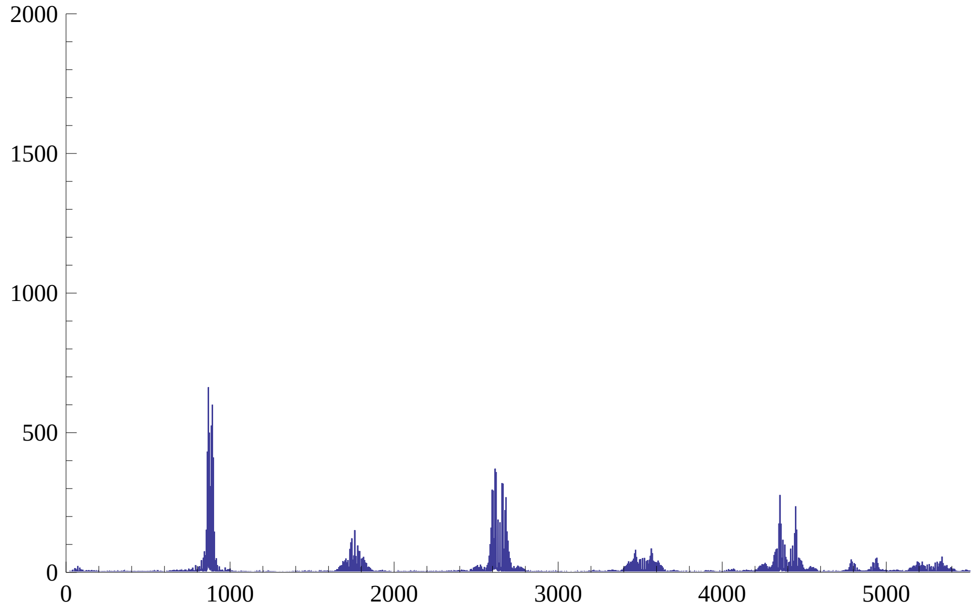


Figure 3.12: Frequency spectrum of the *High Vibrato* output tone whose input was an 880 Hz pure tone

3.2.2 Chorus

The term "chorusing" comes from the fact that each unison voice singing the same part in a choir varies slightly in pitch with respect to the other voices. Chorus is is a modulation effect that splits a signal into two and modulates the pitch of one of the two signals with a low frequency oscillator (or other modulation source). When the the two signals signals are mixed back together, the result is the "chorusing" effect of two parts "singing" together at slightly different pitches. The amount of pitch shift is typically very small. As the amount changes, the two signals go in and out of tune with each other producing a "swooshing" effect. Notice that the effect applied to the "wet" copy is basically vibrato.

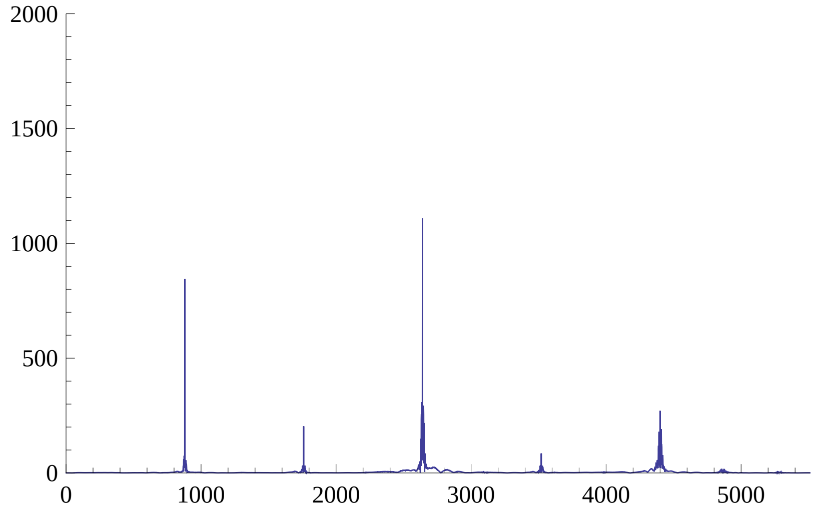


Figure 3.13: The frequency spectrum of the *Chunky Chorus* output tone whose input was an 880 Hz pure tone. Notice that $S_{2640} > S_{880}$.

After observing the frequency spectra of 30 output tones of the three chorus effects that we used in our analysis, we determined that by definition, since there exist m and j such that $\alpha(m,j) < \beta(m,j)$ or $\alpha(m,j) < \eta(m,j)$ or $\alpha(m,j) < \zeta(m,j)$, chorus is not a pure tone effect.[7] See Figure 3.13. Therefore we define it simply as an effect.

Definition. Chorus is an effect $\omega : \mathcal{P} \to \mathcal{S}$ such that

$$\begin{aligned}\omega : \sin(2\pi mt) \mapsto \sum_{j=1}^{\infty} \alpha(m,j)\sin(2\pi jmt) + &\sum_{j=-m+1}^{m-1} \beta(m,j)\sin(2\pi(m+j)t) \\ + &\sum_{j=-m+1}^{m-1} \eta(m,j)\sin(2\pi(3m+j)t) \\ + &\sum_{j=-m+1}^{m-1} \zeta(m,j)\sin(2\pi(5m+j)t)\end{aligned}$$

[7]For instance, the maximum S_n value of the frequency spectrum for the output tone of the Chunky Chorus effect whose input is an 880 Hz pure tone is not S_{880}.

where m is the frequency of the input pure tone,

$$\alpha(m,j) = \begin{cases} f(m,j) & : f(m,j) \geq 0 \\ 0 & : otherwise, \end{cases}$$

$$\beta(m,j) = \begin{cases} g(m,j) & : g(m,j) \geq 0 \\ 0 & : otherwise, \end{cases}$$

$$\eta(m,j) = \begin{cases} h(m,j) & : h(m,j) \geq 0 \\ 0 & : otherwise, \end{cases}$$

$$\zeta(m,j) = \begin{cases} l(m,j) & : l(m,j) \geq 0 \\ 0 & : otherwise, \end{cases}$$

and

$$f(m,1) \approx 0.5 - 0.0005m + 4.0 \times 10^{-7}m^2 - 7.0 \times 10^{-11}m^3,$$

$$f(m,2) \approx 0.04 - 0.0002m + 3.0 \times 10^{-7}m^2 - 1.0 \times 10^{-10}m^3,$$

$$f(m,3) \approx 0.0002m + 1.0 \times 10^{-7}m^2 - 1.0 \times 10^{-10}m^3,$$

$$f(m,5) \approx -0.1 + 0.001m - 2.0 \times 10^{-6}m^2 + 1.25 \times 10^{-9}m^3,$$

$$f(m,j) \approx 0, \forall j \notin \{1,2,3,5\},$$

$$g(m,1) \approx g(m,2) \approx g(m,3) \approx 0.02 - 0.000016m + 4.0 \times 10^{-8}m^2 - 2.4 \times 10^{-11}m^3,$$

$$g(m,j) \approx 0, \forall j \notin \{1,2,3\},$$

$$h(m,1) \approx h(m,2) \approx h(m,3) \approx -0.01 + 0.0001m - 7.0 \times 10^{-8}m^2 + 1.0 \times 10^{-11}m^3,$$

$$h(m,j) \approx 0, \forall j \notin \{0,1,2,3\},$$

$$l(m,1) \approx l(m,2) \approx l(m,3) \approx -0.1 + 0.0009m - 1.6 \times 10^{-6}m^2 + 9.0 \times 10^{-10}m^3,$$

$$l(m,j) \approx 0, \forall j \notin \{0,1,2,3\}.$$

According to our traditional description, chorus splits the original sound signal into two copies and then varies the pitch of one of the two with a small pitch shift before mixing them back together. According to our mathematical definition, the output of chorus is a tone whose most prominent pure tone components are of frequency m, $2m$, $3m$, $5m$, $m \pm j$, $3m \pm j$, and $5m \pm j$, where $j \in \{0, 1, 2, 3\}$. Clearly the component of frequency m represents the dry copy of the input pure tone, and the components of frequency $m \pm j$represent the results of the small pitch shift applied to the wet copy. But what about the components of frequency $2m$, $3m$, $3m \pm j$, $5m$, and $5m \pm j$?

Recall the meaning of the term "chorusing" and imagine constructing a chrous effect. Should the output of chorus be similar to that of vibrato except that only half of the input signal is affected? Perhaps overtones of the fundamental, like those that would be heard singing in a choir, should be included. With a proper adjustment to our traditional definition that would account for these chorusing components, our definitions would agree. Notice that our mathematical definition includes these chorusing overtone components, but our traditional definition does not account for them. We claim that it is our traditional definition that is lacking, not our mathematical definition, and we suppose that either the overtone components are the result of a much larger pitch shift or, as we will see in our analysis of the phaser effect (below), they are the natural result of phase shifting.

3.2.3 Flanger

The name "flanging" comes from the outer edge, or flange, of a tape reel. The flanging effect was produced by playing the same sound simultaneously on two tape decks with a finger applied to the flange of a tape reel of one of the two tape decks. One signal was delayed with respect to the other by different amounts depending on the finger pressure applied. The result when the two signals are mixed together is another "swooshing" effect. While the chorus effect is based on a modulating pitch shift, the flange effect is based on a modulating time delay.

Flanger is a modulation effect that splits a signal into two, then modulates the delay time applied to one of the two signals with a low frequency oscillator (or other modulation source), and then mixes the two signals back together. Because the delay times are very short, typically between 1 and 15 ms, the *individual* delays are not noticeable to the human ear. The result that is heard is a changing phase relationship between the delayed and dry signals after they are mixed back together. The changing phase relationship is due to one copy being delayed by an amount of time that is continually varying.

After observing the frequency spectra of 30 output tones of the three flanger effects that we used in our analysis, we determined that by definition, since for all m and j such that $j \neq 0$, $\alpha(m, 0) > \alpha(m, j)$, flanger is a pure tone effect.

Definition. Flanger is a pure tone effect $\omega : \mathcal{P} \to \mathcal{S}$ such that

$$\omega : \sin(2\pi mt) \mapsto \sum_{j=1}^{m-1} \alpha(m,j)\sin(2\pi(m-j)t) + \sum_{j=1}^{\infty} \alpha(m,j)\sin(2\pi(m+j)t)$$

where m is the frequency of the input pure tone,

$$\alpha(m,j) = \begin{cases} f(m,j) & : f(m,j) \geq 0 \\ 0 & : otherwise \end{cases}$$

and

$$f(m,0) \approx 0.2 - 0.000162m + 1.5 \times 10^{-7}m^2 - 1.75 \times 10^{-11}m^3,$$

$$f(m,1) \approx f(m,2) \approx f(m,3) \approx 0.00017m - 2.6 \times 10^{-7}m^2 + 1.13 \times 10^{-10}m^3,$$

$$f(m,j) \approx 0, \forall j \notin \{0,1,2,3\}.$$

Our traditional and mathematical definitions agree. According to our traditional definition, flanger splits the original sound signal into two 'copies' and then continually varies the delay time applied to one of the two before mixing them back together. According to our mathematical definition, the output of flanger is a tone whose most prominent component is a pure tone of frequency m and whose less prominent components are of frequency $m \pm j$ where $j \in \{1, 2, 3\}$. The most prominent component of all 30 frequency spectra is a pure tone of frequency m. This component represents the dry copy of the input pure tone of frequency m, while the less prominent components of frequency $m \pm j$ represent the results of the modulating delay time applied to the wet copy.

3.2.4 Phaser

Like chorus and flanger, phaser is a modulation effect that splits a signal into two, alters one of the two, and then mixes the two signals back together, producing an effect based upon the relationship betwen the wet and dry signals. Phaser modulates the phase shift applied to one of the two signals by sending it through a series of *all-pass filters* that shift the phase of each harmonic component of the signal by amounts that depend on the frequency of each component, with continuously varying frequency-dependent relationships. The filters are modulated by a low frequency oscillator (or other modulation source) that continually changes the frequency-dependent phase-shifting relationships. The output is a tone with an audible sweeping effect as a series of peaks and troughs in the frequency spectrum vary over time.

After observing the frequency spectra of 30 output tones of the three phaser effects that we used in our analysis, we determined that by definition, since for all m and j such that $j \neq 0$, $\alpha(m,0) > \alpha(m,j)$, phaser is a pure tone effect.

Definition. Phaser is a pure tone effect $\omega : \mathcal{P} \to \mathcal{S}$ such that

$$\omega : \sin(2\pi mt) \mapsto \sum_{j=1}^{m-1} \alpha(m,j)\sin(2\pi(m-j)t) + \sum_{j=1}^{\infty} \alpha(m,j)\sin(2\pi(m+j)t)$$

where m is the frequency of the input pure tone,

$$\alpha(m,j) = \begin{cases} f(m,j) & : f(m,j) \geq 0 \\ 0 & : \textit{otherwise} \end{cases}$$

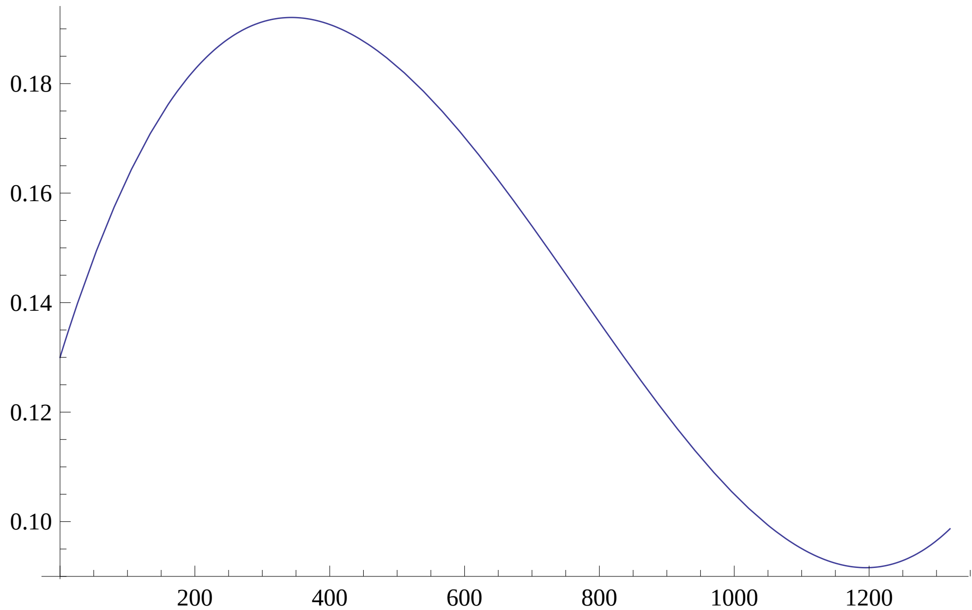
and

$$f(m,0) \approx 0.13 + 0.0004m - 7.5 \times 10^{-7}m^2 + 3.25 \times 10^{-10}m^3,$$

$$f(m,1) \approx f(m,2) \approx f(m,3) \approx 0.00028m - 5.0 \times 10^{-7}m^2 + 2.25 \times 10^{-10}m^3,$$

$$f(m,j) \approx 0, \forall j \notin \{0,1,2,3\}.$$

Our traditional and mathematical definitions agree. According to our traditional definition, phaser splits the original sound signal into two and then continually varies the phase shift of one of the two by continuously varying amounts that depend on the frequency of each component according to a continuously varying frequency-dependent relationship. According to our mathematical definition, the output of phaser is a tone whose most prominent component is a pure tone of frequency m and whose less prominent components are of frequency $m \pm j$ where $j \in \{1,2,3\}$. See Figure 3.14. Clearly the most prominent output component is of frequency m and represents the dry copy of the input pure tone of frequency m, while the less prominent components of frequency $m \pm j$ represent the results of the modulating frequency-dependent phase-shifting relationships. This is very similar to the flanger effect, but the magnitude of



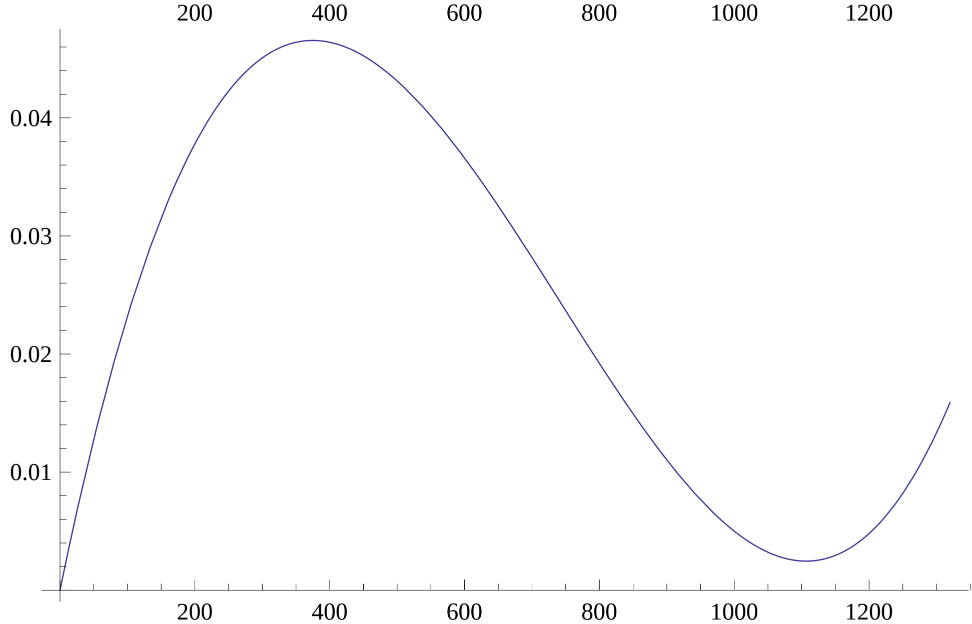


Figure 3.14: Using *Mathematica*'s `Plot` command, these two figures display a plot of the two models we used for the two prominent components, of frequency $n = f$ (above) and of frequency $n = f \pm p$ where p is 1, 2 or 3 (below). Notice that the magnitude of the latter components appears to be greater than the magnitude of the $n = f$ component in the 'neighborhood' of $f = 1200Hz$. However, notice the vertical values are not equally scaled. It is apparent from these two figures that phaser is a pure tone effect, by definition, since for all f, $S_f > S_{f \pm p}$ (with the possible exception of frequencies of approximately $f = 1200Hz$).

the components of the output tones of the flanger and phaser effects have a differing frequency dependent relationship, which of course is expected.

3.2.5 Overdrive

Overdrive is an effect that increases the amplitude of a sound or signal in order to produce distortion at various volume levels. It is the only effect in our analysis that is not a modulation effect. Its name comes from the technique of slightly to moderately overdriving an amplifier to achieve *clipping* through increased *gain*. Gain, a measure of the amplification of a signal, is the ratio of the output level of a signal to the input level of a signal, usually measured in *decibels*, where a ratio of 2 corresponds to an increase of 3 decibels. Clipping occurs when an amplifier is "driven" to produce a signal at a level that is beyond its capabilities. In the context of sound, clipping occurs when an amplifier is driven beyond the maximum level at which it is able to produce a sound wave without distorting its waveform. The result is a "clipped" sound wave (or in electrical terms, sound signal); the peaks of the wave are cut-off. More precisely, the wave/signal is clipped at the maximum capacity of the amplifier.

Some use "overdrive" and "distortion" synonymously, but there is a clear difference between the two. With distortion, although the original structure of the sound wave is still noticeable, it is largely, audibly distorted. Overdrive is typically effected through "soft clipping" where, although there is some noticeable distortion that is slightly "crunchy" or 'fuzzy," the original audible harmonic structure of the sound is nearly intact. The overdrive effect is used to produce a relatively mild level of distortion at a lower volume level. To accomplish this, a *preamp*, accompanied by a gain control, is typically used.

A preamp is an amplifier that receives an electrical signal at a lower voltage level in preparation for further subsequent amplification by the *power amplifier.* The preamp provides the power amplifier with relatively small voltage gain but no significant current gain, while the power amplifier provides the signal with sufficient current gain for output. When the gain is increased enough to achieve clipping of the preamp, the desired overdrive clipping is achieved at a lower volume level, as opposed to the higher voltage/volume needed to clip the power amp. Finally, it should be noted that the overdrive effect is also often accomplished at higher volume levels that clip the power amplifier.

After observing the frequency spectra of 30 output tones of the three overdrive effects that we used in our analysis, we determined that by definition, since there exist m and j such that $\alpha(m, j) > \alpha(m, 1)$, overdrive is not a pure tone effect. Therefore we define it simply as an effect.

Definition. Overdrive is an effect $\omega : \mathcal{P} \to \mathcal{S}$ such that

$$\omega : \sin(2\pi mt) \mapsto \sum_{j=1}^{\infty} \alpha(m, j) \sin(2\pi jmt)$$

where m is the frequency of the input pure tone,

$$\alpha(m, j) = \begin{cases} f(m, j) & : f(m, j) \geq 0 \\ 0 & : otherwise \end{cases}$$

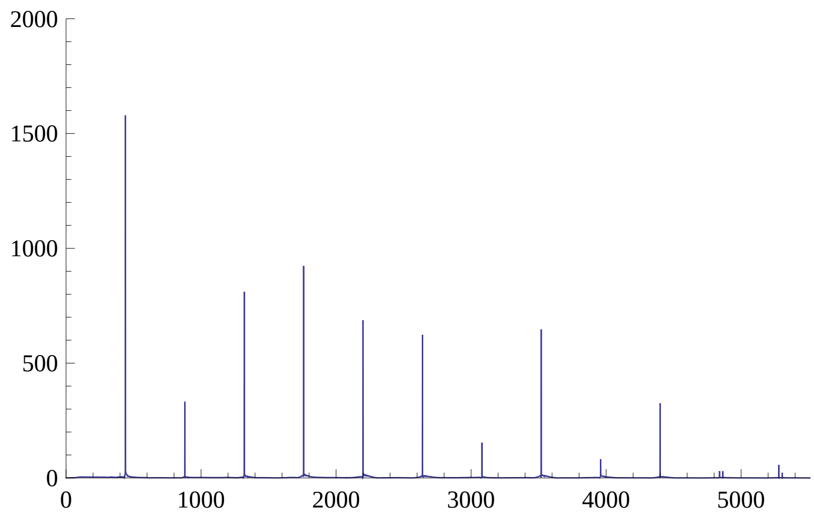


Figure 3.15: In this frequency spectrum of the *Fat Stack* output tone whose input was a 440 Hz pure tone, notice that for all $n, S_{440} > S_n$. Also notice that S_f, S_{3f}, S_{4f}, and S_{5f} are the most prominent components, where $f = 440$ Hz.

and

$$f(m,1) \approx 1.0 - 0.003m + 3.5 \times 10^{-6}m^2 - 1.25 \times 10^{-9}m^3,$$

$$f(m,3) \approx 0.0001m + 7.0 \times 10^{-7}m^2 - 5.0 \times 10^{-10}m^3,$$

$$f(m,4) \approx -0.05 + 0.0005m - 6.0 \times 10^{-7}m^2 + 2.1 \times 10^{-10}m^3,$$

$$f(m,5) \approx -0.6 + 0.005m - 0.00001m^2 + 6.0 \times 10^{-9}m^3,$$

$$f(m,7) \approx -0.00004m + 7.0 \times 10^{-7}m^2 - 8.0 \times 10^{-10}m^3,$$

$$f(m,j) \approx 0, \forall j \notin \{1,3,4,5,7\}.$$

Again our traditional and mathematical definitions agree. According to our traditional definition, although overdrive produces a tone that is typically mildly distorted, the original harmonic structure of the tone is nearly intact. According to our mathematical definition, the output of overdrive is a fairly complex tone whose most prominent components are of frequency m, $3m$, $4m$, $5m$, and $7m$. We conjecture that the noticeable presence of these components is laregly, if not entirely, due to an "across the board" boost in amplitude of all present components, making the otherwise imperceivable components perceivable, noticeable. Hence another result that is expected is that the output tones of overdrive have greater amplitudes than those of the output tones of the other four effects.

Before confirming this, we have already noticed that the average S_n values of the output tones of vibrato are greater than those of chorus, flanger and phaser. We conjecture that this is due to an effect designer choice to boost the ampltiudes of the vibrato output tones in order to make the $m \pm j$ components more audible.

Using the 30 S_f values for each of the five effects, where f is the frequency of the input pure tone, we found the following approximate S_f averages:

1. vibrato: 1360
2. chorus: 1308
3. flanger: 864
4. phaser: 835
5. overdrive: 1494

This data confirms our conjecture.

Note: From the above data, we see that the average S_f values for flanger and phaser are noticeably smaller than the rest. We conjecture that this is because these two effects are constructed with a time delay and a phase shift, respectively. This implies no need to boost any particular components to produce the desired effect. The description of both vibrato and chorus involve frequency shift, so perhaps there was a designer need to boost particular components, resulting in an overall increase in amplitude. Another conjecture is that the averages for flanger and phaser are low because values of only the fundamental were used to compute the averages and the time delay and pitch shift of flanger and phaser shift the strength in the output tone away from the fundamental.

3.3 Analysis limitations

We first note that we analyzed only *GarageBand* effects. Certainly this limited the scope of our analysis, yet we found that these effects served as a fair representation of the standard effects of our study.

Although *Mathematica* is a highly powerful computational software program, it has its limitations. We can not choose whatever *sampling rate* with which we want to work. As is the case with any software program, *Mathematica* can only process so much data in finite time. We will see below that the sampling rate, the number of samples taken per second, directly determines the maximum frequency that we can detect or analyze. The sampling rate (measured in samples per second) is twice the maximum frequency that the discrete Fourier transform can recognize.

We must answer two fundamental questions when analyzing sound wave data. How many frequency values, and which frequency values, can we use? There are two fundamental parameters that must be considered when answering these questions: the length of the time interval on which we collect the data and the sampling rate. For example, if we collect data for 2 seconds, the longest complete oscillation that we can resolve has a period of precisely 2 seconds. The frequency of this wave is $\frac{1}{2}$ cycle per second. This would be the lowest frequency that we could resolve.

To determine which frequency values to use in our analysis, we must first consider the waves that have an integer number of periods on the time interval $[0, I]$. The wave with the largest period has one full period of I units on this interval. Then the lowest frequency associated with the interval $[0, I]$ is $\frac{1}{I}$. We'll denote the fundamental unit of frequency $\Delta\omega = \frac{1}{I}$. All other waves that have an integer[8] number of periods on $[0, I]$ have a frequency that is an integer multiple of $\frac{1}{I}$. These are the frequencies "recognized by" the discrete Fourier transform. The maximum frequency of these waves is $N\Delta\omega = \frac{N}{I}$. We will show, however, that this is *not* the maximum frequency that can be textitaccurately resolved on $[0, I]$ by the discrete Fourer transform.

Notice that an increase in the length of the sample time I results in a longer maximum period over the extended interval $[0, I]$. This results in a decrease in frequency domain because an increase in I results in a decrease in the minimum frequency $\Delta\omega$, which results in a decrease in the maximum frequency $N\Delta\omega = \frac{N}{I}$, of the frequency domain.

Proposition 3.3.1. *The maximum frequency* ***associated with*** *the interval* $[0, I]$ *is* $N\Delta\omega$ *where* $\Delta\omega$ *is the fundamental unit of frequency.*

Proof. The wave with the largest period on $[0, I]$ has period I and frequency $\frac{1}{I}$. The fundamental unit of frequency $\Delta\omega$ is the lowest frequency associated with $[0, I]$. Therefore $\Delta\omega = \frac{1}{I}$. Now notice that since the number of samples along the interval I is N and the maximum number of waves that could be sampled is the wave that has

[8] We are unable to consider frequencies such as 260.5 Hz in our analysis simply because 260.5 is not an integer.

one cycle per sample, the maximum frequency associated with the interval $[0, I]$ is $\frac{N}{I}$. Hence, since $\frac{N}{I} = N\Delta\omega$, the maximum frequency associated with the interval $[0, I]$ is $N\Delta\omega$. $\square$

Proposition 3.3.2. *Suppose that a tone is sampled every Δt units. Then the maximum frequency that can be* ***accurately resolved*** *on the interval $[0, I]$ by the discrete Fourer transform is*

$$\omega_{max} = \frac{1}{2\Delta t} = \frac{1}{2I/N} = \frac{N}{2} \cdot \frac{1}{I} = \omega_{\frac{N}{2}}.$$

Proof. If a tone or signal is sampled every Δt units, then a wave with a period of less than $2\Delta t$ units can not be resolved accurately because the most detail that we can detect in a Δt time interval is a wave that has a peak at one sample time, a valley at the next, and a peak again at the next. This wave has a period of $2\Delta t$ units and a frequency of $\frac{1}{2\Delta t}$ units^{-1}. Since $\Delta t = \frac{I}{N}$ where N is the number of data points,

$$\omega_{max} = \frac{1}{2\Delta t} = \frac{1}{2I/N} = \frac{N}{2} \cdot \frac{1}{I} = \omega_{\frac{N}{2}}.$$

$\square$

Chapter 4: Summary

We consider this a unique and successful study. We have shown how mathematics can be used to analyze effects on sound while using the discrete Fourier transform to compute models for five standard effects on sound. Whether or not this inspires student interest in mathematics is yet to be determined. We know that we have only skimmed the surface of the potential power of the discrete Fourier transform, and there are many more ways in which we can choose to analyze sound waves and effects on sound waves. We are confident that opportunities to share this knowledge will only continue to increase, knowing that we have learned a great deal from these experiences, and that there is much more to learn and to share.

4.1 Educational intent

During my sophomore year in the College of Engineering at Ohio State in 1989, I realized that I wanted to study audio engineeering. I wasn't certain what this might mean, even if I could find a university that might offer such a degree. Ohio State wasn't one of the few on the planet in those years that did. In 1994 I received a degree in music technology at The Evergreen State College in Olympia, Washington, and although I had the opportunity to study such things as MIDI-based composition, digital audio recording and editing, sampling, signal flow, and speaker design, my ultimate desire was to learn anything about any relationship between mathematics and sound. Though it sometimes seemed to be, it was not a secret, how mathematics could be used to design, synthesize, and analyze sound. But even in a music technology based program, there was no real emphasis in applying mathematics to any of the studies. I later received a second bachelors degree in applied mathematics at California State University - Chico, but for the educational opportunity to study and eventually work in a field that most certainly was growing and flourishing exponentially at the time, it was a seemingly futile search. Strangely enough, even though the field has blossomed, the educational opportunities have not.

Bart Snapp is the only individual I have ever known who has been willing and able to direct my attention in such a way that I could learn about the discrete Fourier transform while using it to analyze sound waves. It was Bart's idea to use Fourier analysis to analyze effects on sound. My contention is that people like Bart are much too few and far between, so I certainly hold the intention of using this study as a solid starting point for further educational purposes in this and related fields (other than and including mathematics), for example, computer and electrical engineering. As much as I know that I have finally gained a good understanding of the spectral analysis of sound using the discrete Fourier transform, I know that I want to see this integrated into high school curriculi as, at minimum, an inspirational device to inspire students to want to understand and *use* mathematics. After all, there is never-ending interest in music. Music technology can be at least as interesting. This would be an excellent way to peak student interest in mathematics, mathematical and computer sciences, and engineering.

Bibliography

[1] W. L. Briggs and V. E. Henson, *The DFT: an owner's manual for the discrete Fourier transform*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1995, pp. 16–44, 181-183.

[2] James S. Walker, *Fourier Series*, Encyclopedia of Physical Science and Technology, Elsevier Science Ltd, 2001, pp. 167–183.

[3] Georgi P. Tolstov, *Fourier Series*, General Publishing Company, Toronto, Ontario, 1962, pp. 1–15, 18–19, 32–34, 41–43.

[4] Scott Wilkinson, *Anatomy of a Home Studio: How Everything Really Works, From Microphones to MIDI*, EM Books, Emeryville, CA, 1997, pp. 164–175.