

Chapter 7: Musical scales, frequencies, and temperament

The next few chapters will quantify certain aspects of sound and music. In this chapter, I will discuss how the musical scale has been constructed in several different ways historically. A background in music is not required, but it will make the following pages a little more familiar. What we want to do specifically is learn how different *frequencies* of sound are connected with different musical *itches* you all are familiar with.

There is no one solution to this problem! I will review three systems of *temperament*, or tuning systems. Each has a slightly different method of assigning frequencies to musical notes.

Musical background information

My discussion will center on the Western musical scale, which is typically divided into twelve notes. The *major scale* is a subset of these. On a musical staff, the scale may look like this:



The major scale starting on middle C

This scale is called the C major scale, and it is simple because it can be played entirely on the white keys of a piano, e. g., there are no notes labeled “sharp” or “flat” in there. If you prefer, you can think of the *solfège* names for the notes (Do, Re, Mi, etc., like those used in [The Sound of Music](#)).

The C major scale has the following notes: C D E F G A B C, where the “C” notes at the start and end are *octaves* of each other. The twelve different labels for notes in most Western systems are:

A A# B C C# D D# E F F# G G#

The sharpened notes also have a flat name, that is, A# and B ♭ (B-flat) are the same note. Each gap between two consecutive notes here is called a half-step or semitone. You can see that B-C and E-F are a semitone apart. Therefore, the gap between the third and fourth note of the major scale is a semitone, as is the gap between the seventh and eighth note. There is a whole tone separation between all the other notes in the major scale.

Another musical term it would help to know is the idea of an *interval*. This is simply the gap between two notes on a scale. A *second interval* is the gap between the first and second notes in a scale. This is like the gap between C-D on a C major scale, or the gap between “Do” and “Re” in any major scale. Similarly, a *major third* is the gap between the first and third notes in a major scale (C-E in the C major scale, or the gap between “Do” and “Mi.”) Minor scales use a different third interval, hence the specification.

Similarly, a *fourth*, *fifth*, *sixth*, or *seventh* corresponds to the jump between the first note in a scale (also called the root) and the fourth, fifth, sixth, or seventh note respectively. The different temperament systems all have different ways of defining just how large these jumps in frequency are.

Whole number ratios, Pythagorean, and Just tuning systems

Pythagorean tuning

Pythagoras, who you might know best for his Pythagorean Theorem in geometry, was also a philosopher who studied music in the 6th century BCE. The result of his studies is that two strings whose lengths are in simple, small whole-number ratios sound harmonious together. The key ratios are as follows:

If the string length ratio is:	The musical interval is:
1:1	Two notes in unison
1:2	One octave (hear it here)
2:3	A perfect fifth (hear it here)
3:4	A perfect fourth (hear it here)

The string length ratio and the frequency ratio are the same since for a given string, length and frequency are inversely proportional.

So what does this mean? Trivially, we can say that two notes of the same frequency are in unison with each other. Also, if a note is double the frequency of a second note, those two form an octave (they represent the beginning and ending note in a major scale.)

Also, *any* two frequencies in a 2:3 ratio form a perfect fifth and any two frequencies in a 3:4 ratio form a perfect fourth. To get a perfect fifth above any given frequency, simply multiply that frequency by 3/2 or 1.5.

The 2:3 and 3:4 ratios are very important in the Pythagorean temperament. All the other notes in a major scale are derived from these ratios. For example, to get the fraction you need to multiply a frequency by to get the *second* note in a scale, one would go up a fifth and down a fourth like so:

$$f(2nd) = f(1st) \times \frac{3}{2} \div \frac{4}{3} = f(1st) \times \frac{9}{8}$$

Therefore, if f represents the frequency of “1st,” $f \times (9/8)$ represents “2nd.” Since the gap between scale notes 1 and 2 is the same size as the gap between 2 and 3, multiply by $9/8$ again to get the 3rd frequency.

$$f(3rd) = f(1st) \times \left(\frac{9}{8}\right)^2 = f(1st) \times \frac{81}{64}$$

As you can see, the fractions get more complicated quickly! If you multiply the $3/2$ fraction used to get the fifth note by $9/8$, you get the more complicated fractions for the 6th and 7th. I will summarize the results in the table below – with it, you can find any frequency in the major scale according to the Pythagorean system:

Note in the scale (“1st” is called f here)	Frequency of the note
1st	f
2nd	$\frac{9}{8}f$
3rd	$\frac{81}{64}f$
4th	$\frac{4}{3}f$
5th	$\frac{3}{2}f$
6th	$\frac{27}{16}f$
7th	$\frac{243}{128}f$
8th (octave higher)	$2f$

The Pythagorean temperament system

Example: One octave’s version of the note A is 220 Hz in standard concert tuning. The 6th note in that scale is an F# - what is its frequency in Pythagorean tuning? Simply check the table – to get the 6th note, multiply your basic root frequency by $27/16$. $220 \text{ Hz} \times (27/16) = 371.25 \text{ Hz}$.

It is possible to construct the other notes in the Pythagorean temperament system using the 3:2 and 4:3 ratios in different ways, and the fractions are complicated for those as well. You can see that not all these notes represent small whole-number ratios with the root note. The 3rd, 6th, and 7th are particularly guilty parties. The next system we’ll look at – *Just tuning* or *Just intonation*, does a better job maintaining the small, whole-number ratio system than followers of Pythagoras envisioned.

Just tuning

A Just tuning system is one where notes in a scale have frequencies that are small, whole-number ratios of the root frequency. Now, this sounds like the Pythagorean system, and it’s true the Pythagorean system starts out this way, but once you venture from Pythagoras’s original fourth and fifth intervals,

things get complicated. The Just system takes the least elegant fractions and tweaks the tuning slightly so the whole number ratios are a little simpler. For example, the 3rd ratio in the Pythagorean system is 81:64. If you flatten this note by a factor 80:81, (multiply the fractions 81/64 x 80/81), you get a resulting ratio of 80:64, which reduces to 5:4—a much simpler fraction. Similarly, the 6th and 7th intervals are changed slightly so they represent cleaner whole-number ratios compared to your root note. See the table here:

Note in the scale (“1st” is called f here)	Frequency of the note
1st	f
2nd	$\frac{9}{8}f$
3rd	$\frac{5}{4}f$
4th	$\frac{4}{3}f$
5th	$\frac{3}{2}f$
6th	$\frac{5}{3}f$
7th	$\frac{15}{8}f$
8th (octave higher)	$2f$

The Just temperament system

The Pythagorean ratios that were already small, whole-number ratios remain the same in the Just system; only the 3rd, 6th, and 7th intervals change.

Example: One octave’s version of the note A is 220 Hz in standard concert tuning. The 6th note in that scale is an F# - what is its frequency in Just tuning? Simply check the table – to get the 6th note, multiply your basic root frequency by 27/16. 220 Hz * (5/3) = 366.7 Hz. If you compare with the example on the previous page, this is almost 5 Hz different than the Pythagorean tuning. It’s a small difference, but it is noticeable, and a keyboard playing the same chord in the two systems can sound slightly different (VIDEO).

A problem with Just tuning, and an introduction to equal temperament

Notes played in a Just temperament naturally sound good together. Human singers often revert to a Just temperament to sound in tune with each other. However, there are problems with Just tuning that have been known for centuries. The biggest such problem occurs if you want to do *modulation*, or a key change. The basic problem is this: *on many instruments, if you want to change keys, you have to retune everything.*

Here's a brief discussion why. Suppose you build a scale around an A-natural of 220 Hz. According to the Just system, the next note in the scale (B) would have a frequency of $220 * (9/8)$ or 247.5 Hz. The following note (C#) would have a frequency of $220 * (5/4)$ or 275 Hz. Now, suppose you wanted to do a key change to the key of B. Suppose your B is 247.5 Hz, like before. According to the Just tuning system, the next note in the scale (still C#) would be $247.5 * 9/8$, or 278.4 Hz. This is a small but noticeable difference from the C# in the previous key. So in summary: *changing keys in a Just temperament requires changing the frequency of some of the notes*. In addition, intervals that sound good in one Just tuning would sound bad in other keys (see an article on [wolf intervals](#) for more information). One solution is to introduce more notes, for example, make D# and E ♭ different notes so both the B/D-sharp and E-flat/G intervals sound equally good. That is very impractical. See the section of this chapter on the Pythagorean comma for more details on the problems created by adding sharps and flats to a Pythagorean or Just tuning system.

So what do we do? We certainly can't go without [key changes](#) in music, so we need a new system. The *equal temperament* system (sometimes called 12-TET for "twelve-tone equal tempered") is the most commonly used system for modern instruments, most especially instruments that can't be tuned on the fly, like pianos. The 12-TET system does not preserve perfectly the ideal interval ratios the Pythagorean and Just system do, but it provides the best compromise that allows for key changes as well.

The 12-TET: To construct a set of twelve frequencies, divide the octave up into twelve equal jumps. The way to get the next semitone up from any frequency mathematically is to multiply that frequency by $\sqrt[12]{2}$ (the twelfth root of two, approximately equal to 1.05946). If you want to move higher up the scale, just keep multiplying by the twelfth root of two that many times. In general, if f is your starting frequency, and you want to move up n semitones, the new frequency will be:

$$f_{new} = f \times 2^{(\frac{n}{12})}$$

If you want to move down the scale instead of up, simply make n negative. Notice that our octave structure is the same – to go up an octave, that's 12 semitones, and $2^{(12/12)} = 2$, which produces a doubling in frequency like we expect.

Example: The major scale's notes are constructed by starting at a root note and moving up the following number of semitones:

1 st	Root note
2 nd	2 semitones higher
3 rd	4 semitones higher
4 th	5 semitones higher
5 th	7 semitones higher
6 th	9 semitones higher
7 th	11 semitones higher
8 th	12 semitones (1 octave) higher

Commonly, the 12-TET scale is centered around a concert pitch of A = 440 Hz. Therefore, to construct an A-major scale, the frequencies are as follows:

Note	# of semitones from root	Formula	Frequency (Hz)
A	0	$440 \cdot 2^{(0/12)}$	440
B	2	$440 \cdot 2^{(2/12)}$	493.9
C#	4	$440 \cdot 2^{(4/12)}$	554.4
D	5	$440 \cdot 2^{(5/12)}$	587.3
E	7	$440 \cdot 2^{(7/12)}$	659.3
F#	9	$440 \cdot 2^{(9/12)}$	740.0
G#	11	$440 \cdot 2^{(11/12)}$	830.6
A	12	$440 \cdot 2^{(12/12)}$	880

The A-major scale according to 12-TET

For another example, suppose you wanted the F# *below* 440 Hz. There are two ways to get it. If you know the F# in the table is 740 Hz, you can divide that by two and get 370 Hz. Alternatively, if you know F# is three semitones below A, you can calculate $440 \cdot 2^{(-3/12)} = 370$ Hz.

Cents and the Pythagorean comma

The purpose of this section is to show some of the problems that come up in trying to introduce semitone intervals in Pythagorean and Just tuning system. Before I go into those specifics, however, I want to introduce the idea of a “cent.” As the name implies, a “cent” is a subdivision of a semitone. We want to divide each semitone into 100 cents each. If our Western musical scale has 12 semitones in an octave, there will be 1200 cents in an octave. The formula for relating two frequencies f_1 and f_2 that are x cents apart is:

$$f_2 = f_1 \times 2^{(\frac{x}{1200})}$$

Most humans can hear differences of about five cents or greater.

We can rearrange this formula to find how many cents different two frequencies are if we know their frequency ratio. First, divide through by one of the frequencies to get that ratio.

$$\frac{f_2}{f_1} = 2^{(\frac{x}{1200})}$$

Then, take the base 2 logarithm (see the math review chapter for more detail about logarithms) of both sides:

$$\log_2 \frac{f_2}{f_1} = \frac{x}{1200}$$

Or, solving for x , the number of cents difference between the two frequencies,

$$x = 1200 \log_2 \frac{f_2}{f_1}$$

Or, if you prefer standard base-10 logarithms,

$$x \approx 3986 \log_{10} \frac{f_2}{f_1}$$

The approximate 3986 figure is about 3.32 times larger than 1200. Suppose $f_2/f_1 = 10$. The base ten logarithm of this ratio is simply 1, while the base two logarithm is about 3.32. Since the base two logarithm of a given number is 3.32 times larger than the corresponding base ten logarithm, the constant in front must be smaller to ensure the entire quantity is the same in both forms.

I mention cents because it allows us to talk about very small differences in frequencies in a way that is consistent across all possible octaves.

With a quick bit of calculation, we can see the *Pythagorean comma* come up. Recall that a 2:1 ratio in frequencies produces an octave, and a 3:2 ratio produces a fifth. We can start at a specific note and proceed upward by both of these intervals. The number next to each note name represents the standard octave number it is associated with. For this example, we will start with the note “A1,” which has a frequency of 55 Hz. In the first table, I go up by octaves, and in the second, I go by fifths.

Note	Frequency (Hz)
A1	55
A2	110
A3	220
A4	440
A5	880
A6	1760
A7	3520
A8	7040

Starting from 55 Hz, proceeding upwards in octaves

Note	Frequency (Hz)	Note	Frequency (Hz)
A1	55	A#5	939.73
E2	82.5	E#6	1409.59
B2	123.75	B#7	2114.38
F#3	185.62	F## (G7)	3171.58
C#4	278.44	C## (D8)	4757.37
G#4	417.66	G## (A8)	7136.04
D#5	626.48		

Starting from 55 Hz, proceeding upward in fifths (3:2 ratios)

Now, the last entry in each of these tables should interest you. Whether we move upward in octaves or perfect fifths, ideally, we would like the A8 note to be defined by the same frequency either way.

However, moving up to A8 two different ways yields two different frequencies. This frequency discrepancy is the *Pythagorean comma*. In cents, the difference between 7136 Hz and 7040 Hz is

$$x \approx 3986 \log_{10} \frac{7136 \text{ Hz}}{7040 \text{ Hz}} \approx 23 \text{ cents}$$

That is a perceptible difference. We get this value (actually, the comma is more like 23.5 cents) also by studying semitones. We will use the charts above.

We would expect all semitones to be 100 cents by definition, but this is not always the case. It turns out there are two different ways of defining a semitone. One is the *diatonic* semitone, so called because the two notes would be written on different staff position, such as the C to D-flat semitone. We have such a semitone depicted in the tables above. Note that the A4 three octaves above 55 Hz is 440 Hz and G#4 located five perfect fifths above 55 Hz is 417.66 Hz. These two notes are 90.2 cents apart, not 100.

Similarly, we can define a *chromatic* semitone as a semitone where the notes would be written in the same staff position and differentiated only by accidentals (sharps or flats) added. A5 and A#5 are a chromatic semitone interval apart. From the tables above, the two frequencies are 880 Hz and 939.75 Hz. This gap (check it!) is 113.7 cents.

If we define a diatonic semitone as 90.2 cents and a chromatic semitone as 113.7 cents, that means that a given D-flat is 90.2 cents above the C next to it, but a C-sharp is 113.7 cents above the C. Therefore, the D-flat and C-sharp are different by 23.5 cents – the Pythagorean comma!

One thing the 12-TET system does is flat perfect fifths ever so slightly from the Just counterpart. A fifth represents seven semitones, and $2^{7/12} = 1.498$, not the 1.5 you may expect. If we start with a base frequency of A1 = 55 Hz, the 12-TET system gives a frequency for E2 equal to 82.41 Hz. The Just and Pythagorean systems use 82.5 Hz for the same note. These two frequencies are different by just under two cents. To get the Pythagorean comma in the tables above, I used twelve perfect fifth intervals, over which a discrepancy of just under 24 cents builds up. Again, that's the Pythagorean comma!

Microtones

Many instruments can be played in ways not restricted to the 12-tone Western system. Stringed instruments without frets can hit any frequency they like, the human voice can glide smoothly up and down or jump in smaller steps than a semitone, and tricks exist on wind instruments (rolling the mouthpiece of a flute in and out, for example) to access these pitches.

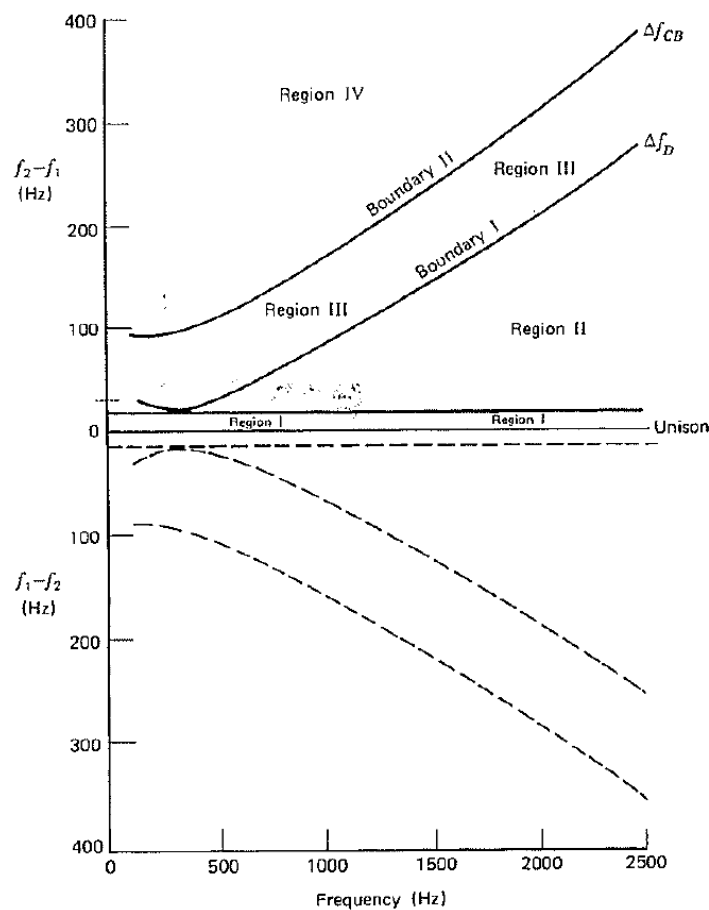
Music that consciously uses these smaller intervals is called *microtonal*. It may use symbols for “half-flat” and “half-sharp” representing fifty-cent differences from standard 12-TET notes. The microtones in that case either use different symbols (see a video about this [here](#)) or are written in standard musical notation with the concert pitch of A-natural adjusted to fit a microtone (see the 5:45 timestamp on [this June Lee transcription](#) of Jacob Collier’s “Moon River” to see such a modulation). It also includes systems where the scale is divided into more pieces. A recent favorite of mine is [this album](#) that uses 17-TET.

Seeing it performed live, it was a headache for keyboard programming, but the stringed instruments (though fretted more frequently) looked to be played the normal way.

Beats revisited

Earlier in these course notes, the phenomenon of *beats* was introduced and demonstrated. To review, if two frequencies f_1 and f_2 are played that are close to one another (we'll nail down what "close" means soon), we perceived a *fused tone* of a frequency equal to the average $\frac{f_1+f_2}{2}$ that exhibits beats with a beat frequency equal to $|f_2 - f_1|$. For example, if 200 Hz and 202 Hz tones are played, a fused tone of frequency 201 Hz is heard that beats two times per second.

This phenomenon does not hold up for any two frequencies. Beats as described earlier only are explicitly perceivable up to a beat frequency of about 15 Hz, and what we perceive changes as the difference between frequencies increases above that. To discuss the details, I refer to the following figure from Rigden's *Physics and the Sound of Music* (Wiley, 1985):



A figure from Rigden (1985) showing four regions (beats, rough fused, rough separate, and clearly distinguished) of perception difference as two frequencies become more different

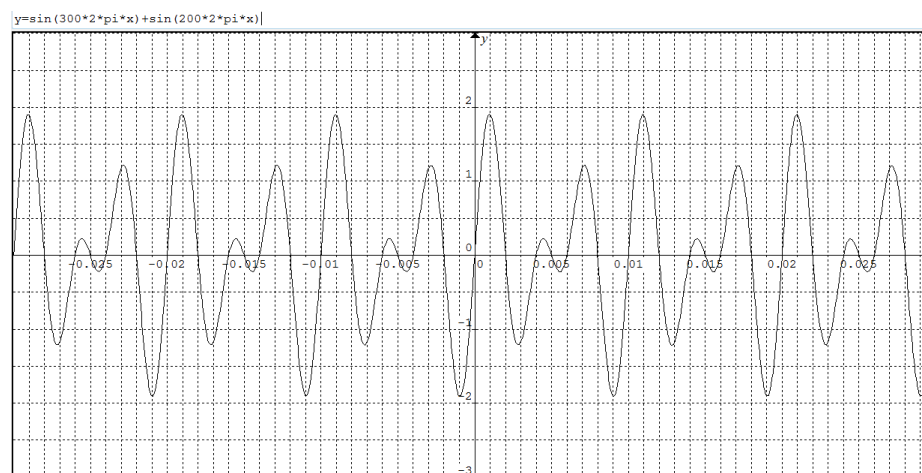
On the x-axis, we see the frequency of a tone f_1 . The y-axis shows the difference between that frequency and a second frequency f_2 . Everything we've discussed about beats up to this point exists in what's labeled here as Region I. Above a beat frequency of 15 Hz (represented by the two horizontal lines above and below the x-axis), we still perceive the fused tone but don't hear individual beats. In the literature, this is called a region of *roughness*, and is labeled above as Region II. "Beating" is still going on, but we perceive this more as a tonal quality rather than explicit swells. This occurs up to what's called Δf_D , or the *limit of frequency discrimination*. If the two frequencies are more different than this, we can hear the two frequencies separately.

Some concrete numbers will help. Suppose f_1 is 1000 Hz. Here, the limit of frequency discrimination is about 80 Hz. So if f_2 is, say, 1060 Hz, the frequency difference is 60 Hz and we'll be in Region II. Therefore, we'll hear a rough fused tone of frequency 1030 Hz. If f_2 is increased to 1120 Hz, the frequency difference is 120 Hz and, being in Region III, we will hear both pitches separately with some roughness.

Looking at another example, suppose f_1 is 2000 Hz. At this frequency, Boundary I happens around a difference of ~ 210 Hz. That means f_2 has to be at least 2210 Hz for the two pitches to be separately perceivable.

The second boundary Δf_{CB} , or the *critical band* shown on the figure, tells us where Region IV is. In this area, we can hear two pitches separately with no roughness. Physiologically what happens here is that if two frequencies are different enough, they excite locations along the basilar membrane of the cochlea far enough apart that no confusion along that membrane occurs.

A further interesting phenomenon happens when two frequencies are played that are harmonics of a common fundamental. For simplicity, I will add identical signals of frequencies 200 Hz and 300 Hz. These two frequencies produce a perfect fifth when played together, but note what happens when the sine waves are combined:



Adding 200 Hz and 300 Hz – this wave has a frequency 100 Hz

The frequency difference is 100 Hz, which according to the Rigden figure above puts us near the boundaries of Regions III and IV – so we hear these distinctly. However, the superposed waveform shown here has a repeating pattern, and the period is 0.01 s! This is equivalent to a frequency of 100 Hz. Practically, what this means is if you play 200 Hz and 300 Hz together and they are loud (>75 dB is enough), you can perceive another 100 Hz difference tone. Other combination tones are possible, but for our purposes here I don't need further detail.

Summary

- Different musical pitches can be assigned specific frequencies.
- There are several different tuning systems or *temperaments*.
- Pythagoras based his system on three ratios (2:3, 3:4, 1:2) and derived the rest of the scale from these.
- A Just tuning system uses small whole-number ratios for all the notes in the scale, making all the intervals in a scale sound harmonious, but it does not allow for simple key changes.
- Attempts to include semitones in the Just system (or even duplicate intervals using only fifths and octaves) leads to a gap called the Pythagorean comma. This makes defining certain semitonal intervals difficult to do consistently (for example, C-sharp and D-flat end up being different notes!)
- The equal-tempered scale (or 12-TET system) does not use simple whole-number ratios, but divides the scale up into twelve equal steps. To get from one semitone to the next, multiply your frequency by the twelfth root of 2.
- The “cent” is one of a hundred subdivisions of each semitone in the 12-TET system.
- Two frequencies played together can exhibit beats, rough fused tones, rough separate tones, or clearly separate tones. If the two frequencies are harmonics of a common fundamental, the fundamental can be heard.

Homework problems

1. Aliens come to Earth and find the 300 Hz frequency most pleasing (to us, that's just a touch sharp of a high D note). Find the eight frequencies (Do-Re-Mi-Fa, etc up to Do) for a scale starting on this note in the Pythagorean scale.
2. Do the same thing in problem 1, but for the Just scale. This will let us show off Western Earth music starting from a note pleasing to them.
3. A standard concert "A" is 440 Hz. (a) Using the equal tempered scale, find the frequency of all the notes in the A minor scale. The scale consists of the following notes: A B C D E F G A, and recall that each of these gaps are two semitones except for the B-C and E-F intervals which are a single semitone. Include at least one digit past the decimal point in each frequency. (b) Middle C has a frequency of 261.6 Hz and it lies one octave below the C you computed in part (a) – use this fact to check your C frequency. (c) The key of A major has three sharps in the key signature, meaning it contains the notes C#, F#, and G# - these are a semitone higher than C, F, and G respectively. Calculate the frequencies of these notes.
4. Use #3 to help. If you play an E above A-440 (equal tempered scale) at the same time as you play an E above A-440 (Just tuning), how many beats will you hear between them in 10 seconds? E is a fifth above A.
5. If I'm tuning my A string (which is supposed to be 110 Hz) on a guitar and the tuner tells me I am 18 cents flat, what frequency am I really producing? How many beats would I hear between this string and a 110 Hz tuning fork every second?
6. An instrument produces a 225 Hz tone when the desired frequency is 220 Hz. How many cents sharp is that?

For the next few problems, refer to the Rigden figure in the coursepack under the "Beats revisited" section.

7. (a) C is 261.6 Hz and D is 293.7 Hz. What do you hear if these are played together? (b) What if you play C with a G (392 Hz)? (c) Are either of these beat frequencies significantly related to the notes being played?
8. (a) Two soft tones of frequencies 1000 Hz and 1040 Hz are played. What is perceived? (b) Two soft tones of frequencies 1500 Hz and 1700 Hz are played. What is perceived? (c) At a frequency of 750 Hz, what are the frequency differences at Boundary I and Boundary II?