Matthew Schwartz

Lecture 7: Music

1 Why do notes sound good?

In the previous lecture, we saw that if you pluck a string, it will excite various frequencies. The amplitude of each frequency which is excited will be proportional to the coefficient in the Fourier decomposition. In this lecture we will start to understand how different frequencies combine to produce music. This lecture is best studied alongside the Mathematica notebook music.nb on the isite.

In the first section "playing notes" of the notebook, you can listen to a pure frequency (C4 = middle C = 261 Hz). It sounds pleasant, but not particularly interesting. Now play the "square wave" version of middle C – you should notice that it sounds somewhat tinny and unpleasant. These are the same notes, but different sounds. Why do they sound different? One way to understand the difference is to compare the Fourier decomposition of the sine wave and the square wave:

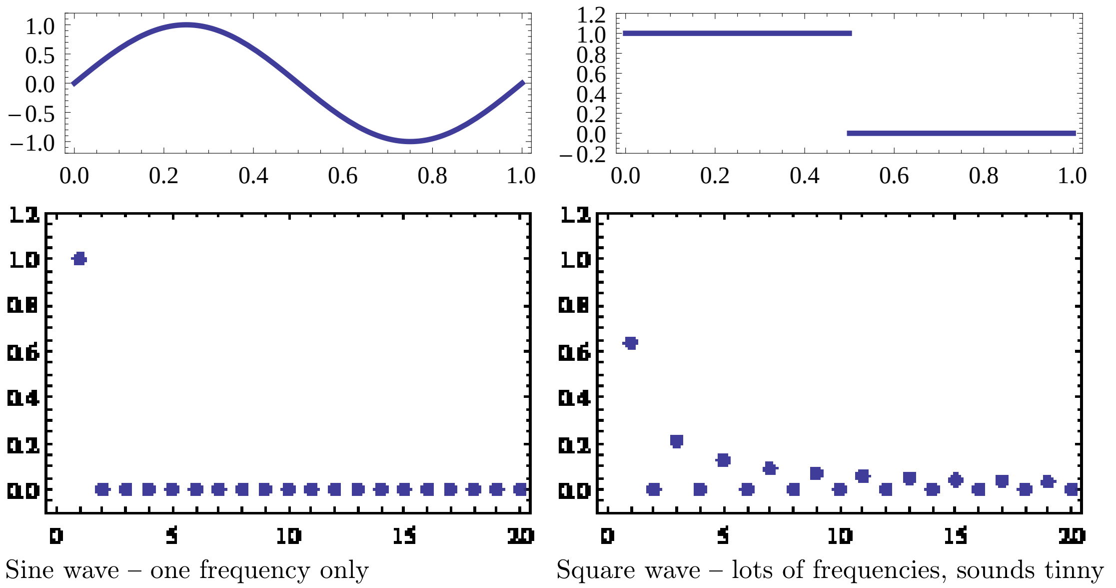

Figure 1. Comparison of Fourier modes for sin and square wave

For both the sine wave and the square wave, the dominant mode is the fundamental. However, the square wave has lots of other modes which make the note sound less pure. One way to understand why the square wave sounds worse is that has many high frequency notes with significant amplitude. It is hard for our brains to process all these high frequency notes, so we find it jarring. In fact, if all frequencies are present at once we get so-called **white noise.** White noise is perhaps as unmusical as you can get.

Now consider playing two notes at once. In the "playing pairs of notes" section, play the 250 Hz and 270 Hz notes at the same time. It doesn't sound great. Why?

The problem is that you hear a rattling around 20 times a second. This 20 Hz rattling is the beat frequency between the 250 Hz and 270 Hz. Indeed,

$$\cos(270\,\mathrm{Hz}\,2\pi t) + \cos(250\,\mathrm{Hz}\,(2\pi t)) = 2\cos(10\,\mathrm{Hz}\,2\pi t)\cos(260\,\mathrm{Hz}2\pi t) \tag{1}$$

This the sum of those two notes oscillates at 10 Hz and at 260 Hz. The 10 Hz oscillation (which we hear as a 20 Hz beat frequency) is jarring – your mind tries to process it consciously. Frequencies as high as 260 Hz do not have this effect.

Thus there seem to be two reasons sounds appear unmusical:

- Too many frequencies are present at once
- Beating occurs at frequencies we can consciously process.

This is a physics class, not a biology class, so we will not try to explain why these facts hold. We merely observe that whenever each criteria is satisfied, sounds appear unmusical. However, now that we have defined the problem, we can start to study music scientifically.

2 Dissonant and consonant note pairs

Now we're ready to study music. If we had a pure sine wave at 300 Hz, then it sounds nasty when played at the same time as a sine wave of 320 Hz. However, if we play it with a sine wave of 580 Hz it does not sound so bad. That is because

$$\cos(300\,\mathrm{Hz}\,2\pi t) + \cos(580\,\mathrm{Hz}\,2\pi t) = 2\cos(140\ \mathrm{Hz}\,2\pi t)\cos(440\ \mathrm{Hz}\,2\pi t) \tag{2}$$

The beat frequency $2 \times 140\,\mathrm{Hz} = 280\mathrm{Hz}$ is not low enough to be harsh – it is just a note (try the Mathematica notebook). On the other hand, if we played 300 Hz and 580 Hz on an actual instrument it would sound horrible.

We can see why from studying our plucked string example. Recall that for a string plucked near the end, the relative Fourier coefficients scale like $\frac{1}{n}$. So the dominant frequency $n = 1$ (the fundamental) has only twice the amplitude of the first harmonic ($n = 2$). Thus playing 580 Hz along side a plucked string would give

$$f(t) = \cos(580\ \mathrm{Hz}\,2\pi t) + \sum_{n=1} \frac{1}{n}\cos(300n\ \mathrm{Hz}\,2\pi t) \tag{3}$$

Writing $T = \mathrm{Hz}\,2\pi t$ to clean up the equation and expanding the sum

$$f(t) = \cos(580\,T) + \cos(300\ T) + \frac{1}{2}\cos(600\,T) + \cdots \tag{4}$$

Let us combine the 580 Hz oscillation with the 600 Hz oscillation using trig sum rules. Using

$$\cos(580T) + \frac{1}{2}\cos(600T) = \frac{3}{2}\cos(10T)\cos(590T) + \frac{1}{2}\sin(10T)\sin(590T) \tag{5}$$

we find

$$f(t) = \cos(300T) + \frac{3}{2}\cos(10T)\cos(590T) + \frac{1}{2}\sin(10T)\sin(590T) + \sum_{n=3}^{\infty} \frac{1}{n}\cos(300nT) \tag{6}$$

Now we see beating at $2 \times 10\,\mathrm{Hz} = 20\mathrm{Hz}$, which is audible and jarring. There is beating between the 580 Hz note and the first harmonic of the plucked string. The point is that with pure sine waves, no harmonics are excited, but with real instruments they are.

In general, instruments will have significant amplitudes for many harmonics. Here is the Fourier spectrum of a flute

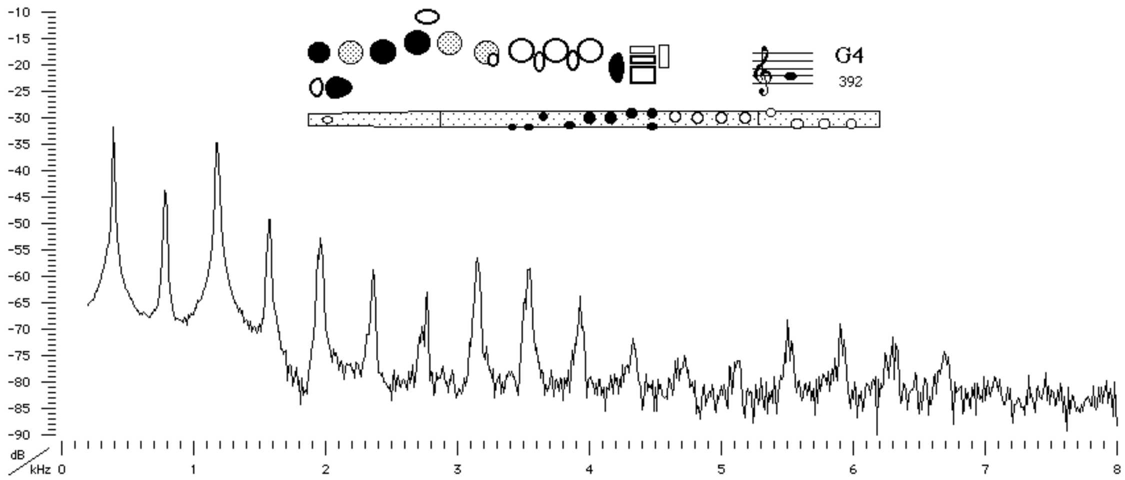

Figure 2. Spectrum of the note G4 on a flute.

This is a flute playing the note G4 which is 392 Hz. You see that not only is the fundamental frequency (the G4) largely excited, but many other modes have significant amplitudes. These modes are all the higher harmonics. The harmonics determine what an instrument sounds like – its **timbre**. Timbre is what a note sounds like when played, as distinguished from its **pitch**, which is the fundamental frequency played, and **intensity**, which is the power going into sound from the instrument. We will study the timbres of different instruments in lecture, in section and on psets. The pitch of a note is the frequency with highest intensity, which is usually the lowest frequency where there is a peak. If the first peak is not the highest, what the pitch is can be somewhat subjective. The relative heights of the different peaks and their Q values is the timbre, and the absolute scale is the intensity. All of this can be read off the Fourier spectrum, as in Figure 2 for a flute. We'll discuss timbre more later, but for now, the main point is that most instruments will have integer multiples of the fundamental frequency excited with significant amplitudes; it is these harmonics which are the key to the scale in Western music.

Key points are:

- On a real instrument, there will be unmusical beating whenever an integer multiple of one harmonic is close but not equal to an integer multiple of another harmonic.
- Conversely, the most consonant notes will have some harmonics which exactly agree.

For example, let's start with middle C. This note is called $C4$ (the C is the note and 4 is the octave) and has a frequency of $\nu_0 = 261$ Hz. Which notes sound good along side C4? Well, the 261 Hz note has harmonics of $\nu_0, 2\nu_0, 3\nu_0, 4\nu_0$, etc.:

$$261\,\text{Hz}, \quad 522\,\text{Hz}, \quad 783\,\text{Hz}, \quad 1044\,\text{Hz}, \quad 1305\,\text{Hz}, \cdots \tag{7}$$

Thus if we play any of those notes along with $C4$ it will sound harmonic. If the fundamental frequency is ν_0 then

$$2\nu_0 = 1\,\text{octave} = 1\text{st harmonic} = C5 \tag{8}$$

Are there more notes which are harmonious? Yes. Consider the note with $\nu_5 = 391\,\mathrm{Hz}$. This note has $2\nu_5 = 783\,\mathrm{Hz}$. Thus the second harmonic of ν_5 matches the 3rd harmonic of ν_0. We call the note $\nu_5 = \frac{3}{2}\nu_0$ the **perfect fifth**

$$\nu_5 = \frac{3}{2}\nu_0 = \text{perfect fifth} = G4 \tag{9}$$

This is the G above middle C. In the same way, consider $\nu_4 = 348\,\mathrm{Hz}$. The 3rd harmonic of ν_4 agrees with the 4rd harmonic of ν_0. We call this the perfect fourth

$$\nu_4 = \frac{4}{3}\nu_0 = \text{perfect fourth} = F4 \tag{10}$$

And so on.

It is easy to see that any rational number ratio of frequencies will be consonant. Many of these ratios have names

$\frac{\nu}{\nu_0}$	1	2	$\frac{3}{2}$	$\frac{4}{3}$	$\frac{5}{4}$	$\frac{6}{5}$	$\frac{5}{3}$	$\frac{8}{5}$
name	fundamental	octave	perfect fifth	perfect fourth	Major third	Minor third	Major sixth	Minor sixth
example	C4	C5	G4	F4	E4	E♭	A5	A♭5

Table 1. Notes names and ratios in the **just intonation scale**

There are an infinite number of rational numbers. So where do we stop? The answer is that the lower the numbers in the ratio (that is, the 3 and 2 in $\frac{3}{2}$ are lower than the 8 and 5 in $\frac{8}{5}$), the more consonant they will be. That's because for something like $\frac{11}{17}$, one would need the 17th harmonic of one note to match the 11th harmonic of another note. By such high harmonics, the amplitudes are no longer large, and the spectrum is messy (as you can see in Figure 2). Also, it is more likely for frequencies with a ratio of $\frac{11}{17}$ to give harmonics which are close but not equal, generating dissonant beating, before generating the harmonic consonance. Thus, for numbers large than about 5 in the ratio, notes are no longer appreciated as harmonic.

3 Scales

If we have a given note, say $C4$, we can define all the other notes so that they will be harmonic with $C4$.

3.1 Just intonation scale

The most harmonic notes will have the smallest integers in the ratio, as in Table 1. This is a particular choice of tuning known as the **just intonation scale**. The just intonation scale is in a sense the most harmonic choice for the frequencies of notes in a scale (it is default tuning for some non-Western instruments, such as the Turkish Baglama). But it is just a choice.

Note that if we pick a set of notes that sound harmoic with $C4$, the same set of notes will generically not sound harmonic with another note, like $D4$. Thus if we're playing a song, the set of notes we want to use is determined relative to some starting note. This starting note is called the **key**. For example, if you are in the key of C, the notes C, G and F sound good. But if you are in the key of D, the notes C, G and F will generally not sound as good.

On some instruments, such as a violin which has no frets, there are no predefined notes. Thus on a violin, if you work in the key of C, you can play all the harmonics in exactly the right place (if you have a skilled enough ear and hand). Thus you can play the just intonation scale in any key. There are an *infinite* number of notes you can play on a violin. The same is true on most instruments actually. For example, even though an oboe has fixed holes corresponding to notes, oboe players can easily move the notes up or down by manipulating the reed. Controlling the precise frequency of a note with your mouth is critical to playing any woodwind instrument well.

On other instruments, like a piano or a stringed instrument with frets like a guitar, the notes are essentially built in to the instrument. You can sometimes tweak the notes if you are skilled, or tune the instrument to a different key, but there are a finite number of notes which can be played in a giving tuning. Unfortunately, it's impossible to have an instrument with a finite number of notes be capable of playing the most harmonic notes in every key.

To see the problem, suppose you want your piano to be in the just intonation scale in the key of C. That means that you want all the other notes to be defined so that they are related to C by rational numbers with low integers. For example, the whole notes can be defined as

note	C	D	E	F	G	A	B	C
$\frac{\nu}{\nu_0}$	1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	2
decimal	1	1.125	1.25	1.333	1.5	1.666	1.875	2

Table 2. Notes names and ratios in **just intonation**

This defines the note D as having the frequency $\frac{9}{8}$ times the frequency of the C. Now, where is the 5th of D? This should be at $\frac{3}{2} \times \frac{9}{8} = \frac{27}{16} = 1.688$ times ν_0. This note is somewhere between the A and the B, but it is not exactly a note in the key of C. It is not hard to see that to get an instrument which could play any note in any key, you would need an enormous number of available notes. Please make sure you understand this point, as it is key to understanding scales.

So what can we do? There are two options: the first is you can tune your instrument to the key you want to play in. Stringed instruments can do this. But it is not so appealing of an option if we want to play music in different keys without retuning every time. The other option is to compromise. Most of our ears are not sensitive enough to distinguish close but slightly different notes. Thus we can choose scales which are not exactly correct in any key, but close to correct in all keys.

3.2 Pythagorean scale

One way to approximate the scale is by choosing the notes to be related by powers of $\frac{3}{2}$ and octaves to be related by factors of 2, For example, $\frac{3}{2}\nu_0$ is the perfect fifth (G4 when C4=ν_0). Then $\left(\frac{3}{2}\right)^2\nu_0$ which is the fifth of the fifth, or the fifth of $G4$ which is $D5$. The next note has $\left(\frac{3}{2}\right)^3\nu_0$ or the fifth of $D5$ which is $A6$ and so on. We can bring any power of $\frac{3}{2}$ back to the interval between 0 and 1 by dividing by 2 to some power. For example, since $D5 = \frac{9}{4}\nu_0$ then $D4 = \frac{9}{8}\nu_0$. We then get

note	C	D	E	F	G	A	B	C
$\frac{\nu}{\nu_0}$	1	$\frac{9}{8}$	$\frac{81}{64}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{27}{16}$	$\frac{243}{128}$	2
decimal	1	1.125	1.266	1.333	1.5	1.688	1.898	2

Table 3. Notes names and ratios in Pythagorean tuning

This is called the **Pythagorean tuning**. Note that the octave, perfect fifth (G) and perfect fourth (F) agree with their values in Table 2. Some notes do not agree: for example, E is $\frac{81}{64}\nu_0 = 1.266\nu_0$ in this tuning. This ratio is close to $\frac{5}{4}$ but not exactly. Thus if we play C and E it will be close to a consonant sounding note, but not exactly.

The advantage of the Pythagorean tuning is that the perfect fifth and perfect fourth of every note is included in the scale.

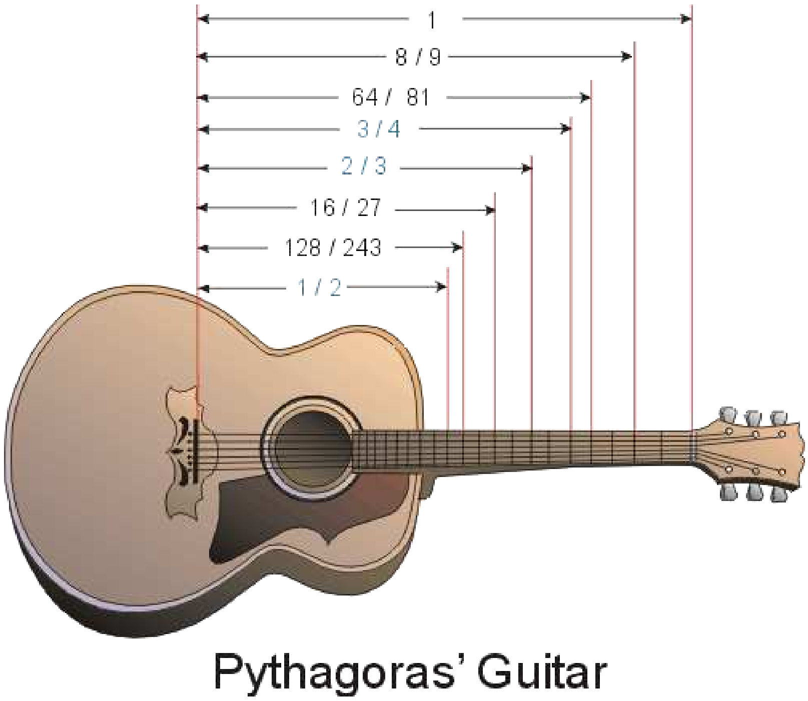

Figure 3. Pythagorean guitar: frets positions are related by powers of $\left(\frac{3}{2}\right)^n 2^m$

3.3 Equal-tempered scale

An interesting feature of the Pythagorean scale is that the 12th fifth is very close to 8 octaves:

$$\left(\frac{3}{2}\right)^{12} = 129.748 \approx 128 = 2^7 \tag{11}$$

Thus another compromise is to say that when we go 12 steps around the circle of fifths, we get back to the note we started at. Then we can devise a scale which does not choose any key. We simply relate the notes by powers of $2^{1/12}$. Each half-step gives another factor of $2^{1/12}$. The result are the frequencies in Table 4.

note	C	C♯	D	D♯	E	F	F♯	G	G♯	A	A♯	B	C
$\frac{\nu}{\nu_0}$	1	$2^{\frac{1}{12}}$	$2^{\frac{2}{12}}$	$2^{\frac{3}{12}}$	$2^{\frac{4}{12}}$	$2^{\frac{5}{12}}$	$2^{\frac{6}{12}}$	$2^{\frac{7}{12}}$	$2^{\frac{8}{12}}$	$2^{\frac{9}{12}}$	$2^{\frac{10}{12}}$	$2^{\frac{11}{12}}$	2
decimal	1	1.059	1.122	1.189	1.260	1.335	1.414	1.498	1.587	1.682	1.782	1.888	2

Table 4. Notes names and ratios in the **equal-tempered scale**.

This tuning is called the **equal-tempered scale**. The equal-tempered scale is the standard tuning for all of Western music.

Because of Eq. (11) if you keep going up by perfect fifths, and normalizing by octaves, you will get back to the note you started at. We can see this in the **circle of fifths**

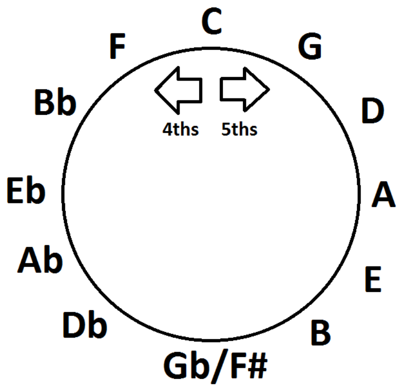

Figure 4. Circle of fifths. Each note going clockwise is a perfect fifth above the previous note. Going counterclockwise, each note is a perfect 4th above the previous note. The circle only closes in the equal-tempered scale.

In this circle, each note is 1 fifth above the note clockwise. So G is a fifth above C, D a fifth above G and so on. Going counterclockwise, the intervals are fourths: C is a fourth above G and F is a fourth above C. Going up a fifth is the same as going down a fourth and adding an octave,

If the notes are defined with the Pythagorean scale, the circle doesn't close: going up by 12 steps, and normalizing back down to the original octave leaves you $\left(\frac{3}{2}\right)^{12}2^{-8}=1.014$ times where you started. Thus the circle doesn't close by 1.4%. In the equal-tempered scale, it *does* exactly close, however, none of the notes have frequency ratios of exactly $\frac{2}{3}$.

3.4 Summary

We discussed 3 scales. The just intonation scale chooses notes to be related by rational number ratios with integers as small as possible in the numerator and denominator. The Pythagorean scale has all notes related by 3^n2^m for some m and n. Both just intonation and the Pythagorean scale require a key to start in. The third scale is the equal-tempered scale. Notes in the equal-tempered scale are related by $2^{\frac{n}{12}}$ for some n.

Here is a comparison between the relative frequencies of the 3 scales in the key of C:

note	C	D	E	F	G	A	B	C
just-intonation	1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	2
Pythagorean	1	$\frac{9}{8}$	$\frac{81}{64}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{27}{16}$	$\frac{243}{128}$	2
equal-tempered	1	$2^{\frac{2}{12}}$	$2^{\frac{4}{12}}$	$2^{\frac{5}{12}}$	$2^{\frac{7}{12}}$	$2^{\frac{9}{12}}$	$2^{\frac{11}{12}}$	2

Table 5. Comparison of scales: exact ratios.

In decimals

note	C	D	E	F	G	A	B	C
just-intonation	1	1.125	1.25	1.333	1.5	1.666	1.875	2
Pythagorean	1	1.125	1.266	1.333	1.5	1.688	1.898	2
equal-tempered	1	1.122	1.260	1.335	1.498	1.682	1.888	2

Table 6. Comparison of scales: decimal approximations.

Here is a graphical comparison of how far off the frequency is in the equal-tempered scale from the frequency in the just intonation scale

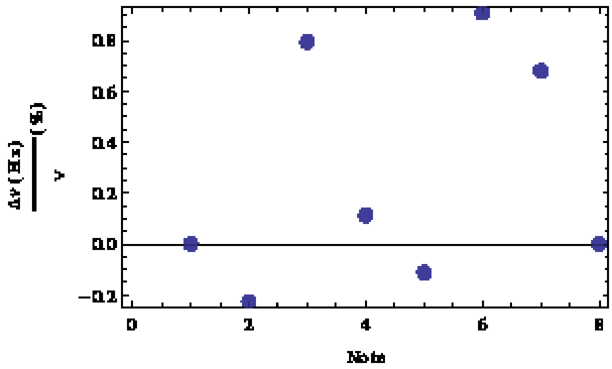

Figure 5. Difference between the equal-tempered frequencies ν_{WT} and the just intonation frequencies for whole notes $C, D, E....C$ labeled as 1 to 8.

One thing we can see is that the perfect 4th and perfect 5th are very close to their optimal values, while the 6th and 7ths are not so close.