This note's for you: A mathematical temperament

Edward G. Dunne

September 15, 2000

It is an old (and well-understood) problem in music that you can't tune a piano perfectly. To understand why takes a tiny bit of mathematics and a smattering of physics.

1 The Physics

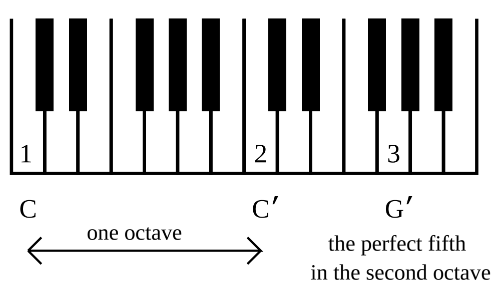
Let me begin by explaining the way a scale is constructed. To avoid sharps and flats (and to make the diagrams easier to draw), I'll use the key of C. So-called middle C represents a particular frequency. There are various standards for fixing the starting frequency. For string players in the U.S., it is commonly done by fixing an A at 440 Hz. Elsewhere, other pitches are common. I will avoid the question entirely (almost) by using the old trick of defining my units so that my middle C has a frequency of **1**.

The area of physics in play here is *acoustics*. There are two rules to begin with:

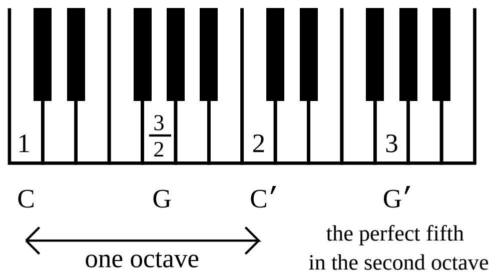
The Piano Axioms

1. Going up one octave doubles the frequency.
2. Tripling the frequency moves to the perfect fifth in the next octave.

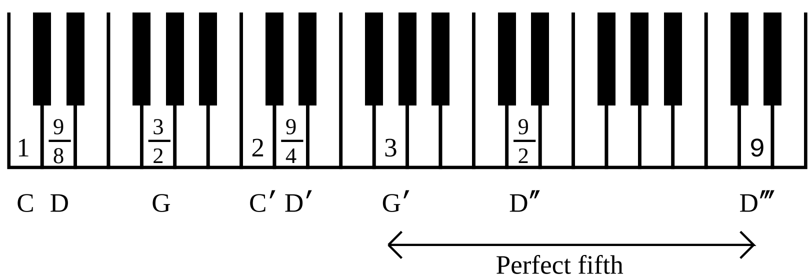
Axiom 1 implies that the C one octave up from middle C has a frequency of 2. Axiom 2 says that in our case the G in the next octave has a frequency of 3. In what follows, I will add a prime for each octave above middle C.



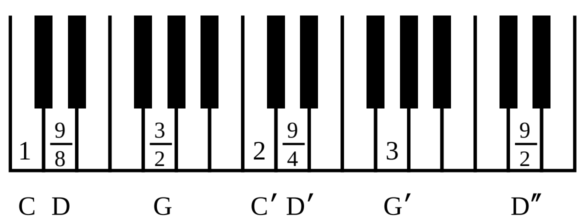
By inverting the rule that says that the note one octave above another must have double the frequency, we can fill-in the perfect fifth in the first octave. It should have half the frequency of the G in the second octave.



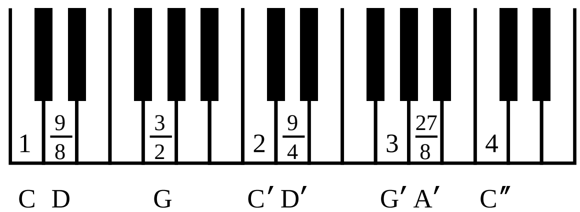
Following Pythagoras, we can now attempt to use these two rules to construct ‘all the notes’, i.e., a complete octave. The perfect fifth in the key of G is D. Thus we have, by tripling then halving, then halving again:



Note that we could have started by tripling G $= 3/2$ to obtain D$'' = 9/2$, which would have saved us a step.



Now repeat again: the perfect fifth in the key of D is A. Thus, by tripling $D = 9/8$, we arrive at $A' = 27/8$.



We can shorten this by looking at the Table of Fifths, also known as the Circle of Fifths (but then you need to know how to draw circular tables in LaTeX).

The Circle of Fifths

Tonic				**Fifth**			
C	D	E	F	**G**	A	B	**C**
G	A	B	C	**D**	E	F#	**G**
D	E	F#	G	**A**	B	C#	**D**
A	B	C#	D	**E**	F#	G#	**A**
E	F#	G#	A	**B**	C#	D#	**E**
B	C#	D#	E	**F#**	G#	A#	**B**
F#	G#	A#	B	**C#**	D#	E#=F	**F#**
C#	D#	E#=F	F#	**G#=A♭**	A#	B#=C	**C#**
G#=A♭	A#	B#=C	C#	**D#=E♭**	E#=F	G	**G#**
D#=E♭	E#=F	G	G#	**A#=B♭**	B#=C	D	**D#**
A#=B♭	B#=C	D	D#	**E#=F**	G	A	**A#**
E#=F	G	A	A#	**B#=C**	D	E	**E#=F**
C	D	E	F	**G**	A	B	**C**

If we use the rule of doubling/halving for octaves, we arrive at the following frequencies for the twelve notes in our basic octave:

Frequency	Tonic
1	**C**
3/2	**G**
9/8	**D**
27/16	**A**
81/64	**E**
243/128	**B**
729/512	**F#**
2187/1024	**C#**
6561/4096	**G#=A♭**
19683/8192	**D#=E♭**
59049/32768	**A#=B♭**
177147/131042	**E#=F**
531441/262144	**C**

However, there are some rules of acoustics that might also be used.

- The frequency of the *perfect fifth* is 3/2 that of the tonic.
- The frequency of the tonic at the end of the *octave* is twice that of the original tonic.
- The frequency of the *perfect fourth* is 4/3 that of the tonic.
- The frequency of the *major third* is 5/4 that of the tonic.
- The frequency of the *minor third* is 6/5 that of the tonic.

Note that the rules for the last three rules are *not* consequences of our two axioms. Using these rules and combining them as efficiently as possible, one arrives at the following list of frequencies for the notes in the C major scale.

Note	Acoustics	Up by fifths, down by octaves
C	1	1
D	$9/8$	$9/8$
E	$5/4$	$81/64$
F	$4/3$	$177147/131042$
G	$3/2$	$3/2$
A	$5/3$	$27/16$
B	$15/8$	$243/128$
C	2	$531441/262144$

Notice that some of these fractions are not equal! In particular, the final C in the scale ought to have frequency twice the basic C. Instead, if we go waaaay up by fifths, then back down again by octaves, we have this strange fraction, $431441/262144$, whose decimal expansion is (exactly): $2.02728652954101 5625$. If we took half of this to return to our starting point, we would have:

$$531441/524288 = 1.0136432647705078125$$

This discrepancy is known as the *Pythagorean (or ditonic) Comma*. So what is the problem? To answer this, it is time to consider some mathematics.

2 The Mathematics

The essence of the comparison is that we went up twelve perfect fifths, which is equivalent to changing the starting frequency from **1** to $\left(\frac{3}{2}\right)^{12} = \frac{5314411}{4096} = 129.746337890625$. This should produce another copy of the note **C**, but seven octaves up. Thus, we should compare this frequency with $2^7 = 128$. The problem is that we are mixing a function based on tripling (for the fifths) with a function based on doubling (for octaves). More abstractly, we are trying to solve an equation of the type: $2^x = 3^y$ where x and y are rational numbers. (With minor finagling, we could restrict to just integers.) Notice that for different notes in our chromatic scale (cf. §5), we will be using different (and inequivalent) values of x and y. The first issue to contend with regarding the difficulty of notes not agreeing with themselves (that is to say, enharmonics that have different frequencies) is to make a choice of where to concentrate the errors. There are ways of tuning an instrument so that some keys have only slight problems, while other keys have rather bad discrepancies. (See Section 4.4 below.)

2.1 Equal Temperament

The method that western music has adopted is to use the system of **equal temperament** (also known as *even temperament*) whereby the ratio of the frequencies of any two adjacent notes of the chromatic scale (i.e. **semitones**) is constant, with the only interval that is acoustically correct being the octave. It is not clear when this was originally developed. Bach certainly went a long way to popularize it, writing two series of twenty-four preludes and fugues for keyboard in each of the twelve major and twelve minor keys. These are known as the *Well-tempered Clavier*. Some people claim that Bach actually invented the system of even temperament. However, guitars in Spain were evenly tempered at least as early as 1482, long before Bach was born. Beethoven also wrote works that took advantage of equal temperament, for instance, his Opus 39 (1803) *Two preludes through the twelve major keys* for piano or organ.

Even temperament spreads the error around in two ways.

1. The errors in any particular key are more or less evenly spread about.

2. No keys are better off than any others. With alternative means of tempering, such as just temperament or **mean temperament** (cf. §4.4), roughly four (out of a possible twelve) major keys are clearly better than the others.

Since western music has settled on a **chromatic scale** consisting of twelve **semitones**, we can compute the necessary ratio, r. Since twelve intervals will make an octave, we must have

$$r^{12} = 2 \qquad \Longrightarrow \qquad r = \sqrt[12]{2} = 2^{1/12}\,.$$

This leaves us with the following values for our C-major scale. (The table includes a comparison with acoustical values.)

Note	Acoustics	Equal Temperament
C	1	1
D	$9/8 = 1.125$	$r^2 \approx 1.12246$
E	$5/4 = 1.25$	$r^4 \approx 1.25992$
F	$4/3 = 1.333\ldots$	$r^5 \approx 1.33484$
G	$3/2 = 1.5$	$r^7 \approx 1.49831$
A	$5/3 = 1.666\ldots$	$r^9 \approx 1.68179$
B	$15/8 = 1.875$	$r^{11} \approx 1.88775$
C	2	$r^{12} = 2$

If we do this with a "440" A and adjust so that the tonic C is acoustically correct with respect to this A, we arrive at the following table:

Note	Acoustics	Equal Temperament
C	264 Hz	$440/r^9 \approx 261.626$
D	297 Hz	$440/r^7 \approx 293.665$
E	330 Hz	$440/r^5 \approx 329.628$
F	352 Hz	$440/r^4 \approx 349.228$
G	396 Hz	$440/r^2 \approx 391.995$
A	440 Hz	440
B	495 Hz	$440r^2 \approx 493.883$
C	528 Hz	$440r^3 \approx 523.251$

It is natural to use a logarithmic scale for measuring intervals in our musical/acoustical setting, since intervals correspond to ratios. The basic unit in our diatonic scale, the semitone, in equal tempering is equal to *100 cents*. Thus, one semitone equals 100 cents and an octave equals 1200 cents. We can measure the Pythagorean comma in terms of cents. The discrepancy was:

$$\begin{aligned} \left(\frac{3}{2}\right)^{12} &\approx 129.746... \\ 2^7 &= 128 \end{aligned}$$

This means that *cents* are measured in a logarithmic scale with base the twelfth root of two. (Really!) Therefore, our discrepancy, in hundreds of cents is:

$$\log_{\sqrt[12]{2}}\left(\frac{(3/2)^{12}}{2^7}\right) = \frac{1}{\ln(\sqrt[12]{2})}\ln\left(\frac{(3/2)^{12}}{2^7}\right) = \frac{12}{\ln 2}\ln\left(\frac{(3/2)^{12}}{2^7}\right) .$$

After a little algebra, we see that this is equal to

$$144\frac{\ln 3}{\ln 2} - 228 \approx 0.23460010385\ldots \approx 23.5 \text{ cents} .$$

This difference is so small that most people cannot hear it. However, many musicians, particularly string players, are sensitive to such differences. Following up on the algebra of the preceding problem, we see that an interval corresponding to the ratio I equals $1200\frac{\ln(I)}{\ln 2} = 1200\log_2(I)$ cents. This will simplify the formulas given below for the other commas.

There are other commas: The *syntonic (or Didymic) comma* is the difference between four perfect fifths and two octaves plus a major third. The syntonic comma occurs more easily than the Pythagorean comma or the schisma, since one doesn't need to go through particularly many chord progressions to move through four perfect fifths. If you would like to hear what a syntonic comma sounds like, I heartily recommend Erich Neuwirth's book [5] and the software that comes with it. In [2], we show how the syntonic comma arises in a song by Madonna. The *schisma* is the difference between eight perfect fifths plus one major third and five octaves. The *diaschisma* is the difference between four perfect fifths plus two major thirds and three octaves. The computations are given in Section 4. It should be clear at this point that most (indeed almost all) of the acoustic intervals will be imperfect in an equally tempered scale. What is gained, though, is that none of the errors is particularly large. Let me come back to some of the details of even temperament after addressing another issue. Namely, why should we have twelve half steps in an octave anyway?

2.2 Continued Fractions

We will use the logarithm base 2, which we denote by $\log_2(x)$. Thus, if $y = \log_2(x)$, then $2^y = x$. The heart of our problem with fifths and octaves is the attempt to solve the equation $2^x = 3^y$, where x and y are integers or rational numbers. Notice, if we're using rational numbers it is an equivalent problem to solve the equation $2^z = 3$.

If we take logarithms base 2 of both sides of the troublesome equation, we are left with the equation

$$x\log_2(2) = y\log_2(3) .$$

Of course, since $\log_2(2) = 1$, the equation reduces to:

$$x = y\log_2(3) \quad \text{or} \quad \frac{x}{y} = \log_2(3) .$$

We then try to solve this for integer (or rational) values of x and y. Unfortunately, $\log_2(3)$ is not a rational number. The best we can do is to try to approximate it by a rational number. Plugging this into my calculator, I get the decimal approximation: 1.584962500721. (Note: this is an approximation; the real decimal goes on and on and on and on.) A good – and well-known – way to approximate an irrational number by a rational number is by continued fractions.

A **continued fraction** is an expression of the form:

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cdots}}}}$$

where $a_0, a_1, a_2, \ldots$ are integers. Using this form (with only 1s in the numerators) means we will only be considering **simple** continued fractions. For notational convenience, write $[a_0, a_1, a_2, ...]$ for the infinite continued fraction above. Of course, it is also possible to consider finite continued fractions. It is an exercise to see that any rational number can be expressed as a finite continued fraction. I refer you to Hardy and Wright's book [3] for a discussion of the uniqueness of such an expression. If we cut off an infinite continued fraction after N terms, we have the N^{th} **convergent**. For the infinite continued fraction given above, this is

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\cdots + \frac{1}{a_N}}}}}$$

which is denoted $[a_0, a_1, a_2, \ldots, a_N]$. This is obviously a rational number, which we write (in reduced form) as $\dfrac{p_N}{q_N}$. There is a convenient algorithm for computing the continued fraction expansion of a given positive number x, called **the continued fraction algorithm**. For any positive number A, let $[A]$ denote the integer part of A. To compute a continued fraction expansion for x, take $a_0 = [x]$. So

$$x = a_0 + x_0$$

and $0 \leq x_0 < 1$. Now write

$$\frac{1}{x_0} = a_1 + x_1 \quad \text{with} \quad a_1 = [1/x_0]$$

and

$$\frac{1}{x_1} = a_2 + x_2 \quad \text{with} \quad a_2 = [1/x_1]$$

and so on. The numbers a_0, a_1, a_2, etc. are non-negative integers, which go into the continued fraction expansion. The 'remainders', x_0, x_1, x_2, etc., are real numbers with $0 \leq x_j < 1$.

2.3 Some examples

In what follows, the notation for a repeating continued fraction is similar to that for a repeating decimal expression. For the continued fraction $[n, m, m, m, m, \ldots]$, we write $[n, \dot{m}]$.

- $\sqrt{2} = [1, 2, 2, 2, ...] = [1, \dot{2}]$ with convergents:

$$1 \quad \frac{3}{2} \quad \frac{7}{5} \quad \frac{17}{12} \quad \frac{41}{29} \quad \ldots$$

- $\dfrac{1+\sqrt{5}}{2} = [1,1,1,1,1,...] = [\dot{1}]$ with convergents:

$$1 \quad 2 \quad \frac{3}{2} \quad \frac{5}{3} \quad \frac{8}{5} \quad \frac{13}{8} \quad \frac{21}{13} \quad \frac{34}{21} \quad \frac{55}{34} \quad \ldots$$

  Notice that the numerators and denominators of the convergents are successive Fibonacci numbers. If you start simplifying the convergents of the continued fraction $[1,1,1,1,...]$ as a rational number, you will soon see why this is so. Try it! Also, it is a well-known fact that the ratio of successive Fibonacci numbers is indeed the **golden mean** $\frac{1+\sqrt{5}}{2}$. That is to say, this particular continued fraction does indeed converge to the irrational number it is supposed to represent.

- $e = [2,1,2,1,1,4,1,1,6,1,1,8,1,1,10,...]$ with convergents:

$$2 \quad 3 \quad \frac{8}{3} \quad \frac{11}{4} \quad \frac{19}{7} \quad \frac{87}{32} \quad \frac{106}{39} \quad \frac{193}{71} \quad \frac{1264}{465} \quad \frac{1457}{536} \quad \frac{2721}{1001} \quad \ldots$$

  Remarkably, the pattern in the continued fraction expansion for e actually persists. Perhaps less remarkably: it is not particularly easy to prove this.

- $\pi = [3,7,15,1,292,1,1,1,2,...]$ with convergents:

$$3 \quad \frac{22}{7} \quad \frac{333}{106} \quad \frac{355}{113} \quad \frac{103993}{33102} \quad \frac{104348}{33215} \quad \frac{208341}{66317} \quad \frac{312689}{99532} \quad \frac{833719}{265381} \quad \ldots$$

  The zero-th convergent is familiar to certain state legislatures. The first convergent is familiar to many school children, at least it was known before the dawn of handheld calculators. Unlike the continued fraction expansion for e, the complete expansion for π is unknown.

There are two theorems that are pertinent to our discussion (cf. [3]):

Theorem 1 If x is an irrational number with continued fraction expansion $[a_0, a_1, a_2, \ldots]$ and convergents $\{p_k/q_k\}$ and $n \geq 1$, then

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}} < \frac{1}{{q_n}^2}\,.$$

Since the denominator of the $(n+1)^{\text{st}}$ convergent is strictly larger than the denominator of the n^{th} convergent (and they are all integers), we see that the continued fraction expansion does indeed converge to the irrational number it is meant to be approximating.

Theorem 2 If x is an irrational number, with continued fraction expansion $[a_0, a_1, a_2, \ldots]$ and convergents $\{p_k/q_k\}$ and $n \geq 1$, and $0 < q \leq q_n$ and $\frac{p}{q} \neq \frac{p_n}{q_n}$ with p, q integers, then

$$\left| x - \frac{p_n}{q_n} \right| < \left| x - \frac{p}{q} \right|\,.$$

That is to say, the n^{th} convergent is the fraction among all fractions with denominators no greater than q_n which provides the best approximation to x. It is common to use the size of the denominator as a measure of the 'complexity' of the rational number. Thus, we have that the n^{th} convergent is optimal for a given complexity.

What does this say for our musical problem? Recall that the troublesome equation $2^x = 3^y$ is equivalent to the equation $2^z = 3$, provided we use rationals, and not just integers. The 'obvious' solution is

$z = \log_2(3)$. We want to approximate this by a rational number. By direct computation, the continued fraction expansion for $\log_2(3)$ is

$$[1, 1, 1, 2, 2, 3, 1, 5, 2, 23, 2, 2, 1, ...] \, .$$

The first few convergents are:

$$1 \quad 2 \quad \frac{3}{2} \quad \frac{8}{5} \quad \frac{19}{12} \quad \frac{65}{41} \quad \frac{84}{53} \quad \frac{485}{306} \quad \ldots$$

Thus, taking the fourth approximation (start counting at zero):

$$3 = 2^{\log_2(3)} \approx 2^{19/12} \, .$$

That is to say, we obtain the perfect fifth, one octave up, by nineteen semitones. Moreover, the denominator, being twelve, forces us to have twelve semitones per octave. Thus, western music has adopted, quite by accident I assume, the fourth best approximation to a Pythagorean scale using equal temperament.

Obviously it is possible to have scales that come from dividing an octave into other than twelve pieces. For instance, one common Chinese scale has five notes to the octave. This corresponds to the third convergent of the continued fraction expansion. An accident? I don't know.

Going in the other direction, we could use the next more accurate continued fraction approximation of $\log_2(3)$, which would lead to an octave consisting of forty-one pieces. Below is a comparison of what happens to some standard intervals in these three systems. The fundamental interval for our standard twelve-tone chromatic scale is the semitone. There is no name for the basic intervals of our other chromatic scales. So I will refer to their basic intervals merely as 'basic intervals'.

In the following computations, we will repeatedly use

$$\log_{\sqrt[n]{2}}(x) = \frac{\ln(x)}{\ln(2^{1/n})} = n\frac{\ln(x)}{\ln(2)} = n \log_2(x) \, ,$$

which includes the formula for computing logarithms in any base in terms of natural logarithms.

Now, if we compute *exactly* in a twelve-tone scale, we find:

- The fifth is $\log_{\sqrt[12]{2}}(3/2) = 12 \log_2(3/2) \approx 7.01955 \ldots \approx 7$ basic intervals (semitones).
- The major third is $12 \log_2(5/4) \approx 3.8631 \ldots \approx 4$ basic intervals (semitones).
- The minor third is $12 \log_2(6/5) \approx 3.1564 \ldots \approx 3$ basic intervals (semitones).

If we used a five-tone scale, the computation of the number of basic intervals corresponding to an interval with a ratio of I is $\log_{\sqrt[5]{2}}(I) = 5 \log_2(I)$. Thus we find:

- The fifth being $5 \log_2(3/2) \approx 2.924812 \ldots \approx 3$ basic intervals.
- The major third being $5 \log_2(5/4) \approx 1.60964 \ldots \approx 2$? basic intervals.
- The fourth being $5 \log_2(4/3) \approx 2.075187 \ldots \approx 2$ basic intervals.
- The minor third being $5 \log_2(6/5) \approx 1.31517 \ldots \approx 1$? basic intervals.

Thus, the major third and the perfect fourth would be indistinguishable in an equal-tempered five-tone scale. Here, the '?' for the major third and minor third indicate that the rounding to the nearest integer is fairly inaccurate.

If we used forty-one semitones per octave, the computation is $\log_{\sqrt[41]{2}}(I) = 41 \log_2(I)$, leading to:

- The fifth being $41 \log_2(3/2) \approx 23.98346253 \ldots \approx 24$ basic intervals.
- The major third being $41 \log_2(5/4) \approx 13.19905189 \ldots \approx 13$ basic intervals.
- The fourth being $41 \log_2(4/3) \approx 17.01653745 \ldots \approx 17$ basic intervals.
- The minor third being $41 \log_2(6/5) \approx 10.78441064 \ldots \approx 11$ basic intervals.

This type of scale has a fairly good separation of the standard acoustically distinct notes. I would guess that if we used such a scale, our ears would be trained to hear the difference between adjacent basic intervals. However, this difference is only $\sqrt[41]{2} = 2^{1/41} \approx 1.0170497 \ldots$, which is 29 cents, only slightly more than the Pythagorean comma.

Interestingly, around 40 B.C., King Fang, in China, discovered the sixth best approximation given above. It is unlikely, of course, that he actually used continued fractions to do this, which makes it all the more remarkable. In particular, Fang noticed that fifty-three perfect fifths are very nearly equal to thirty-one octaves. This leads to what is sometimes called the *Cycle of 53*. It can be represented by a spiral of fifths, replacing the more usual circle of fifths.

3 The Pythagorean Hammers

Western music has adopted certain intervals as basic to acoustics. The legend about the source of some of these intervals involves Pythagoras. The story has him listening to the sound of the hammers of four smiths, which he found to be quite pleasant. Upon investigation, the hammers weighed 12, 9, 8, and 6 pounds. From these weights, Pythagoras derived the intervals:

The octave:	12:6 = 2:1
The perfect fifth:	12:8 = 9:6 = 3:2
The perfect fourth:	12:9 = 8:6 = 4:3
The whole step:	9:8

I don't know. Maybe. It's hard to say what really happened twenty-six centuries ago. But this certainly seems lucky. Maybe he was sitting in the same bath tub that Archimedes was sitting in four hundred years later. Also, it is not clear whether the hammers control the tone or if it's the anvils that matter. In the present, we can look to see what might be natural intervals to construct. Firstly, the octave is quite natural, as a doubling of frequency. As usual, we will also take its inverse, halving of frequency, as equally natural. The next integral multiplication of frequency is tripling, which leads to the perfect fifth when combined with halving. Multiplying the frequency by four is just going up two octaves, so we already have that in our system. The next natural operation is to multiply the frequency by five. To remain in the original octave, we need to combine this with two halvings, leading to the interval of the major third.

Now it is not simply a preference for integers that leads to these intervals. There is also the phenomenon of *overtones*. A vibrating string has a fundamental tone, whose frequency f can be calculated from its length L, density ρ and tension T according to a basic formula of acoustics:

$$f_1 = \frac{1}{2L}\sqrt{\frac{T}{\rho}} = \frac{c}{2L}$$

where $c = \sqrt{T/\rho}$ is the speed with which the wave travels along the string.

The string also vibrates in other modes with less intensity. The existence of these other modes can be deduced mathematically, by looking at the eigenvalues of differential operators. This is discussed in almost PDE textbook. You could also consult Knobel's little book [4]. From either the mathematics or the physics, we discover that these other modes are vibrations at integer multiples of the fundamental frequency. The increasing sequence of such frequencies is called the *harmonic series* based on the given fundamental frequency. The fundamental frequencey is also called the *first harmonic* or the *first mode of vibration*. The frequency of the octave (twice that of the fundamental) is called the *second harmonic* or the *second mode of vibration*. The *third harmonic* is the perfect fifth one octave up from the fundamental. And so it goes. Thus, the argument for preferring intervals based on doubling, tripling and multiplying by five is actually based on acoustics, not just a fondness for the numbers 2, 3 and 5.

The phenomenon of overtones is an important factor in the quality of the sound of any particular instrument. Now, in theory, it may appear that the harmonic series for a particular fundamental frequency continues through all the integers. However, this would surely produce unbearable dissonance. What actually happens is that the intensity of the higher harmonics decreases quite rapidly. Indeed, on some instruments it is difficult to discern beyond the third harmonic. (My guitar, for instance.) Violins and oboes have strong higher harmonics, leading to a 'bright' tone. Flutes and recorders have weak higher harmonics. Apparently the clarinet has strong odd-numbered harmonics, which is why it has a 'hollow' tone. Before valves were added to brass instruments, it was only notes corresponding to harmonics that could be played on these instruments.

When defining the basic acoustical intervals, after the intervals based on multiplying by two, three and five, our choices become more arbitrary.

- The perfect fourth. Should we go down a perfect fifth then up an octave, resulting in an interval of $(3/2)(2) = 4/3$? Or should we do something else? (Question: How is the perfect fourth related to the perfect fifth?)
- The whole tone. Why is it better to go up two perfect fifths and down an octave, that is, $(3/2)(3/2)(1/2) = 9/8$, rather than, say, up two fifths and down three major thirds: $(3/2)(3/2) \cdot (4/5)(4/5)(4/5) = 144/125$? (There is a difference of about 41 cents here.)
- The minor third. Should we use $(4/5)(3/2) = 6/5$, i.e. down a major third and up a perfect fifth, or $(2/3)(2/3)(2/3)(2)(2) = 32/27$, i.e. down three fifths and up two octaves?
- The major third. One could even argue that $(3/2)(3/2)(3/2)(3/2)(1/2)(1/2) = 81/64$ is preferable to $5/4$, as the former is obtained by going up four perfect fifths then down two octaves, thus using only the doubling and tripling rules.

For the sake of curiosity, we could investigate what we obtain using the *major third* as the basis for our computations. The acoustic major third is $5/4$. Thus, the critical quantity is $\log_2(5/4) = \log_2(5) - \log_2(4)$. Since $\log_2(4)$ is an integer, the crux of the approximation is that of $\log_2(5)$. The continued fraction expansion is

$$[2, 3, 9, 2, 2, 4, 6, 2, 1, 1, 3, 1, 18]\,.$$

The convergents are:

$$\frac{7}{3} \quad \frac{65}{28} \quad \frac{137}{59} \quad \frac{339}{146} \quad \frac{1493}{643} \quad \ldots$$

Since three notes are certainly too few for an octave, we would have been stuck with octaves of twenty-eight notes! I think I'll stick with the perfect fifth and twelve tones per octave.

4 Some Other Commas

4.1 Syntonic Comma

The *syntonic (or Didymic)* comma is the difference between four perfect fifths and two octaves plus a major third. Four perfect fifths correspond to $(3/2)^4$. In the key of C, this is C $\longrightarrow$ E$''$. Two octaves plus a major third correspond to $2^2\,(5/4)$. In the key of C this corresponds to C $\longrightarrow$ C$''$ $\longrightarrow$ E$''$. Compare the two frequencies using the logarithmic scale, to obtain

$$\log_{\sqrt[12]{2}}\left(\frac{(3/2)^4}{2^2(5/4)}\right) \approx 0.215062895967\ldots \approx 21.5 \text{ cents.}$$

4.2 Schisma

The *schisma* is the difference between eight perfect fifths plus one major third and five octaves. Eight perfect fifths plus one major third correspond to $(3/2)^8\,(5/4)$. In the key of C this is

$$C \to G \to D' \to A' \to E'' \to B'' \to F'''\,\# \to C''''\,\# \to G''''\,\# \to C'''''$$

versus just the jump of five octaves from C to C''''', which corresponds to $2^5 = 32$. If we compare the two frequencies using our logarithmic scale, we obtain

$$\log_{\sqrt[12]{2}}\left(\frac{(3/2)^8\,(5/4)}{32}\right) \approx 0.019537207879\ldots \approx 2 \text{ cents.}$$

4.3 Diaschisma

The *diaschisma* is the difference between four perfect fifths plus two major thirds and three octaves. In the key of C this is $C \to G \to D' \to A' \to E'' \to G''\,\# \to C'''$. The computation of the comparison boils down to:

$$\log_{\sqrt[12]{2}}\left(\frac{2^3}{(3/2)^4(5/4)^2}\right) \approx 0.1955\ldots \approx 20 \text{ cents.}$$

4.4 Mean-tone system

One alternative to equal temperament is the *mean-tone system*, which seems to have begun around 1500. In mean temperament, the fifth is 697 cents, as opposed to 700 cents in equal temperament or 701.955 cents for the acoustically correct interval. The mean-tone system for tuning a piano is satisfactory in keys that have only one or two sharps or flats. But there are problems. For instance, G$^{\#} = 772$ cents and A$^{\flat} = 814$ cents. But they ought to be the same! This discrepancy is called *the wolf*. While the Pythagorean comma, at 23.5 cents, is not discernible by most listeners, the wolf, at 52 cents is quite noticeable.

Before equal temperament was widely accepted, keyboards had to accommodate these problems. One solution was only to play simple pieces in the keys your instrument could handle. A second solution, which was certainly necessary for large and important organs, was to have divided keyboards. Thus, the single key normally used today for G$^{\#}$ and A$^{\flat}$ would be split into two keys. Often, the back of one key would be slightly raised to improve the organist's ability to play by touch. The most extraordinary keyboard I was able to find a reference to was Bosanquet's 'Generalized Keyboard Harmonium' built in 1876, which had 53 keys per octave!

5 Definitions

Chromatic Scale The chromatic scale contains all the possible pitches in an octave, as opposed to a diatonic scale, which contains combinations of whole tones and semitones. When using octaves divided into other than twelve intervals, the chromatic scale contains all the microtones in the subdivision.

Enharmonics An interval less than a half step.

Major Third For the purposes of this discussion, we take a *major third* to be defined as the interval corresponding to a change in frequency by a factor of $5/4$. In the key of C, this is the interval (approximated by!) C $\rightarrow$ E.

Minor Third For the purposes of this discussion, we take a *minor third* to be defined as the interval corresponding to a change in frequency by a factor of $6/5$. In the key of C, this is the interval (approximated by!) C $\rightarrow$ E$^\flat$.

Octave The interval corresponding to doubling the frequency.

Perfect Fifth The interval corresponding to a change of frequency by a factor of $3/2$. It is the interval separating the fifth note of a major scale from the tonic.

Semitone A *semitone* is the basic interval of the standard octave of western music. That is to say, it is an interval of $2^{1/12} = \sqrt[12]{2}$. For the scales of five, twelve and forty-one notes that are also considered here, the semitone is not quite as useful. Instead, we speak of the 'basic interval'. For the scale obtained by dividing the octave into five pieces, the basic interval is $2^{1/5}$. Generally, intervals that are not obtained from semitones are called *microtones*.

Temperament For our purposes, *temperament* refers to any system of defining the frequencies of the notes in a scale, be it chromatic, diatonic or some other sort of scale.

Tonic The tonic is the first note in a key or scale. It is also the note after which the scale is named, hence, the keynote.

Acknowledgments

My thanks to Mark McConnell for his help with this. An alternative version of this article [2], written jointly with him, appears in *Mathematics Magazine*.

References

[1] Eric Blom (ed.) *Grove's Dictionary of Music and Musicians*, Fifth Edition, St Martin's Press, Inc., 1955.

[2] Edward Dunne and Mark McConnell, Pianos and Continued Fractions, *Mathematics Magazine*, vol. 72, no. 2 (1999), pp. 104-115.

[3] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, Fourth Edition, Oxford University Press, 1975. See, in particular, Chapters X and XI concerning continued fractions.

[4] Roger Knobel, *An Introduction to the Mathematical Theory of Waves*, American Mathematical Society, 2000. See, in particular, Chapters 7, 12, and 14, and Section 11.3.

[5] Erich Neuwirth, *Musical Temperaments*, Springer-Verlag, Vienna, 1997.

[6] Don Michael Randel, *Harvard Concise Dictionary of Music*, Harvard University Press, 1978. See the sections titled 'Comma' and 'Temperament'.

[7] Percy Scholes, *The Oxford Companion to Music*, Ninth Edition, Oxford University Press, 1955. See the section titled 'Temperament'.