Mathematics and Music

In the ancient world, music was considered to be an important subject for theoretical studies, on a level with pure mathematics itself. Boethius' famous *de institutione musica* is hardly concerned with the practical performance of music. It has much more to do with describing (in a very roundabout way) various rather trivial ideas about numbers. For example, he makes much of the fact that the difference of the squares of two adjacent positive integers is the sum of those numbers. Once you have worked your mind around all those words — and Boethius goes on and on through many paragraphs in this fashion — then you see that he is really just describing the simple relationship

$$(n+1)^2 - n^2 = 2n + 1 = (n+1) + n.$$

In more recent centuries, Euler published a system of musical ratios which, although it may have had a certain mathematical elegance, was found to be musically useless.

A modern book which attempts to reconstruct something of the *practical* music theory of the ancient world is "Ancient Greek Music", by M.L. West. Also to be recommended in this connection is "The Mathematics of Plato's Academy", by D.H. Fowler, which discusses the importance of the idea of ratio throughout Greek mathematics — not just in music theory. As far as the physical and mathematical basis of music is concerned, a good reference is the internet website of the Physics Department of the University of New South Wales in Australia. It is "http://www.phys.unsw.edu.au/music/".

What I would like to do here is to explain why music is based on the diatonic scale; that is, the white keys on the piano. Is this convention something which is only peculiar to our culture[1], or is it something which arises naturally from the basic rules of arithmetic? One suspects that the later is the case.

Now it is true that there are "primitive" societies (such as ancient Greece) which use musical scales which are very different from the diatonic scale. But surely these strange scales represent a cultural departure from the normal, natural system given to us by the rules of arithmetic. The simplest way to verify this fact is to observe the reactions of a cat[2] to different kinds of music. Given a simple tune on the diatonic scale, or even the practice of scales on an instrument, then the cat will go to sleep and begin purring if the intonation is good. However if the intonation is off, or if the music leaves the diatonic scale in an unmotivated way, then the cat will fold it's ears back uncomfortably and try to leave the room as quickly as possible.

[1] Many composers in the 20th century seem to have become irritated by the fact that there are lots of black keys in there as well. In a kind of "democracy of the keys" of the piano, they created 12-tone music, where the different keys were depressed on the average for about an equal period of time. Most people find such "music" to be somewhat offensive to the ear.

[2] Dogs are not good subjects for this experiment since, in order to please their masters, they might pretend to like something which they don't.

1 The Harmonic Series

The difference between noise and music is that a musical tone is a *periodic* sound wave. This sound wave can be represented as a continuous function, and therefore we can use the theorem on Fourier series from the Analysis I lecture, which states that any continuous periodic function can be represented as a sum of sine (and cosine) functions whose periods are whole number fractions of the period of the function. The relative strengths of these "overtones" determine the sound of the tone.

Put another way, one can say that the tone can be thought of as being produced by a number of "pure" instruments (which produce sine waves) which are all playing at frequencies which are exact integer multiples of a basic lowest frequency, say $\mathcal{F}$. Then the first overtone is at a frequency of $2\mathcal{F}$, the second is at $3\mathcal{F}$, and so on. On the other hand, if the intonation of these pure instruments is bad, so that the frequencies are not exact multiples of a basic lowest frequency, then the combined sound tends to degenerate into an unpleasantly jarring noise.

So to summarise: the harmonic series is simply

$$\mathcal{F},\ 2\mathcal{F},\ 3\mathcal{F},\ 4\mathcal{F}, \dots,$$

and two or more "pure" instruments playing together at exactly these frequencies produce a tone which the ear accepts as being something different from noise, hopefully pleasant. To simplify our thoughts, we can normalise things, taking $\mathcal{F} = 1$. Thus the harmonic series is the set $\mathbb{N}$ of natural numbers.

2 Symphona - Diaphona

Let us now consider two "impure" (more human) instruments, each of which produces it's own harmonic series. Let us say that one of the instruments has as it's basic frequency $\mathcal{F}_1$, and the basic frequency of the other is $\mathcal{F}_2$. The simplest case is that $\mathcal{F}_1$ say, is an integer multiple of $\mathcal{F}_2$. So we let $\mathcal{F}_1 = k\mathcal{F}_2$, for some $k \in \mathbb{N}$. Then the harmonic series for the first instrument is

$$n\mathcal{F}_1 = (nk)\mathcal{F}_2,$$

for $n = 1, 2, \dots$. That is, it fits in perfectly with the harmonic series of the second instrument; we have a pure sound.

Things become more complicated if neither $\mathcal{F}_1$ nor $\mathcal{F}_2$ is an integral multiple of the other. Let's consider the case that they are both multiples of some smaller number $\mathcal{F}$. Say $\mathcal{F}_1 = a\mathcal{F}$ and $\mathcal{F}_2 = b\mathcal{F}$, so that

$$\frac{\mathcal{F}_1}{\mathcal{F}_2} = \frac{a}{b}.$$

Obviously it is sensible to assume that this fraction is reduced, so that $gcd(a, b) = 1$. That is, $\mathcal{F}$ is the largest number with this property. But then the combined wave produced by both instruments has a frequency of $\mathcal{F}$. This is the reason that two tones which are close together produce an irritating "beating" noise when played together. $\mathcal{F}$ is the frequency of the beats, and it can take on the feeling of a rapid hammering if is $\mathcal{F}$ is small enough, say around 10 Hertz. When $\mathcal{F}$ is somewhat greater, say around 30 or 40 Hertz, then one simply has the feeling of a clash of tones. Thus, even in the case that the ratio of $\mathcal{F}_1$ to $\mathcal{F}_2$ is a rational number, we still can have a disharmony[3]; if the ratio is irrational, harmony is impossible.

The ancient Greeks identified two distinct cases: symphona and diaphona, that is harmony and disharmony. For them, the ratio 1:2 was clearly symphona (and since one instrument is as good as another, we will say that the ratio 2:1 is symphona as well). But they also realized that a music consisting of nothing but octaves is rather too boring for the human ear. Thus the next simplest ratio, namely 2:3 (and thus 3:2) was also taken to be symphona. All else was declared to be diaphona. Disharmonious.

3 Harmonious Numbers

Given that our basic starting point is some frequency $\mathcal{F}$, which we will just normalise to the number $\mathcal{F} = 1$, then 2, 4, 8, and so on, can all be built up from the simple ratio of 1:2. By the same token, we accept the fractions 1/2, 1/4, and so on. All of these numbers — musical octaves — are to be admitted into the category of symphona.

But then we have 3/2, 9/4, 27/8, ..., owing to the disturbing influence of the ratio 3:2 in our symphona definition. And then if we remember the 1/2 interval, we must also admit the ratios 3/4, 9/8, and so on. In the end, we end up with all ratios of the form $2^m/3^n$ and $3^n/2^m$.

Obviously this gives us some rather close ratios which would not be admitted into practical music. On the other hand, it does allow us to make the rule that we will only think about the musical ratios which lie between 1 and 2. A ratio such as 9/4 can be reduced into this interval by dividing by 2. Thus, in this way of thinking, the harmonious number 9/4 is "really" the number 9/8. (Of course, if we multiply 9/4 by 2/3 to bring us back into the interval 1 to 2, then we return to the simple symphona interval 3/2, which we already had.) So the rule we have is that a given ratio should be brought into the range 1 to 2 by multiplying with some 2^n, or 3^n, for some $n \in \mathbb{Z}$.

4 The Tetrachord

Actually, the ancient Greeks didn't think in terms of all those numbers which we have seen in the last section. In fact, they only thought of the three numbers: 1, 3/2, and 2. This gives the basis of the famous Greek tetrachord. All of Greek music was based on the tetrachord.

[3] One sees that the numbers in a ratio cannot be small if one of them has a large prime factor. In fact, all the prime numbers greater than 5 are bad for music.

Now the ratio of the interval from 1 to 3/2 is obviously just the number 3/2. On the other hand, the interval from 3/2 to 2 is

$$\frac{2}{\frac{3}{2}} = \frac{4}{3}.$$

Therefore, we now have four numbers, 1, 4/3, 3/2 and 2. These give us two large intervals, namely 1 to 4/3, and 3/2 to 2. But one immediately sees that the higher interval here is again

$$\frac{2}{\frac{4}{3}} = \frac{3}{2},$$

so that we have a nice symmetry. The ratio 4/3 defines the bounds of the Greek tetrachord. Within this interval, two further notes are to be placed.

But before doing that, we should note that the smaller interval from 4/3 to 3/2 gives us a new number, namely

$$\frac{\frac{3}{2}}{\frac{4}{3}} = \frac{9}{8}.$$

So already we have four musical intervals: 9/8, 4/3, 3/2 and 2. In the Theory of Music, these are called the second (or a whole note), the fourth, and the fifth, and finally the octave.

Looking at the white keys on the piano, you should try to locate the c-key somewhere around the middle of the keyboard. (It has a group of two black keys directly to the right of it.) Then if we give this key the number 1, we find that the adjacent white key to the right (the *second* one) has the number 9/8, then the *fourth* white key to the right has the number 4/3, and the *fifth* white key has the number 3/2. Finally, the eighth white key to the right has the number 2, and it is simply the c-key an octave higher.

5 What About the Other Three White Keys?

Well, they come into the picture if we allow ourselves to go slightly beyond the bounds of symphona and accept the next prime number, namely 5.[4] Admittedly, we have now entered the realms of diaphona, and so we are constructing the diatonic scale of music. On the other hand, looking at reduced ratios of positive integers, it is clear that after 3/2 and 4/3, the smallest ratios must involve the number 5. If we reduce the number 5 into the interval between 1 and 2 by dividing by 4, then we obtain the number 5/4. Obviously another way to go about things is to divide by 3, thus giving us 5/3.

So now we have the *third* white key to the right of our starting point on the c-key. This is the number 5/4. Similarly, the *sixth* key is 5/3.

[4] It is interesting to note that jazz recognises the "blue" notes, which contain prime numbers greater than 5. For example the interval 7/4 is played in jazz: a note somewhere between the sixth and seventh notes (but not really on a black key).

What about the remaining *seventh* key, at the end of the scale? To understand which number it has, we need only notice that a new ratio has already appeared in the numbers which we have obtained so far. It is namely the case that the third key is 5/4 and the fourth key is 4/3. This gives the ratio

$$\frac{\frac{5}{4}}{\frac{4}{3}} = \frac{15}{16}$$

which when multiplied by 2 gives 15/8. Therefore the seventh key is 15/8.

Note that there are two pairs of white keys which are adjacent to one another, with no black key between them. These are the pair of keys (third, forth) and (seventh, eighth). In numbers, these are the pairs (5/4, 4/3) and (15/8, 2). But we have

$$\frac{\frac{4}{3}}{\frac{5}{4}} = \frac{16}{15} = \frac{2}{\frac{15}{8}},$$

so everything is Ok.

Or perhaps not quite. A little bit of arithmetic shows that there are two kinds of adjacent white keys which have a black key in between. Namely the pairs (1, 9/8), (3/2, 4/3), and (5/3, 15/8) have the ratio 9/8. The other two, namely (9/8, 5/4) and (3/2, 5/3) have the ratio 10/9. Within the Theory of Music, one says that the ratio 9/8 is a *major* whole tone, and the ratio 10/9 is a *minor* whole tone.

So let's summarise the work so far.

The diatonic scale of music.

note	musical interval	ratio
c	—	1
d	second	9/8
e	third	5/4
f	fourth	4/3
g	fifth	3/2
a	sixth	5/3
b	seventh	15/8
c	octave	2

6 Other Scales of Music

One puzzling question which arises here is:

"Why doesn't the scale start with the letter a, rather than c?"

The simple answer to this question is that the scale *does*, in fact, start with a! That is, if we start looking at the white notes on the piano starting at the a-key, then we have the scale of a-*minor*. What we have been thinking about up to now is the scale of c-*major*. In the old days, the minor scales were the big thing. The major scales were not so important.

More generally, in medieval music, before the baroque period, people were thinking about the *Gregorian Modes* of music. They were called things like the "Dorian" mode, which starts at d; the "Phrygian" mode, which starts at e; the "Lydian" mode, which starts at f, and so forth. But all the time, this is staying on the white keys of the piano — the diatonic scale.

Actually, the reason those medieval people gave their musical modes such funny-sounding Greek names was that they had become confused by the elaborate arithmetical contortions in Boethius' *de institutione musica*. In reality classical Greek music departed strongly from our simple diatonic scale.[5]

As we have seen, the Greek scale started with the intervals 1, 4/3, 3/2 and 2. And then both the lower half (from 1 to 4/3) and the upper half (from 3/2 to 2) were subdivided using two further notes, just as is done in our diatonic scale. However the tetrachords of the Greeks were arranged according to the *enharmonic* scale, where the two additional notes are squashed together in tiny quarter-notes, right at the bottom of the interval, and also the *chromatic* scale, where they are still rather squashed at the bottom, but not quite so much as in the enharmonic scale. According to the classical writers, the enharmonic scale was the true, hard, masculine, heroic form of Greek music. The chromatic scale was really just for softies — for women (although I suspect that many men also secretly preferred it!) Eventually, the Romans, who found other ways to parade their heroic masculinity to the world, dropped the enharmonic scale altogether, returning to the natural diatonic scale.

7 What About the Black Keys?

It seems to be a characteristic of the human mind that it seeks the new and the bizarre, becoming dissatisfied with what has naturally been given to us. I am reminded of a passage in the book *The Silverado Squatters*, by Robert Lewis Stevenson, where he describes the idea of "two bits" in the monetary system of the United States in the 19th century. Two bits was 25 cents. Therefore one bit must be 12 1/2 cents. Yet the smallest unit of money in the United States is one cent. This led to endlessly interesting conflicts in the bars of San Francisco as to whether one bit was 12 or else 13 cents.

In a similar way, we see that the scale of music consists of seven intervals, such that five of the intervals are approximately equal to twice the length of the two shorter ones. Therefore, in order to add a bit of spice into things, it was often found to be interesting to play a note about half-way between an adjacent pair of notes which are a whole note apart. In particular, b-flat was often added somewhere in between a and b. Given that, one can then make a new, but slightly "imperfect" diatonic scale by starting at f and then playing b-flat, rather than b. (I leave it as an exercise for the reader to determine what number should best be assigned to this new note.) In any case, most of the

[5] Perhaps it was a good method of keeping the ancient Greek households free of stray cats!

written music from the renaissance period seems to be either in the "natural" scale starting with c (with no sharps or flats), or else in the scale of f, with one flat (namely b-flat). Only seldom does one see music written with two flats. The keys with sharps are even more seldom.

On the other hand, to cover all eventualities, the organs from this period had black keys between all the pairs of white keys which are a whole note apart.[6] Given that we accept the principle of the democracy of all the keys of the organ, then we would wish to place the black keys in such a way that the whole and half notes, starting from a given key, will produce a pure diatonic scale. However the circumstances of simple arithmetic are such that this wish can never be fulfilled.

The question of how to tune the intervals of an organ, or in fact any keyboard instrument with fixed tuning, was a subject of much theoretical debate. The mean-tone system was widely followed. This attempted to optimise the possible thirds which might come up in various keys. Not only Euler; many other mathematicians gave much thought to this problem. One practical baroque system was described by Andreas Werckmeister (1645-1706). Another, somewhat later system is that of Johann Philipp Kirnberger (1721-1783). It is thought that he may have studied for a short period in Leipzig with the great Johann Sebastian Bach. In any case, he became the private music master of Princess Anna Amalie of Prussia, and at her request he wrote a book of music instruction, including a system of keyboard tuning. The Kirnberger system — which may well approximate the "well tempered" Klavier of Bach — kept the intervals for the white keys (except for a) precisely at their "natural harmonic" values, as we have described them. The black keys are somewhere in between. A good reference for these developments is the book *Zur musikalischen Temperatur*, by Herbert Kelletat.

Much of the charm of baroque music stems from the fact that the intervals, starting at different keys, are all different. Therefore music written in one key has an entirely different flavour from music written in another key. A theory of the emotional moods of the various keys of music was developed, and it's practical application was brought to a sublime culmination in the music of Bach.[7]

8 Equal Temperament

If, according to Kirnberger's system, music written in the key of c-major — that is, without sharps or flats — is nearly perfectly diatonic music, then it follows that music written in other keys departs more or less strongly from the perfect diatonic scale. In an age of democracy, this inequality of the scales is thought by

[6] It is interesting to note here that the ancient Romans built gigantic organs, in particular to increase the splendour of the games in the Colosseum. How strange it is that the Christian Church, whose early martyrs went to their deaths accompanied by blaring organ music, has adopted the organ as the central instrument of it's music.

[7] But it is interesting to remember that some of the greatest composers of Broadway musicals — when creating the famous melodies which are a part of our modern cultural tradition — composed their works entirely on the white keys, with one finger.

many people to be unfair, if not intolerable. Therefore a movement developed to secure the equality of all musical scales.

Given that there are 12 intervals if we count both the black and the white keys on the keyboard, then it was decided to make all these intervals equal. How can this be done? Clearly, each of the intervals must be exactly $\sqrt[12]{2}$. Therefore, a half note is $\sqrt[12]{2} = 1.0594630944\ldots$. A whole note is $\sqrt[6]{2} = 1.1224620483\ldots$, and so on. These are all irrational numbers, so the intervals can never be pure.

Still, people say that it's good enough. Democracy is worth this small price to pay. The intervals are so near to their theoretical values that nobody will ever be able to hear the difference anyway. But is this true? For example consider the musical interval which is a third. It is the ratio 5/4=1.25, exactly. On the other hand, if you count the number of black and white notes you need to reach a third along the keyboard of a piano, you will find that it is 4 equal intervals. That is $2^{4/12} = 1.2599210499\ldots$. The difference between this and 5/4 is just shy of 1/100.

Put another way, two instruments tuned to the equal temperament system, playing notes a third apart (which is an extremely common interval in practical music) are out of tune by a factor of approximately 1/100. Let us say that they are playing somewhere in the middle of the treble staff, where most music is found. That is, somewhere around 500 Hertz. But now, being 1/100th out of tune means that the dissonance produces a beating of about 5 beats per second. A beating of this kind gives a most unpleasant sound. Anybody would say that it sounds terribly out of tune. Why is it that we are generally not aware of this problem when listening to music?

There are a number of reasons. Most music involves many instruments, playing various intervals simultaneously. Thus the intonation problems of two of these instruments playing a third apart become submerged in the whole vibrating sound of the music. On the other hand, things become more critical when only two or three instruments are playing together. In order to hide these disharmonies, the usual practice is to play with a more or less exaggerated vibrato technique. That is, each individual instrument is played in such a way that the frequency produced wobbles back and forth through an interval of something less than 1/100th. The frequency of the vibrato is 4 or 5 wobbles per second. The modern ear experiences this as a kind of “warmth” in the music.

But why is it that a piano, tuned to the equal temperament system does not sound out of tune? There are a number of reasons which could be given. Perhaps it is just that we are accustomed to the sound of the piano in this tuning. But more to the point, a note played on the piano is not a continuous tone. It starts with the impact of the hammer on the strings, producing an initial complex wave. This progresses into a steadier vibration of the strings. But then the fact that the strings, particularly in the low notes, are very thick, means that the frequency changes as the amplitude of vibration of the strings decreases.

9 Just Intonation

There are other ways of producing a feeling of warmth in music. Not only in classical music, even more so in the more popular forms of music, a true and clear intonation is of the greatest importance. We are moved by a voice which is just right for the song. Nobody likes to hear a voice which wobbles all over the place!

Thus, in reality, a musician playing an instrument whose pitch can be varied — that is, not a fixed, keyboard instrument — always tries to play in such a way that everything fits together as perfectly as possible. This is something practical, not theoretical! When playing in this way, it is said that the instruments are playing according to the system of *just intonation*. In other words, the musicians adjust their playing at each moment of the music so that the ratios of the frequencies of their instruments are given by small numbers. They try to make the music sound "good", not "bad".

Perhaps the clearest example of just intonation is given by Indian music. It is played against the background of a plucked, twangy, droning sound. This sound consists of a precisely given low frequency $\mathcal{F}$, whose overtones $k\mathcal{F}$ are very strong, even for larger $k \in \mathbb{N}$. Then the melody instruments play with this given, basic sound, finding one and another overtone. Music theoreticians speak of *microtonal* ornaments. Such music is the very opposite of the equal temperament system based on the irrational number $\sqrt[12]{2}$.

With the advent of electronic music synthesisers, even keyboard instruments today can enjoy the virtues of unequal intonation. At the touch of a button, the system of tuning can be instantly altered. It is even possible to imagine a keyboard synthesiser which analysises the sounds of other instruments with which it is playing, just as if it were itself a musician. Then it would be programed to produce sounds which fit in perfectly — following the just intonation system — with the other instruments.