Musical Filters: The Linear Algebra Perspective

Evan Wagner

October 31, 2022

Contents

* * *

CAUTION: This paper includes links to audio files. Some of the sounds are abrasive and can cause discomfort when played loudly. Make sure your audio is turned to a very low setting before you click a link, then adjust up if necessary. It's much better to not hear it the first time than to take a loud sawtooth wave point blank to the eardrum.

If you are extra curious, here are links to the entire Drive folder of audio files and the MATLAB code used to generate example audio and figures.

I would like to thank Profs. Marie Snipes, Adam Lizzi, Judy Holdener, Bob Milnikel, Pamela Pyzza, and Carol Schumacher for their indispensable guidance throughout the capstone process.

1 Introduction: Sound in the Language of Linear Algebra

The core essence of musicianship has been hotly debated for millennia. Is technique the essential ingredient? Is emotion? Is it simply raw talent?

No. The core essence of musicianship is looking cool, and there are few better ways to look cool than turning knobs on your instrument. Confident knob-turning signals competence to one's audience. A musician may practice scales and arpeggios until their eyes cross, but knob-turns truly make or break their performance.

What are musicians really doing to their sound when they turn knobs? That question is the motivation for this paper. It turns out that they are applying linear transformations called **filters** to periodic functions, the properties of which are very useful for the music-making process. I will guide the reader from intuitive concepts of sound to a concrete mathematical understanding of the use of filters in music.

1.1 Mechanics of Sound

Sound waves are vibrations which travel through the air as molecules bump into each other, eventually hitting our ear drums and causing them to vibrate. We can represent this process mathematically with a **signal**, a function that describes some phenomenon. A simple vibration from a single source is called a **pure tone**, and its signal is a periodic function of pressure applied to the eardrum against time. Figure 1 demonstrates how the graph of this function takes a sinusoidal shape.[1] Thus, the signal for a pure tone is $f(t) = a\sin(\frac{2\pi t}{T})$, or perhaps $f(t) = a\cos(\frac{2\pi t}{T})$.

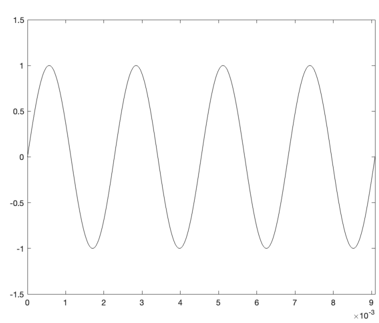


Figure 1: A [▶] pure tone with amplitude $a = 1$ and period $T = \frac{1}{440}$ seconds

Since the function $f(t) = \sin(t)$ has a period of 2π, it is useful to use pure tones whose frequencies are multiples of 2π, which "standardizes" the period. The pure tone $\sin(2\pi(1)t)$ has period $T = 1$, while $\sin(2\pi(\frac{1}{2})t)$ has period $T = 2$, and so on.

A pure tone can be described when two[1] variables are known:[1]

- The **amplitude** a of the pure tone is the vertical distance between 0 and the peak of the sine wave, or the maximum pressure. Amplitude determines how loud the pure tone is, with higher values corresponding to louder sounds.

- The **period** T of the pure tone is how long the sine function takes to complete one cycle. Alternatively, we can use the **frequency** $\nu = \frac{1}{T}$ to describe this aspect of the pure tone, in which case the equation is $f(t) = a\sin(2\pi\nu t)$. Frequency determines the **pitch** we hear from the pure tone. For example, the pitch of a pure tone with frequency $\nu = 440$ Hertz (Hz) is known as A4.

All around us, vibrations generate pure tones of different amplitudes and frequencies.[2] When they reach our ears simultaneously, the vibrations compound each other, and the signal for the pressure on our eardrum is the sum of the pure tone signals.[1] If enough pure tones are happening, their sum can sound like any continuous periodic function imaginable. We can even use pure tones to construct sounds that represent non-continuous functions, such as the square and sawtooth waves shown in Figure 2. How is this possible without breaking the laws of physics? Stay tuned.

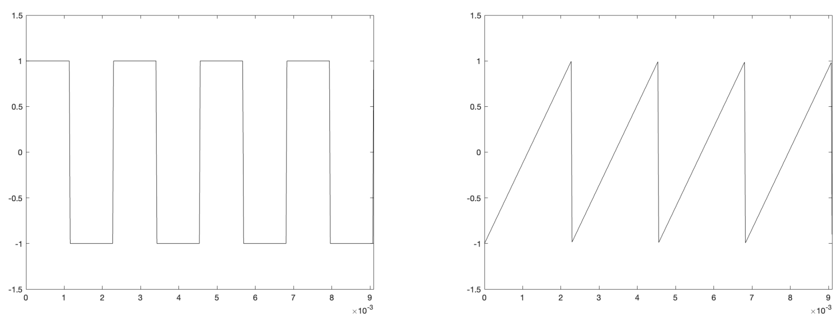


Figure 2: [▶] Square and [▶] sawtooth waves with amplitude $a = 1$ and period $T = \frac{1}{440}$

[1]Strictly speaking, a third variable is necessary to describe a pure tone completely: its **phase** is the location of its peak amplitude. For a few examples, the phase of $\cos(2\pi t)$ is 0, the phase of $\sin(2\pi t)$ is 0.25, and the phase of $\sin(2\pi(t - 0.6))$ is $.25 - (-0.6) = 0.85$.

Since a pure tone is a periodic function, its overall form remains the same as it is repeated over and over. Thus, a change of phase does not affect our auditory experience of a pure tone, and it is not very relevant for the following.

[2]After the mechanical force producing them is done happening, sounds decay in amplitude over time. We get "sustained" periodic tones from sustained vibrations. Wind instruments and traditional organs sustain tones by blowing air at a consistent rate through a pipe, while bowed instruments vibrate strings for the same effect. In contrast, "plucked" tones are produced by short vibrations from a guitar or piano or some such, decaying quickly thereafter. Though plucked tones can be filtered, we will focus on the easier task of filtering sustained tones.

1.2 Sound as a Linear System

A set of pure tones with specified frequencies can be thought of as a vector that represents a sound. It follows that the collection of all sounds has the properties of a vector space. To verify this, we must rigorously define the following operations on vectors of pure tones:

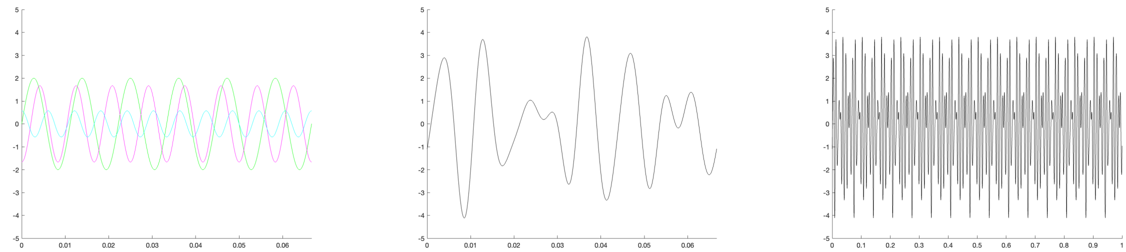
- **Addition:** When two pure tones $\mathbf{v}$ and $\mathbf{w}$ hit the eardrum simultaneously, their forces combine to apply the pressure-over-time function $f(t) = \mathbf{v} + \mathbf{w}$. In other words, we do not experience the two pure tones separately, but as one wave with varying amplitude across the period.
- **Scalar multiplication:** Scaling a pure tone $c \cdot \mathbf{v}$ makes it c times louder.
- **Negation:** Scaling a pure tone $\mathbf{v}$ by -1 reflects the sine or cosine function across the time axis. This negated pure tone has the same frequency and amplitude, but its peak is displaced by $\frac{T}{2}$, so our auditory experience is virtually unchanged.

Finally, let the vector $\mathbf{0}$ correspond to a silent wave with amplitude 0. Under these definitions, the set of pure tones qualifies as a vector space. Table 1 provides intuitive explanations of the associative, commutative, and distributive properties of this pure tone vector space. The mathematical truth of each statement can be verified easily by replacing each vector in the left column with $\sin(\frac{2\pi t}{T})$ or $\cos(\frac{2\pi t}{T})$ for some $T \in \mathbb{R}$.

Axiom	Intuitive Explanation
$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$	Sustaining the pure tone $\mathbf{v}$ and then playing $\mathbf{w}$ produces the same sound as playing $\mathbf{w}$ and then $\mathbf{v}$.
$(\mathbf{u} + \mathbf{v}) + \mathbf{w} =$ $\mathbf{u} + (\mathbf{v} + \mathbf{w})$	Playing the sound that results from combining $\mathbf{u}$ and $\mathbf{v}$, then adding $\mathbf{w}$ is the same as playing $\mathbf{u}$, then the combination of $\mathbf{v}$ and $\mathbf{w}$.
$\mathbf{v} + \mathbf{0} = \mathbf{v}$	Playing $\mathbf{v}$ and then silence is the same as just playing $\mathbf{v}$.
$\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$	If one pure tone is pushing the eardrum one way while the other pushes it the opposite way the same amount, they will entirely cancel each other out.
$1\mathbf{v} = \mathbf{v}$	Amplifying a pure tone by a factor of 1 does not change its loudness.
$c(d\mathbf{v}) = (cd)\mathbf{v}$	Amplifying $\mathbf{v}$ by d, then by c is the same as amplifying $\mathbf{v}$ by cd.
$c(\mathbf{v} + \mathbf{w}) = c\mathbf{v} + c\mathbf{w}$	Amplifying pure tones after combining them is no different than amplifying them individually, then combining them.
$(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$	Clearly, $\mathbf{v}$ has the same frequency as itself, so adding it to itself simply increases the pressure applied to the ear.

Table 1: Axioms of the sound wave vector space

We can think of the sound that results from hearing multiple pure tones simultaneously as a *linear combination of pure tones*. For example, if we hear $2\sin(\frac{2\pi t}{90})$, $-\frac{5}{3}\cos(\frac{2\pi t}{120})$, and $\frac{4}{7}\cos(\frac{2\pi t}{165})$ all at once, the pressure on our ear over time is given by $f(t) = 2\sin(\frac{2\pi t}{90}) - \frac{5}{3}\cos(\frac{2\pi t}{120}) + \frac{4}{7}\cos(\frac{2\pi t}{165})$.[3] Readers who have worked with audio processing software will likely recognize the look of the third plot in Figure 3: amplitude-over-time graphs are common visual displays for audio clips.



Pure tones over $\frac{1}{15}$ seconds | $f(t)$ over $\frac{1}{15}$ seconds | $f(t)$ over 1 second

Figure 3: [▶] Three pure tones and their [▶] linear combination

1.3 Musical Sounds

Listen to the audio clips linked in Figure 3. You may notice that the three pure tones sound pleasant, while their linear combination sounds like garbage. Unfortunately, not every linear combination of pure tones produces a musical note. What property distinguishes a musical note from random noise?

The answer lies in the ratios of pure tone frequencies. A musical note is a linear combination of pure tones whose frequencies are each equal to the lowest frequency, called the **fundamental frequency**, multiplied by an integer.[1] When such a family of frequencies is played simultaneously, our ears hear the pitch of the fundamental frequency. When x unique frequency families are played together, we hear x distinct notes. When the frequencies don't have nice ratios, we hear only muddled noise.

Notes composed of fewer pure tones sound more pleasant and simple. This is a desirable quality for placid, soothing music, but musicians aim to evoke a much wider variety of emotions with their music. The relative amplitudes of the frequencies change the note's emotional characteristics, which together are called the **quality of sound**. We will refer to notes with the same pitch but different sound qualities as different **tones**. As we shall see, filters alter a note's sound quality without changing the frequencies of its pure tones, i.e. they change tone without changing pitch, making them quite useful for musicians.

[3]The period of this function is however long it takes for all the pure tones to return to their starting point simultaneously; that is, 1 over the greatest common divisor of their frequencies. In this case, the three frequencies are 90, 120, and 165, so the period is $\frac{1}{15}$.

2 The Real Fourier Basis

The amplitude-over-time function for a tone with fundamental frequency $\frac{1}{T}$ lives in $C[0, T]$, the space of continuous-time functions with period T. This information is not so helpful for musicians. Changing the function by adding terms or adjusting frequencies might accidentally make the tone a new note or destroy it entirely.

To fulfill their purpose, filters must be applied using some space of tones instead. The previous section gives us a sense of what such a space will look like: tones with fundamental frequency $\frac{1}{T}$ are composed of pure tones with frequencies ranging from $\frac{1}{T}$ up to $\frac{N}{T}$ for some positive integer N.

Definition 2.0.1. The **frequency space** $V_{N,T}$ is the space of tones composed of the fundamental pure tone (with frequency $\frac{1}{T}$) and the first $N - 1$ harmonic pure tones. For instance, $V_{20,\frac{1}{440}}$ contains all possible tones of the note A4 using the first 20 pure tones.

Luckily for us, the terms of the **Fourier series** are a basis for $V_{N,T}$. This section will discuss the Fourier series for real T-periodic functions, which uses sines and cosines and thus connects intuitively with our characterization of pure tones so far. In the next section, the complex Fourier basis will be introduced and used for the remainder of the paper; as will be explained, it is much more convenient for implementing filters.

2.1 The Real Fourier Series

The terms of a Fourier series of order N are the constant function 1 followed by N sines and cosines, for a total of $2N + 1$ terms. For our purposes, the sine and cosine terms are given by $\sin(\frac{2\pi nt}{T})$ and $\cos(\frac{2\pi nt}{T})$, respectively, for each whole number $n \leq N$.[2]

Definition 2.1.1. The N'th-order real Fourier series for a function $f(t)$ with period T is

$$f_N(t) = a_0 + \sum_{n=1}^{N} \left[a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right) \right].$$

Any function $f(t)$ with period T can be approximated by finding the coefficients of its Fourier series. The approximation $f_N(t)$ improves as N increases and thus more terms are added.[2] These coefficients are found using an inner product for continuous T-periodic functions given in Definition 2.1.2.

Definition 2.1.2. The inner product for T-periodic functions

$$\langle f, g\rangle = \frac{1}{T}\int_0^T f(t)g(t)dt.$$

Briefly,[4] an inner product on a linear space V must map each and every pair of elements $f, g \in V$ to some real number c without violating any of the properties outlined in Table 2.[4] Note that we defined our inner product with a normalizing factor, $\frac{1}{T}$, which scales the inner product's output by the length of the period. The importance of doing so will become clear when we prove the orthonormality of the complex Fourier basis.

Property	Proof for $\frac{1}{T}\int_0^T f(t)g(t)dt$ on $V_{N,T}$
$\langle f, g\rangle = \langle g, f\rangle$	Clearly $\frac{1}{T}\int_0^T f(t)g(t)dt = \frac{1}{T}\int_0^T g(t)f(t)dt$.
$\langle f+g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$	$\frac{1}{T}\int_0^T (f(t)+g(t))h(t)dt = \frac{1}{T}\int_0^T f(t)h(t)+g(t)h(t)dt$ $= \frac{1}{T}\int_0^T f(t)h(t)dt + \frac{1}{T}\int_0^T g(t)h(t)dt$.
$\langle cf, g\rangle = c\langle f, g\rangle$	$\frac{1}{T}\int_0^T (cf(t))g(t)dt = c(\frac{1}{T}\int_0^T f(t)g(t)dt)$.
For all nonzero $f \in V$, $\langle f, f\rangle > 0$	Where $f(t)$ is nonzero, $f(t)^2$ is positive. Where $f(t)$ is zero, $f(t)^2$ is zero. So $f(t)^2$ is either positive or zero at any point on T, and thus $\frac{1}{T}\int_0^T f(t)^2dt$ is positive whenever f(t) is nonzero for some t.

Table 2: Properties of an inner product space, proved for $V_{N,T}$

Definition 2.1.3. The coefficients for the real Fourier series of a function $f(t)$ with period T are given by

$$\begin{aligned}
a_0 &= \langle f, 1\rangle = \frac{1}{T}\int_0^T f(t)dt,\\
a_n &= 2\langle f, \cos\left(\frac{2\pi nt}{T}\right)\rangle = \frac{2}{T}\int_0^T f(t)\cos\left(\frac{2\pi nt}{T}\right)dt, \text{ and}\\
b_n &= 2\langle f, \sin\left(\frac{2\pi nt}{T}\right)\rangle = \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2\pi nt}{T}\right)dt.
\end{aligned}$$

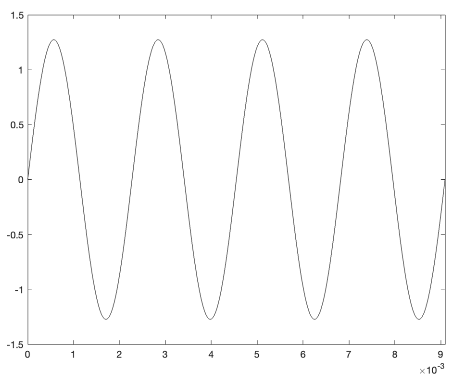
The coefficients $\{a_0, a_n, b_n, 1 \leq n \leq N\}$ provide the amplitudes of a tone's constituent pure tones, so changing these coefficients results in a different tone. We can use the Fourier series not only to approximate periodic functions, but to construct new waveforms from scratch. As N increases, the variety of possible sound qualities greatly increases as more harmonics are added to the sound.

[4]For a more complete review of inner products and inner product spaces, see Bretscher §5.5.

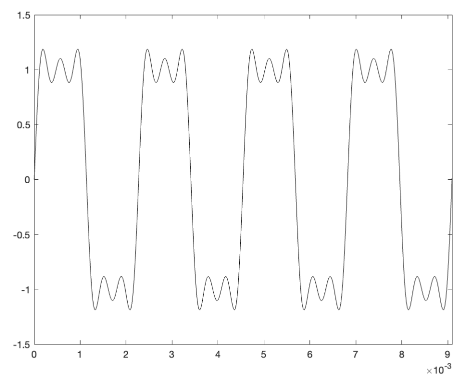
2.2 Example: Approximating the Square Wave

A square wave is a non-continuous function, so the “sound” it represents can’t occur in the real world. However, its Fourier approximation certainly can. Figure 4 shows amplitude-over-time graphs of Fourier approximations of a square wave with pitch A4.

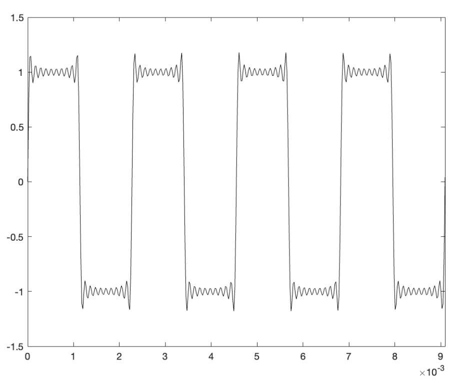
- The first-order approximation is simply the fundamental pure tone of A4, which has period $T = \frac{1}{440}$.
- The 5th-order approximation begins to exhibit a square shape. The function is a tight squiggle which leaps between the amplitudes 1 and -1 every $\frac{T}{2}$ seconds. The sound is somewhat more angular.
- The 20th-order approximation already looks and sounds quite square.
- Once N is sufficiently large, the approximation is visually indistinguishable from the original piecewise square wave.



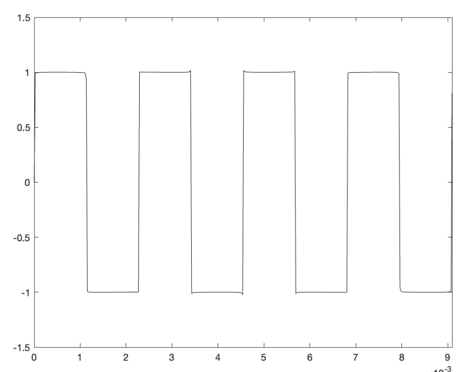
[▶] $N = 1$



[▶] $N = 5$



[▶] $N = 20$



[▶] $N = 1000$

Figure 4: Fourier approximations of the square wave for A4

2.3 Linear Independence of the Real Fourier Terms

The real Fourier terms span $V_{N,T}$ because they are linearly independent. Proving the linear independence of Fourier terms for a general order N is difficult. For now, we will prove linear independence for the case $N = 2$. Later, we will prove rigorously that the complex Fourier terms are linearly independent given any value of N.

Theorem 2.3.1. Consider the second-order Fourier terms $\{1, \cos(\frac{2\pi t}{T}), \cos(\frac{4\pi t}{T}), \sin(\frac{2\pi t}{T}), \sin(\frac{4\pi t}{T})\}$. Let $a_0 + a_1\cos(\frac{2\pi t}{T}) + a_2\cos(\frac{4\pi t}{T}) + b_1\sin(\frac{2\pi t}{T}) + b_2\sin(\frac{4\pi t}{T}) = 0$. Then $a_0 = a_1 = a_2 = b_1 = b_2 = 0$. In other words, the second-order Fourier terms are linearly independent.

Proof. Let $t_1, t_2, t_3, t_4, t_5 \in [0, T]$. Then we have the following system of equations:

$$\begin{aligned}
a_0 + a_1\cos(\frac{2\pi t_1}{T}) + a_2\cos(\frac{4\pi t_1}{T}) + b_1\sin(\frac{2\pi t_1}{T}) + b_2\sin(\frac{4\pi t_1}{T}) &= 0\\
a_0 + a_1\cos(\frac{2\pi t_2}{T}) + a_2\cos(\frac{4\pi t_2}{T}) + b_1\sin(\frac{2\pi t_2}{T}) + b_2\sin(\frac{4\pi t_2}{T}) &= 0\\
a_0 + a_1\cos(\frac{2\pi t_3}{T}) + a_2\cos(\frac{4\pi t_3}{T}) + b_1\sin(\frac{2\pi t_3}{T}) + b_2\sin(\frac{4\pi t_3}{T}) &= 0\\
a_0 + a_1\cos(\frac{2\pi t_4}{T}) + a_2\cos(\frac{4\pi t_4}{T}) + b_1\sin(\frac{2\pi t_4}{T}) + b_2\sin(\frac{4\pi t_4}{T}) &= 0\\
a_0 + a_1\cos(\frac{2\pi t_5}{T}) + a_2\cos(\frac{4\pi t_5}{T}) + b_1\sin(\frac{2\pi t_5}{T}) + b_2\sin(\frac{4\pi t_5}{T}) &= 0.
\end{aligned}$$

Put in matrix form, this system is

$$\begin{bmatrix} 1 & \cos(\frac{2\pi t_1}{T}) & \cos(\frac{4\pi t_1}{T}) & \sin(\frac{2\pi t_1}{T}) & \sin(\frac{4\pi t_1}{T}) \\ 1 & \cos(\frac{2\pi t_2}{T}) & \cos(\frac{4\pi t_2}{T}) & \sin(\frac{2\pi t_2}{T}) & \sin(\frac{4\pi t_2}{T}) \\ 1 & \cos(\frac{2\pi t_3}{T}) & \cos(\frac{4\pi t_3}{T}) & \sin(\frac{2\pi t_3}{T}) & \sin(\frac{4\pi t_3}{T}) \\ 1 & \cos(\frac{2\pi t_4}{T}) & \cos(\frac{4\pi t_4}{T}) & \sin(\frac{2\pi t_4}{T}) & \sin(\frac{4\pi t_4}{T}) \\ 1 & \cos(\frac{2\pi t_5}{T}) & \cos(\frac{4\pi t_5}{T}) & \sin(\frac{2\pi t_5}{T}) & \sin(\frac{4\pi t_5}{T}) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} = 0.$$

If we can find 5 unique values of t such that the row-reduced form of the square matrix above is the identity matrix, then all 5 coefficients must be 0.

Let $t_1 = 0$, $t_2 = \frac{T}{8}$, $t_3 = \frac{T}{4}$, $t_4 = \frac{T}{2}$, and $t_5 = \frac{3T}{4}$. Software tells us that

$$\operatorname{rref}\left(\begin{bmatrix} 1 & \cos(0) & \cos(0) & \sin(0) & \sin(0) \\ 1 & \cos(\frac{\pi}{4}) & \cos(\frac{\pi}{2}) & \sin(\frac{\pi}{4}) & \sin(\frac{\pi}{2}) \\ 1 & \cos(\frac{\pi}{2}) & \cos(\pi) & \sin(\frac{\pi}{2}) & \sin(\pi) \\ 1 & \cos(\pi) & \cos(2\pi) & \sin(\pi) & \sin(2\pi) \\ 1 & \cos(\frac{3\pi}{2}) & \cos(3\pi) & \sin(\frac{3\pi}{2}) & \sin(3\pi) \end{bmatrix}\right) = \operatorname{rref}\left(\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 1 \\ 1 & 0 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 \end{bmatrix}\right) = I.$$

∎

We can view the process of approximating $f(t)$ as a change of basis from the continuous-time domain, where coordinates are given by values of t, to the frequency domain $V_{N,T}$, where coordinates are given by the Fourier coefficients found using the formulas in Definition 2.1.3 (i.e. the amplitudes of the pure tones that constitute $f(t)$).

Definition 2.3.2. The N'th-order real Fourier basis is the set of pure tones with frequency $\frac{n}{T}$ where $0 \leq n \leq N$, notated as

$$\mathcal{D}_{N,T} = \left\{ \begin{array}{l} 1, \cos(\frac{2\pi t}{T}), \cos(\frac{2\pi 2t}{T}), ..., \cos(\frac{2\pi Nt}{T}), \\ \sin(\frac{2\pi t}{T}), \sin(\frac{2\pi 2t}{T}), ..., \sin(\frac{2\pi Nt}{T}) \end{array} \right\}. [2]$$

3 The Complex Fourier Basis

Because the real Fourier series employs sines and cosines, its connection to sound waves is clear and intuitive. However, it is impractical to work with. For one thing, the coefficients can only be derived by calculating three different inner products. For another, the terms of the real Fourier basis $\mathcal{D}_{N,T}$ are given in a peculiar order and cannot be elegantly made into a vector. Finally, though $\mathcal{D}_{N,T}$ is orthogonal,[5] it is not orthonormal.

We solve all these problems by using the complex Fourier basis $\mathcal{F}_{N,T}$. Its coefficients can be calculated with one inner product, and its terms neatly correspond to a vector of length $2N+1$. Most importantly, as I will prove, $\mathcal{F}_{N,T}$ is orthonormal under the complex inner product defined in 3.0.4.

Definition 3.0.1. The N'th-order complex Fourier basis is the set of pure tones with frequency $\frac{n}{T}$ where $-N \leq n \leq N$, notated as

$$\mathcal{F}_{N,T} = \left\{ \begin{array}{l} e^{-\frac{2\pi iNt}{T}}, e^{-\frac{2\pi i(N-1)t}{T}}, ..., e^{-\frac{2\pi it}{T}}, \\ 1, e^{\frac{2\pi it}{T}}, ..., e^{\frac{2\pi i(N-1)t}{T}}, e^{\frac{2\pi iNt}{T}} \end{array} \right\}. [2]$$

To show that $\mathcal{F}_{N,T}$ spans $V_{N,T}$ just as $\mathcal{D}_{N,T}$ does, we take advantage of Euler's Formulas, swapping in $\frac{2\pi nt}{T}$ as the continuous variable:

Definition 3.0.2. Euler's Formulas are

$$e^{\frac{2\pi int}{T}} = \cos\left(\frac{2\pi nt}{T}\right) + i\sin\left(\frac{2\pi nt}{T}\right) \qquad \text{and} \qquad e^{-\frac{2\pi int}{T}} = \cos\left(\frac{2\pi nt}{T}\right) - i\sin\left(\frac{2\pi nt}{T}\right).$$

[5]In the interest of getting to the good stuff, the orthogonality of $\mathcal{D}_{N,T}$ will not be proved in this paper. It would proceed much like the upcoming Lemma 3.1.1.

These equations allow us to turn the cosine terms of the real Fourier series into complex exponential terms:

$$\cos(\tfrac{2\pi nt}{T}) = e^{\frac{2\pi int}{T}} - i\sin(\tfrac{2\pi nt}{T}) = e^{\frac{2\pi int}{T}} + e^{-\frac{2\pi int}{T}} - \cos(\tfrac{2\pi nt}{T})$$
$$2\cos(\tfrac{2\pi nt}{T}) = e^{\frac{2\pi int}{T}} + e^{-\frac{2\pi int}{T}}$$
$$\cos(\tfrac{2\pi nt}{T}) = \tfrac{1}{2}(e^{\frac{2\pi int}{T}} + e^{-\frac{2\pi int}{T}})$$

The sine terms become complex exponential terms as well:

$$\sin(\tfrac{2\pi nt}{T}) = \tfrac{1}{i}(e^{\frac{2\pi int}{T}} - \cos(\tfrac{2\pi nt}{T})) = \tfrac{1}{i}(e^{\frac{2\pi int}{T}} - e^{-\frac{2\pi int}{T}}) - \sin(\tfrac{2\pi nt}{T})$$
$$2\sin(\tfrac{2\pi nt}{T}) = \tfrac{1}{i}(e^{\frac{2\pi int}{T}} + e^{-\frac{2\pi int}{T}})$$
$$\sin(\tfrac{2\pi nt}{T}) = \tfrac{1}{2i}(e^{\frac{2\pi int}{T}} + e^{-\frac{2\pi int}{T}})$$

Therefore, since $\mathcal{D}_{N,T}$ spans $V_{N,T}$, $\mathcal{F}_{N,T}$ must also span $V_{N,T}$. Definition 3.0.3 gives the terms of the N'th-order complex Fourier series for some T-periodic function.

Definition 3.0.3. The N'th-order complex Fourier series for a function $f(t)$ with period T is

$$f_N(t) = \sum_{n=-N}^{N} c_n e^{\frac{2\pi int}{T}}.[2]$$

Finding the complex coefficients c_n is a much simpler task than finding real Fourier coefficients. To do so, we employ an inner product on the space of complex T-periodic functions, which uses the complex conjugate $\overline{g(t)}$ of the function $g(t)$. Taking the complex conjugate of a function leaves its real part alone while flipping the sign of its imaginary part. For complex exponentials, the sign of the exponent is flipped, i.e. $e^{\frac{2\pi int}{T}}$ becomes $e^{-\frac{2\pi int}{T}}$.

Definition 3.0.4. The complex inner product for T-periodic functions

$$\langle f, g\rangle = \frac{1}{T}\int_0^T f(t)\overline{g(t)}dt$$

Definition 3.0.5. The coefficients of the complex Fourier series of a function $f(t)$ with period T are given by

$$c_n = \langle f, e^{\frac{2\pi int}{T}}\rangle = \frac{1}{T}\int_0^T f(t)e^{-\frac{2\pi int}{T}}dt.$$

We established earlier that Fourier approximation is a change of basis from the continuous-time domain to the frequency domain. Now we have a way to change the basis in the other direction: simply use the components of the amplitude vector $\{c_{-N}, c_{-N+1}, ..., c_{-1}, c_0, c_1, ..., c_{N-1}, c_N\}$ as the coefficients of $\mathcal{F}_{N,T}$ to obtain $f_N(t)$.

3.1 Orthonormality of the Complex Fourier Basis

Proving that a basis is orthonormal takes two steps:

1. Prove that each vector in the basis is orthogonal to the rest. To do so, show that the inner product of any two different basis vectors is 0.
2. Prove that each vector in the basis is normal. To do so, show that the inner product of any basis vector with itself is 1.

Taken together, Lemmas 3.1.1 and 3.1.2 prove that the complex Fourier basis is indeed orthonormal.

Lemma 3.1.1. The inner product of any two unique vectors in $\mathcal{F}_{N,T}$ is 0.

Proof. Consider arbitrary $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{F}_{N,T}$ such that $\mathbf{v}_1 \neq \mathbf{v}_2$. Then $\mathbf{v}_1 = e^{\frac{2\pi i n_1 t}{T}}$ and $\mathbf{v}_2 = e^{\frac{2\pi i n_2 t}{T}}$ for some $-N \leq n_1, n_2 \leq N$ such that $n_1 \neq n_2$. We see that

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \frac{1}{T}\int_0^T e^{\frac{2\pi i n_1 t}{T}} \overline{e^{\frac{2\pi i n_2 t}{T}}} dt = \frac{1}{T}\int_0^T e^{\frac{2\pi i (n_1 - n_2) t}{T}} dt = \frac{e^{\frac{2\pi i (n_1 - n_2) t}{T}}}{2\pi i (n_1 - n_2) T}\Bigg]_0^T$$

$$= \frac{e^{2\pi i (n_1 - n_2)} - 1}{2\pi i (n_1 - n_2) T} = \frac{cos(2\pi(n_1 - n_2)) + i \sin(2\pi(n_1 - n_2)) - 1}{2\pi i (n_1 - n_2) T} = \frac{1 + i \cdot 0 - 1}{2\pi i (n_1 - n_2) T} = 0.$$

∎

Lemma 3.1.2. The inner product of any vector in $\mathcal{F}_{N,T}$ with itself is 1.

Proof. Consider arbitrary $\mathbf{v}_0 \in \mathcal{F}_{N,T}$. Then $\mathbf{v}_0 = e^{\frac{2\pi i n_0 t}{T}}$ for some $-N \leq n_0 \leq N$. We see that

$$\langle \mathbf{v}_0, \mathbf{v}_0 \rangle = \frac{1}{T}\int_0^T e^{\frac{2\pi i n_0 t}{T}} \overline{e^{\frac{2\pi i n_0 t}{T}}} dt = \frac{1}{T}\int_0^T e^{\frac{2\pi i (n_0 - n_0) t}{T}} dt$$

$$= \frac{1}{T}\int_0^T e^0 dt = \frac{1}{T}\int_0^T 1 dt = \frac{t}{T}\Bigg]_0^T = 1.^6$$

∎

As promised, we will use $\mathcal{F}_{N,T}$'s orthonormality to prove its linear independence.

Theorem 3.1.3. Let arbitrary $e^{\frac{2\pi i n_1 t}{T}}, e^{\frac{2\pi i n_2 t}{T}}, \ldots, e^{\frac{2\pi i n_k t}{T}} \in \mathcal{F}_{N,T}$ for some $k \leq N$. Assume $c_1 e^{\frac{2\pi i n_1 t}{T}} + c_2 e^{\frac{2\pi i n_2 t}{T}} + \cdots + c_k e^{\frac{2\pi i n_k t}{T}} = 0$. Then $c_1 = c_2 = \cdots = c_k = 0$. In other words, $e^{\frac{2\pi i n_1 t}{T}}, e^{\frac{2\pi i n_2 t}{T}}, \ldots, e^{\frac{2\pi i n_k t}{T}}$ are linearly independent.

Proof. Take the inner product of both sides of $c_1 e^{\frac{2\pi i n_1 t}{T}} + c_2 e^{\frac{2\pi i n_2 t}{T}} + \cdots + c_k e^{\frac{2\pi i n_k t}{T}} = 0$ with $e^{\frac{2\pi i n_j t}{T}}$ for some arbitrary $j \in \{1, 2, \ldots, k\}$.

On the left, we see that $\langle c_1 e^{\frac{2\pi i n_1 t}{T}} + c_2 e^{\frac{2\pi i n_2 t}{T}} + \cdots + c_k e^{\frac{2\pi i n_k t}{T}}, e^{\frac{2\pi i n_0 t}{T}} \rangle = c_1 \langle e^{\frac{2\pi i n_1 t}{T}}, e^{\frac{2\pi i n_j t}{T}} \rangle + c_2 \langle e^{\frac{2\pi i n_2 t}{T}}, e^{\frac{2\pi i n_j t}{T}} \rangle + \cdots + c_j \langle e^{\frac{2\pi i n_j t}{T}}, e^{\frac{2\pi i n_j t}{T}} \rangle + \cdots + c_k \langle e^{\frac{2\pi i n_k t}{T}}, e^{\frac{2\pi i n_j t}{T}} \rangle$. Since $\mathcal{F}_{N,T}$ is orthonormal, this expression is equal to $c_1 \cdot 0 + c_2 \cdot 0 + \cdots + c_j \cdot 1 + \cdots + c_k \cdot 0 = c_j$.

On the right, we see that $\langle 0, e^{\frac{2\pi i n_j t}{T}} \rangle = 0$. Then $c_j = 0$, and since c_j is arbitrary, all the coefficients equal 0. ∎

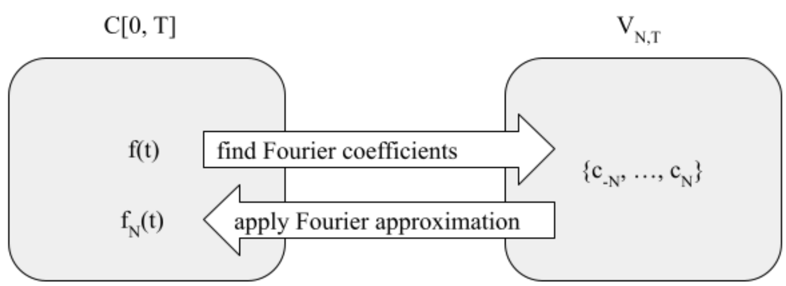


Figure 5: Moving between continuous time and the frequency space

Since each of the pure tones in $\mathcal{F}_{N,T}$ is linearly independent from the rest, the frequency information of a tone f is completely preserved when taken from the time domain to the frequency domain and back again. Thus, a musician can obtain Fourier coefficients for the T-periodic sound function of their choice, modify them, and return the new function to continuous time without eliminating or muddling any of the harmonics, nor changing the pitch.

Meanwhile, since each pure tone in $\mathcal{F}_{N,T}$ is a unit vector, its amplitude is unaffected by changes in coordinates. A musician can alter the amplitudes of harmonic frequencies to their heart's content without worrying about calculating some denominator for each Fourier coefficient when changing bases.

Now that we have a nice and ergonomic setting to mess around with sound quality, we can finally start discussing how the messing-around is actually accomplished. At long last, the next section mathematically defines filters that operate on continuous-time functions. In this setting, they are called **analog filters**, while **digital filters** (discussed later) operate on discrete vectors of sampled amplitudes.

4 Analog Filters

Say we have a continuous-time function that represents some tone. The complex Fourier coefficients of that function are the amplitudes of the pure tones that constitute the tone. To turn the tone into a new one while preserving the pitch, all we must do is find those amplitudes (i.e. apply the Fourier approximation), change them, and go back to a continuous-time function (i.e. set the new amplitudes as the coefficients of $\mathcal{F}_{N,T}$). This process is diagrammed in Figure 6.

A linear function $s : V_{N,T} \to V_{N,T}$ will do what we want when *all the pure tones in* $\mathcal{F}_{N,T}$ *are eigenvectors of* s. Recall that putting an eigenvector through s returns that same eigenvector, but scaled by some constant (also known as its eigenvalue). From this, a definition for analog filters arises:

Definition 4.0.1. An analog filter is a linear function $s : V_{N,T} \to V_{N,T}$ such that, for all pure tones $e^{\frac{2\pi int}{T}}$, $-N \leq n \leq N$,

$$s(e^{\frac{2\pi int}{T}}) = \lambda_s\left(\frac{n}{T}\right)e^{\frac{2\pi int}{T}}$$

where $\lambda_s(\frac{n}{T})$ is the **frequency response** of s with respect to frequency $\frac{n}{T}$.[2]

Applying the filter s to each pure tone in the N'th-order complex Fourier series and then summing the results is equivalent to multiplying each element of the frequency response vector $\lambda_s = \{\lambda_s(0), \lambda_s(\frac{1}{T}), \lambda_s(\frac{2}{T}), ..., \lambda_s(\frac{n}{T}), ..., \lambda_s(\frac{N-1}{T})\}$ by $\mathcal{F}_{N,T}$.[7]

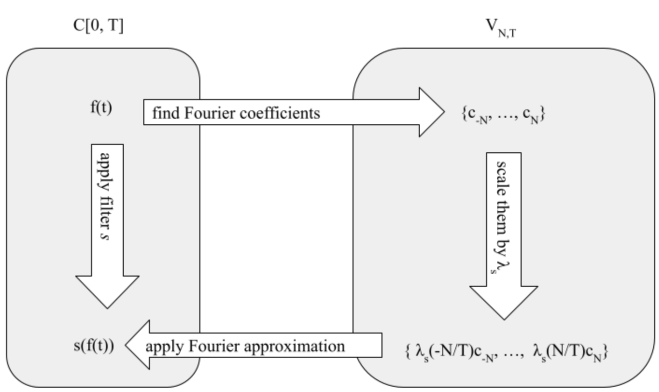


Figure 6: Moving within and between continuous time and the frequency space

[7]Note that λ_s gives the coordinates of s in $V_{N,T}$ under the basis $\mathcal{F}_{N,T}$, meaning $s \in V_{N,T}$ and so must itself be some tone of the note with fundamental frequency $\frac{1}{T}$.

4.1 Properties of Analog Filters

Filters have a few key properties that allow them to aid music production.[2]

- *The composition of multiple filters is itself a filter.* In other words, for any sequence of filters applied one after the other, one can find a single filter that is equivalent to the entire sequence.
  - Additionally, *the frequency response for the composition of multiple filters is equal to the product of the frequency responses of the constituent filters.* That means for arbitrary filters s_1 and s_2, $\lambda_{s_1 s_2}(\frac{n}{T}) = \lambda_{s_1}(\frac{n}{T})\lambda_{s_2}(\frac{n}{T})$.
- *Filters commute.* In other words, the order of application does not change the sound that is ultimately produced. Also, if we apply one filter, apply a second, then remove the first, we get the same sound as if we had only applied the second.

Theorem 4.1.1. The set of filters has the above properties and therefore qualifies as a commutative algebra.

Proof. Consider arbitrary filters s_1 and s_2 with frequency responses $\lambda_{s_1}(\frac{n}{T})$ and $\lambda_{s_2}(\frac{n}{T})$. We see that

$$\begin{aligned} s_1\left(s_2\left(e^{\frac{2\pi int}{T}}\right)\right) = s_1\left(\lambda_{s_2}\left(\frac{n}{T}\right)e^{\frac{2\pi int}{T}}\right) = \lambda_{s_2}\left(\frac{n}{T}\right)\left(s_1\left(e^{\frac{2\pi int}{T}}\right)\right) \\ = \lambda_{s_2}\left(\frac{n}{T}\right)\lambda_{s_1}\left(\frac{n}{T}\right)\left(e^{\frac{2\pi int}{T}}\right) \\ = \lambda_{s_2}\left(\frac{n}{T}\right)\left(\lambda_{s_1}\left(\frac{n}{T}\right)e^{\frac{2\pi int}{T}}\right) = \lambda_{s_2}\left(\frac{n}{T}\right)s_1\left(e^{\frac{2\pi int}{T}}\right) = s_2\left(s_1\left(e^{\frac{2\pi int}{T}}\right)\right). \end{aligned}$$

Therefore, s_1 and s_2 commute. Additionally, since the composition of s_1 and s_2 is equal to $\lambda_{s_2}(\frac{n}{T})\lambda_{s_1}(\frac{n}{T})(e^{\frac{2\pi int}{T}})$, it is a filter with frequency response $\lambda_{s_1}(\frac{n}{T})\lambda_{s_2}(\frac{n}{T})$. ∎

4.2 Example: Low-Pass and High-Pass Filters

Two of the most commonly-used types of filter are called **low-pass filters (LPFs)** and **high-pass filters (HPFs)**. Low-pass filters remove pure tones with frequencies above some threshold and allow the frequencies below it to pass through. High-pass filters do the opposite.

To see how this works, consider the signal $f(t) = e^{\frac{2\pi i(1)t}{T}} + e^{\frac{2\pi i(2)t}{T}} + e^{\frac{2\pi i(3)t}{T}} + e^{\frac{2\pi i(4)t}{T}} + e^{\frac{2\pi i(5)t}{T}} + e^{\frac{2\pi i(6)t}{T}}$. This function is the sum of the pure tones with frequency $\frac{n}{T}$, $1 \leq n \leq 6$, and its Fourier coefficients are clearly $\{0, 1, 1, 1, 1, 1, 1\}$. Figure 7 shows an amplitude-over-time graph of the tone represented by this signal and its constituent pure tones.

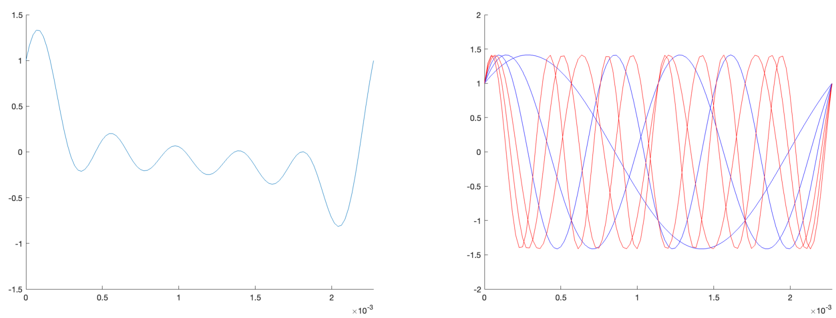
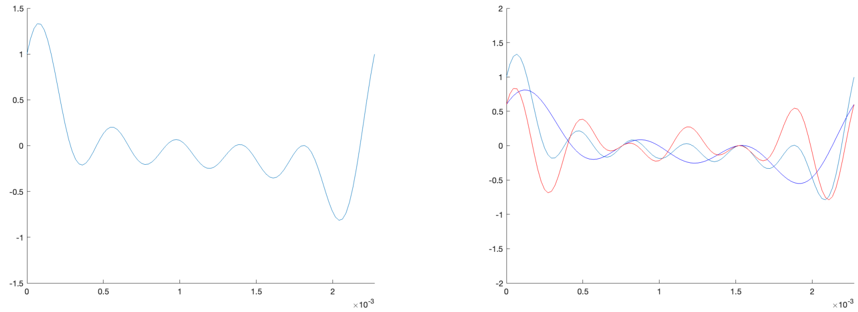


Figure 7: A sixth-order tone and its constituent pure tones

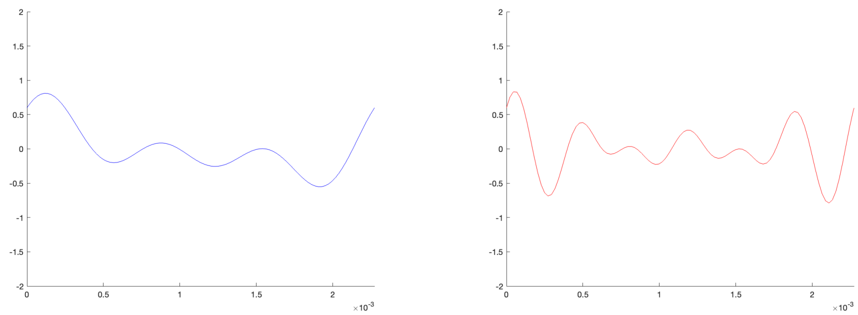
When the LPF is applied, the amplitudes of the three higher pure tones ($n \in \{1, 2, 3\}$, shown in red) are set to 0, while the three lower pure tones ($n \in \{4, 5, 6\}$, shown in blue) are left alone. The HPF removes the lower and keeps the higher.

$$\lambda_{LPF} = \{0, 1, 1, 1, 0, 0, 0\} \qquad \lambda_{HPF} = \{0, 0, 0, 0, 1, 1, 1\}$$

Listen to the audio files linked in Figures 7 and 8 to hear the effect of these two filters on our sixth-order pure tone. The LPF produces a more gentle and well-rounded tone, while the HPF produces a tinnier, more aggressive tone.



[▶] Unfiltered tone



[▶] Tone after LPF is applied [▶] Tone after HPF is applied

Figure 8: The LPF and HPF applied to a basic sixth-order tone

5 The Discrete Fourier Transform

Analog filters are useful for conceptualization, but one more consideration is necessary to make filtering practical. Today's musicians work in the realm of digital sound, where discrete units of digital information must somehow be substituted for the continuous function $f(t)$. To do so, we can make N measurements of the amplitude $f(t)$ uniformly across the period T, starting at $t = 0$ and ending at $t = \frac{(N-1)T}{N}$.[8] These samples form the N terms of a **sample vector**.

Example 5.0.1. The sample vector for some T-periodic function $f(t)$ is

$$\mathbf{v} = \begin{bmatrix} f(0) \\ f(\frac{T}{N}) \\ f(\frac{2T}{N}) \\ \vdots \\ f(\frac{kT}{N}) \\ \vdots \\ f(\frac{(N-2)T}{N}) \\ f(\frac{(N-1)T}{N}) \end{bmatrix}.$$

[8]Note that $f(\frac{NT}{N}) = f(T) = f(0)$, so $t = \frac{(N-1)T}{N}$ is the last meaningful timestamp to sample.

In order to bring a sample vector into the frequency domain and apply filters to it, we must perform some sort of **discrete Fourier transform (DFT)**. We are looking for an $N \times N$ matrix that a vector of N samples can be multiplied by to produce a vector of N complex Fourier coefficients, also known as a **frequency response vector**.

First, we find the sample vector for the pure tone $f(t) = e^{\frac{2\pi i n t}{T}}$ and normalize it with $\frac{1}{\sqrt{N}}$:

Example 5.0.2. The sample vector for the n'th pure tone in $\mathcal{F}_{N,T}$ is

$$\phi_n = \frac{1}{\sqrt{N}}\{1, e^{\frac{2\pi i n}{N}}, e^{\frac{2\pi i 2n}{N}}, ..., e^{\frac{2\pi i k n}{N}}, ..., e^{\frac{2\pi i (N-2)n}{N}}, e^{\frac{2\pi i (N-1)n}{N}}\}.$$

Together, the pure tone vectors $\{\phi_n\}_{n=0}^{N-1}$ form the N**-point Fourier basis** $\mathcal{F}_N$. To prove than $\mathcal{F}_N$ is an orthonormal basis for $\mathbb{C}^n$, we use the complex Euclidean inner product:

Definition 5.0.3. The complex Euclidean inner product is

$$\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\mathbf{w}}^T \mathbf{v} = \sum_{k=0}^{N-1} v_k \overline{w_k} \qquad \text{where } \overline{\mathbf{w}}^T \text{ is the conjugate transpose of } \mathbf{w}.[3]$$

Lemma 5.0.4. The Euclidean inner product of any two unique vectors in $\mathcal{F}_N$ is 0.

Proof. Consider arbitrary $\phi_{n_1}, \phi_{n_2} \in \mathcal{F}_{N,T}$ such that $\phi_{n_1} \neq \phi_{n_2}$. Then $\phi_{n_1} = e^{\frac{2\pi i n_1 k}{N}}$ and $\phi_{n_2} = e^{\frac{2\pi i n_2 k}{N}}$ for some $0 \leq n_1, n_2 \leq N-1$ such that $n_1 \neq n_2$. We see that

$$\langle \phi_{n_1}, \phi_{n_2} \rangle = (\frac{1}{\sqrt{N}})^2 \sum_{k=0}^{N-1} e^{\frac{2\pi i n_1 k}{N}} \overline{e^{\frac{2\pi i n_2 k}{N}}} = \frac{1}{N} \sum_{k=0}^{N-1} e^{\frac{2\pi i (n_1 - n_2) k}{N}}.$$

Since $\sum_{k=0}^{N-1} e^{\frac{2\pi i (n_1 - n_2) k}{N}}$ is a geometric series with common ratio $e^{\frac{2\pi i (n_1 - n_2)}{N}}$, we see that

$$\frac{1}{N} \sum_{k=0}^{N-1} e^{\frac{2\pi i (n_1 - n_2) k}{N}} = \frac{1}{N} \left(\frac{1 - e^{\frac{2\pi i (n_1 - n_2) N}{N}}}{1 - e^{\frac{2\pi i (n_1 - n_2)}{N}}} \right) = \frac{1}{N} \left(\frac{1 - e^{2\pi i (n_1 - n_2)}}{1 - e^{\frac{2\pi i (n_1 - n_2)}{N}}} \right)$$
$$= \frac{1}{N} \left(\frac{1 - (\cos(2\pi(n_1 - n_2)) + i \sin(2\pi(n_1 - n_2)))}{1 - e^{\frac{2\pi i (n_1 - n_2)}{N}}} \right) = \frac{1}{N} \left(\frac{1 - (1 + 0)}{1 - e^{\frac{2\pi i (n_1 - n_2)}{N}}} \right) = 0.$$

∎

Lemma 5.0.5. The Euclidean inner product of any vector in $\mathcal{F}_N$ with itself is 1.

Proof. Consider arbitrary $\phi_{n_0} \in \mathcal{F}_{N,T}$. Then $\phi_{n_0} = e^{\frac{2\pi i n_0 k}{N}}$ for some $0 \leq n_0 \leq N-1$. We see that

$$\begin{aligned}\langle \phi_{n_0}, \phi_{n_0} \rangle = (\frac{1}{\sqrt{N}})^2 \sum_{k=0}^{N-1} e^{\frac{2\pi i n_0 k}{N}} \overline{e^{\frac{2\pi i n_0 k}{N}}} &= \frac{1}{N} \sum_{k=0}^{N-1} e^{\frac{2\pi i (n_0 - n_0) k}{N}} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} e^0 = \frac{1}{N} \sum_{k=0}^{N-1} 1 = \frac{1}{N}(N) = 1.\end{aligned}$$

∎

Since $\mathcal{F}_N$ is orthonormal, its vectors form a basis for $V_{N,T}$. Consider the matrix A whose columns are the pure tone vectors in $\mathcal{F}_N$:

Example 5.0.6. The matrix of pure tone vectors $\{\phi_n\}_{n=0}^{N-1}$ is

$$A = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & \dots & 1 \\ 1 & e^{\frac{2\pi i}{N}} & e^{\frac{4\pi i}{N}} & \dots & e^{\frac{2\pi i n}{N}} & \dots & e^{\frac{2\pi i (N-1)}{N}} \\ 1 & e^{\frac{4\pi i}{N}} & e^{\frac{8\pi i}{N}} & \dots & e^{\frac{4\pi i n}{N}} & \dots & e^{\frac{4\pi i (N-1)}{N}} \\ \vdots & \vdots & \vdots & \ddots & & & \vdots \\ 1 & e^{\frac{2\pi i k}{N}} & e^{\frac{4\pi i k}{N}} & & e^{\frac{2\pi i n k}{N}} & & e^{\frac{2\pi i (N-1) k}{N}} \\ \vdots & \vdots & \vdots & & & \ddots & \vdots \\ 1 & e^{\frac{2\pi i (N-1)}{N}} & e^{\frac{4\pi i (N-1)}{N}} & \dots & e^{\frac{2\pi i n (N-1)}{N}} & \dots & e^{\frac{2\pi i (N-1)^2}{N}} \end{bmatrix}.$$

Multiplying a vector of amplitudes by this matrix A is a change of coordinates from $V_{N,T}$ to discrete time. Then its inverse A^{-1} changes coordinates in the other direction, from the discrete-time domain to the frequency domain. This latter matrix we will call F_N, or the $N \times N$ Fourier matrix.

To find $F_N = A^{-1}$, we recognize that multiplying A by its conjugate transpose $\overline{A}^T$ produces a matrix whose entries are the complete set of inner products of the pure tone vectors in $\mathcal{F}_N$. Lemmas 5.0.4 and 5.0.5 tell us that this matrix is simply the $N \times N$ identity matrix:

Example 5.0.7.

$$A\overline{A}^T = \begin{bmatrix} \langle\phi_0,\phi_0\rangle & \langle\phi_1,\phi_0\rangle & \langle\phi_2,\phi_0\rangle & \dots & \langle\phi_n,\phi_0\rangle & \dots & \langle\phi_{N-1},\phi_0\rangle \\ \langle\phi_0,\phi_1\rangle & \langle\phi_1,\phi_1\rangle & \langle\phi_2,\phi_1\rangle & \dots & \langle\phi_n,\phi_1\rangle & \dots & \langle\phi_{N-1},\phi_1\rangle \\ \langle\phi_0,\phi_2\rangle & \langle\phi_1,\phi_2\rangle & \langle\phi_2,\phi_2\rangle & \dots & \langle\phi_n,\phi_2\rangle & \dots & \langle\phi_{N-1},\phi_2\rangle \\ \vdots & \vdots & \vdots & \ddots & & & \vdots \\ \langle\phi_0,\phi_n\rangle & \langle\phi_1,\phi_n\rangle & \langle\phi_2,\phi_n\rangle & & \langle\phi_n,\phi_n\rangle & & \langle\phi_{N-1},\phi_n\rangle \\ \vdots & \vdots & \vdots & & & \ddots & \vdots \\ \langle\phi_0,\phi_{N-1}\rangle & \langle\phi_1,\phi_{N-1}\rangle & \langle\phi_2,\phi_{N-1}\rangle & \dots & \langle\phi_n,\phi_{N-1}\rangle & \dots & \langle\phi_{N-1},\phi_{N-1}\rangle \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & & & \vdots \\ 0 & 0 & 0 & & 1 & & 0 \\ \vdots & \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix} = I.$$

Thus $\overline{A}^T = A^{-1} = F_N$. In other words, A is unitary.[3] Now we can say that $F_N = \overline{A}^T$ and find it easily:

Definition 5.0.8. The $N \times N$ Fourier matrix is

$$F_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & \dots & 1 \\ 1 & e^{-\frac{2\pi i}{N}} & e^{-\frac{4\pi i}{N}} & \dots & e^{-\frac{2\pi ik}{N}} & \dots & e^{-\frac{2\pi i(N-1)}{N}} \\ 1 & e^{-\frac{4\pi i}{N}} & e^{-\frac{8\pi i}{N}} & \dots & e^{-\frac{4\pi ik}{N}} & \dots & e^{-\frac{4\pi i(N-1)}{N}} \\ \vdots & \vdots & \vdots & \ddots & & & \vdots \\ 1 & e^{-\frac{2\pi in}{N}} & e^{-\frac{4\pi in}{N}} & & e^{-\frac{2\pi ink}{N}} & & e^{-\frac{2\pi in(N-1)}{N}} \\ \vdots & \vdots & \vdots & & & \ddots & \vdots \\ 1 & e^{-\frac{2\pi i(N-1)}{N}} & e^{-\frac{4\pi i(N-1)}{N}} & \dots & e^{-\frac{2\pi i(N-1)k}{N}} & \dots & e^{-\frac{2\pi i(N-1)^2}{N}} \end{bmatrix}.$$

The matrix given in Example 5.0.6 is, of course, F_N^{-1}. Multiplying a frequency response vector by F_N^{-1} to produce a sample vector is called the **inverse discrete Fourier transform (IDFT)**, which is just as important for the digital filtering process.

6 Digital Filters

With our DFT and IDFT matrices in hand, the process for digital filtering becomes clear: take a sample vector $\mathbf{x}$, apply the DFT to obtain a frequency vector λ, change the amplitudes of this vector, then apply the IDFT to obtain a new sample vector that generates the filtered sound. This leads us to the following definition for digital filters:

Definition 6.0.1. A digital filter S is diagonalizable to a diagonal matrix D_{λ_S} whose diagonal terms are the terms of a frequency response vector λ_S. In other words, S is a filter if

$$S = F_N^{-1} D_{\lambda_S} F_N \qquad \text{for some } D_{\lambda_S} = \text{diag}(\lambda_S).$$

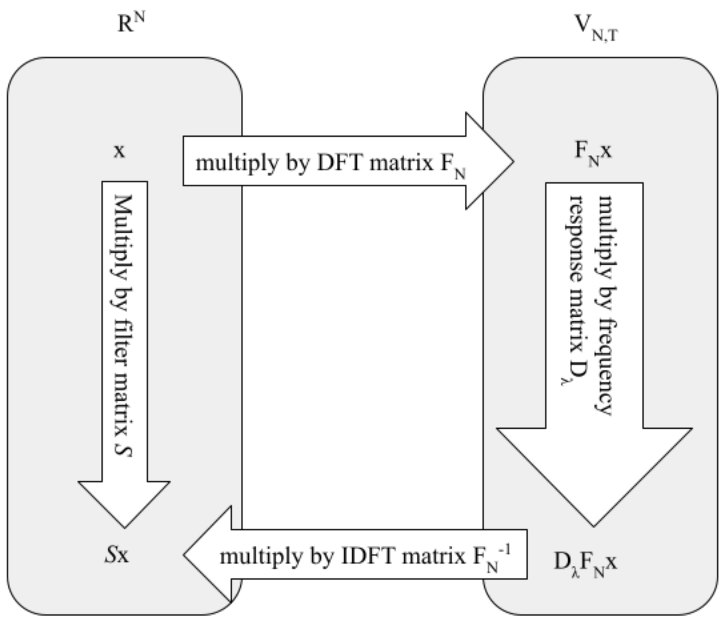


Figure 9: Moving within and between discrete time and the frequency space

Once we use the DFT and IDFT to find S for a desired frequency response, any sample vector $\mathbf{x}$ can be filtered to a new sample vector $S\mathbf{x}$.

Example 6.0.2. The term in the j'th row and k'th column[9] of a digital filter S is given by

$$s_{j,k} = (F_N^{-1} D F_N)_{j,k} = (\frac{1}{\sqrt{N}})^2 \sum_{n=0}^{N-1} e^{-\frac{2\pi ikn}{N}} \lambda_S(\frac{n}{T}) e^{\frac{2\pi ijn}{N}} = \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i(j-k)n}{N}}.$$

Note that:

$$s_{j+1,k+1} = \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i(j+1-k-1)n}{N}} = \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i(j-k)n}{N}} = s_{j,k}.$$

That means the entries of S are equal along each diagonal. To put this another way, let $l = j - k$ and note that:

$$s_{\{l=j-k\}} = \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi iln}{N}}.$$

In other words, the entries of S are equal to $\frac{1}{N} \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi iln}{N}}$ along the diagonal given by l, where $l = 0$ refers to the central diagonal, $l = 1$ the diagonal just below it, $l = -1$ the diagonal just above it, etc. Such a matrix is called a Toeplitz matrix.

Note also that:

$$\begin{aligned} s_{j,N-1} = \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S\left(\frac{n}{T}\right) e^{\frac{2\pi i(j+1-N)n}{N}} &= \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S\left(\frac{n}{T}\right) (e^{\frac{2\pi i(j+1)n}{N}} \cdot e^{-2\pi in}) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S\left(\frac{n}{T}\right) e^{\frac{2\pi i(j+1-0)n}{N}} = s_{j+1,0}. \end{aligned}$$

What this entails for S is that the last entry in row j is equal to the first entry in row $j + 1$. To put this in our l-based notation from earlier:

$$\begin{aligned} s_{\{l-N\}} = \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S\left(\frac{n}{T}\right) e^{\frac{2\pi i(l-N)n}{N}} &= \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S\left(\frac{n}{T}\right) [e^{\frac{2\pi iln}{N}} \cdot e^{-2\pi in}] \\ = \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S\left(\frac{n}{T}\right) [e^{\frac{2\pi iln}{N}} \cdot 1] &= \frac{1}{N} \sum_{n=0}^{N-1} \lambda_S\left(\frac{n}{T}\right) e^{\frac{2\pi iln}{N}} = s_{\{l\}}. \end{aligned}$$

In short, this means the entries in diagonal l are equal to the entries in diagonal $N - l$. That makes S a **circulant Toeplitz matrix**, defined in 6.0.3.

[9] For simplicity's sake, we will index filter matrices starting at 0 rather than 1 because the first row and column correspond to $n = 0$ and $k = 0$, respectively.

Definition 6.0.3. A circulant Toeplitz matrix has equivalent terms along each diagonal and "wraps around" such that the first term in each row is equal to the last term in the previous row. In other words, an $N \times N$ circulant Toeplitz matrix S takes the form

$$\begin{bmatrix} s_0 & s_{N-1} & s_{N-2} & \cdots & s_2 & s_1 \\ s_1 & s_0 & s_{N-1} & s_{N-2} & \cdots & s_2 \\ s_2 & s_1 & s_0 & s_{N-1} & \ddots & \vdots \\ \vdots & s_2 & s_1 & s_0 & \ddots & s_{N-2} \\ s_{N-2} & \vdots & \ddots & \ddots & \ddots & s_{N-1} \\ s_{N-1} & s_{N-2} & \cdots & s_2 & s_1 & s_0 \end{bmatrix}.[3]$$

Circulant Toeplitz matrices are nice because, as you may have noticed, all the information about them is contained in their first column $s_{j,0}$, which we will henceforth call $\mathbf{s}$. In fact:

Theorem 6.0.4. The frequency response vector for a filter S can be found with $\lambda_S = \sqrt{N} F_N \mathbf{s}$. If we instead know λ_S but not S, the first column of S can be found with $\mathbf{s} = \frac{1}{\sqrt{N}} F_N^{-1} \lambda_S$.[2]

Proof. Consider the frequency response vector λ_S for an arbitrary filter S.

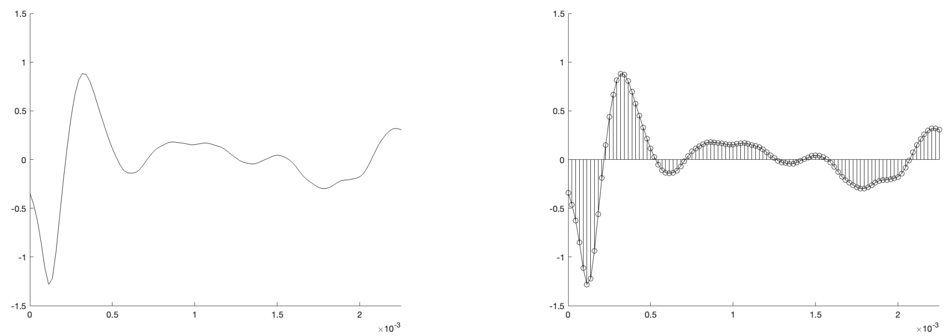
$$\frac{1}{\sqrt{N}} F_N^{-1} \lambda_S = \left(\frac{1}{\sqrt{N}}\right)^2 \begin{bmatrix} \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i n}{N}} \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{4\pi i n}{N}} \\ \vdots \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i j n}{N}} \\ \vdots \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i (N-1) n}{N}} \end{bmatrix} = \frac{1}{N} \begin{bmatrix} \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i (0-0) n}{N}} \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i (1-0) n}{N}} \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i (2-0) n}{N}} \\ \vdots \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i (j-0) n}{N}} \\ \vdots \\ \sum_{n=0}^{N-1} \lambda_S(\frac{n}{T}) e^{\frac{2\pi i ((N-1)-0) n}{N}} \end{bmatrix}.$$

From Example 6.0.2, we see that the j'th term of this column vector is equal to $s_{j,0}$. Therefore, the vector must be equal to $\mathbf{s}$, the first column of S.

Then also $\sqrt{N} F_N \mathbf{s} = \sqrt{N} F_N (\frac{1}{\sqrt{N}} F_N^{-1} \lambda_S) = \lambda_S$. ∎

6.1 Example: Filtering a Sampled Trumpet

Digital filters expand our creative options immensely. We no longer have to know a waveform's amplitude-over-time function $f(t)$; we can apply the DFT to any recording of a periodic sound we wish. For instance, we can record a trumpet sustaining the note A4 at a rate of 44,000 samples per second.[10]



The trumpet's unknown waveform $f(t)$ $\qquad$ $f(t)$ sampled 100 times

Figure 10: Sampling the waveform of a [▶] trumpet playing A4

The first 100 of these samples cover the period of A4's fundamental frequency, $\frac{1}{T} = 440$ Hz. We take those samples as a vector $\mathbf{x}$ and filter it with 100×100 matrices.

First, we create the frequency response vectors for an LPF and an HPF. We know that the frequencies in $V_{N,T}$ are lowest near the two extremes, $n = 0$ and $n = N$, while frequencies increase as they approach $n = \frac{N}{2}$.[2] To see which frequencies are contributing meaningfully to the tone, we look at the plot of $F_N\mathbf{x}$:

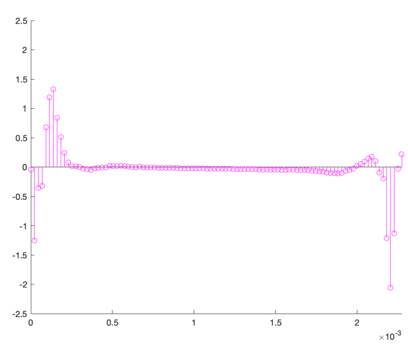


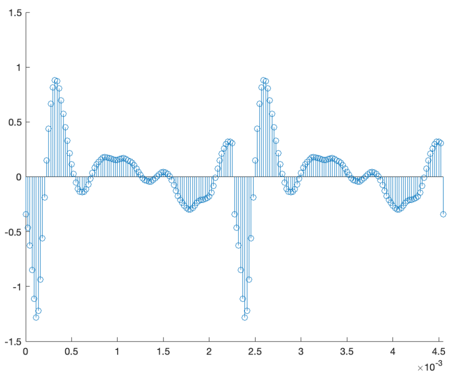
Figure 11: $DFT\mathbf{x}$

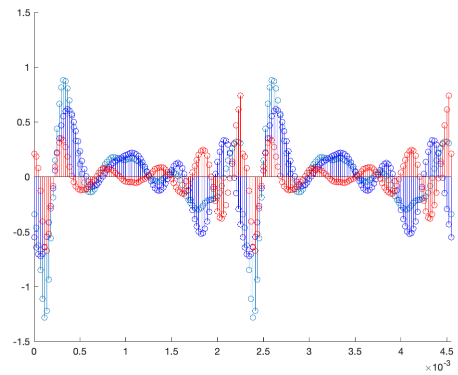
[10]This is quite close to the most common sampling rate, 44,100. Our slight adjustment greatly simplifies the math for this example.

Seeing that the frequencies become insignificant between $n = 10$ and $n = 90$, we set the frequency responses of the LPF and HPF such that they preserve a similar number of important frequencies:

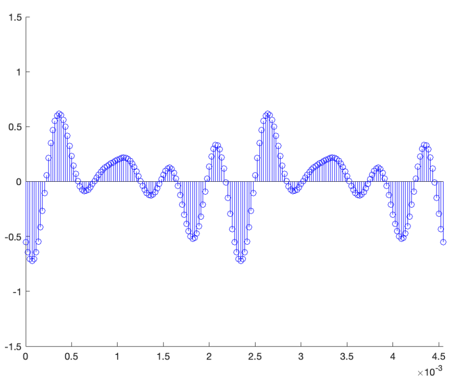
$$\lambda_{LPF} = \{\underbrace{1, \ldots, 1}_{5}, \underbrace{0, \ldots\ 0}_{90}, \underbrace{1, \ldots\ 1}_{5}\} \qquad \lambda_{HPF} = \{\underbrace{0, \ldots, 0}_{5}, \underbrace{1, \ldots\ 1}_{90}, \underbrace{0, \ldots\ 0}_{5}\}$$

Setting $D_{\lambda_{LPF}} = \text{diag}(\lambda_{LPF})$ and $D_{\lambda_{HPF}} = \text{diag}(\lambda_{HPF})$, we obtain $S_{LPF} = F_N^{-1} D_{\lambda_{LPF}} F_N$ and $S_{HPF} = F_N^{-1} D_{\lambda_{HPF}} F_N$. Finally, we can generate two new tones from our sampled trumpet, given by the sample vectors $\mathbf{x}_{LPF} = S_{LPF}\mathbf{x}$ and $\mathbf{x}_{HPF} = S_{HPF}\mathbf{x}$. Figure 12 plots our three vectors over two periods, with their audio realizations linked below.

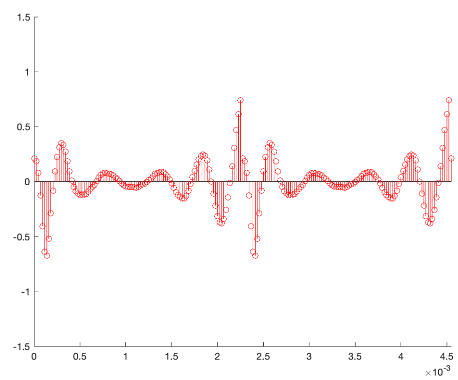




[▶] $\mathbf{x}$



[▶] $\mathbf{x}_{LPF}$



[▶] $\mathbf{x}_{HPF}$

Figure 12: Applying the digital LPF and HPF to the trumpet sample

6.2 Conclusion: Tip of The Iceberg

Digital filters enable music production techniques far beyond manipulating the sound quality of a single sustained note.

Filters can be applied to an increasing or decreasing degree over time. A series of filters artfully applied to create an effect is called an **envelope**. Suppose we want to create a "plucked" note, which quickly decays in amplitude. Then we can simulate a host of instruments which don't sustain notes, such as guitars, pianos, and percussion instruments. By applying a series of low-pass filters over time, we can eliminate frequencies one-by-one from highest to lowest. This replicates how sound decays in the physical world: the higher a frequency, the more quickly it fades away after the vibration that created it has stopped.[1]

Filters can also be generalized to non-periodic wave functions, opening a whole new can of worms. Such filters operate on irrational vectors called angular frequencies, allowing them to have infinite dimensions. General filters can be applied to any audio clip to smooth the waveform, add reverb and delay, and warp the sound quality of irregular musical sounds like vocals, among many other capabilities.[2]

Though we must leave much of the world of linear algebra in music unexplored, hopefully this paper has helped the reader build the necessary confidence to turn knobs like the professionals. For added benefit, the author recommends wearing sunglasses onstage. As a great man once said, "Let me put my sunglasses on here / So I can see what I'm doing."

References

[1] J. Pierce, "Chapter 4: Sound Waves and Sine Waves." From *Music, Cognition, and Computerized Sound.* MIT Press, Cambridge, MA, 1999.

[2] O. Ryan, *Linear Algebra, Signal Processing, and Wavelets – A Unified Approach.* Springer, London, UK, 2019.

[3] F. Zhang, *Matrix Theory.* Springer, New York, NY, 2011.

[4] O. Bretscher, *Linear Algebra with Applications.* Fifth ed., Pearson, Upper Saddle River, NJ, 2013.