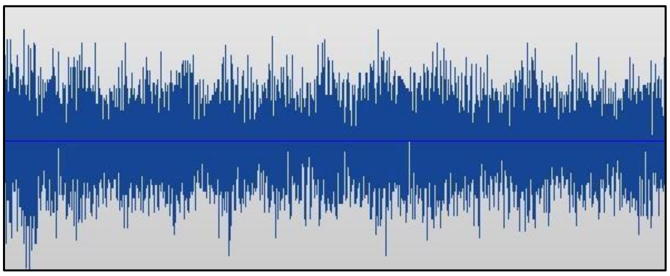
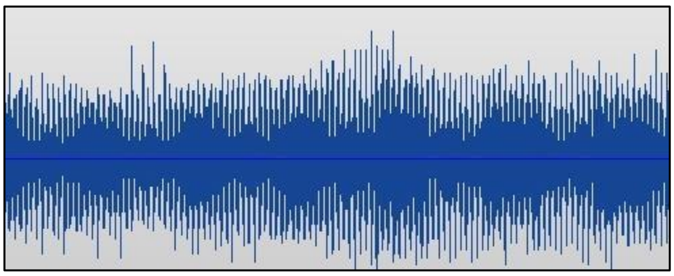
The Sound of Mathematics
Professor Sarah Hart

17 November 2020

This is the second lecture in my series of three on the links between Mathematics and Music. In the first lecture, we talked about the mathematics of musical composition. This time, we are going to look more closely at the mathematics of how sounds are made, what kinds of sounds fit harmoniously together, and the mathematical conundrum at the heart of harmony that threatens to destroy the whole edifice.

<u>What is a sound?</u>
Sounds arise from variations in air pressure. The greater the variation in pressure, the louder the sound. Human ears can detect remarkably small deviations in air pressure, as little as 0.00000002%. We can represent these fluctuations pictorially. Interestingly, it is almost impossible to tell with the naked eye the difference between even the most structured sound and random noise.


Beethoven's 5th Symphony


White noise

Usually, though not always (think of the cannon in Tchaikovsky's 1812 overture), the sounds we hear in musical compositions are from instruments producing notes with defined pitches. When we look at the sound made by a musical instrument, we see the same shape repeating. That regularity is what gives it a pitch, making it a note. The frequency with which the pattern repeats determines how low or high the corresponding pitch will be – the higher the frequency, the higher the pitch.

Frequencies are usually measured in Hertz (Hz), or repetitions per second. A higher frequency is perceived as a higher sound. We can hear for ourselves how changing frequency changes pitch – for example in a simple instrument like a siren whistle that contains two discs, one of which is moved at a pace determined by how hard you blow into it. The discs each have small holes, and when they line up air is pushed through. So, in blowing it, you create a repeating pattern of changes in pressure. The harder you blow, the faster the disc spins, and so the higher the frequency of the pattern. When we hear the sound, we experience this as an increase in pitch. Another phenomenon that allows us to observe this is the Doppler effect. When a sound is moving towards us (such as the siren on a moving fire-engine), the waves of sound are squashed closer together, thus increasing the frequency, which we perceive as the sound getting higher in pitch as it approaches. When the sound passes us and starts to move away, the reverse effect happens. The waves are stretched out, the frequency decreases and the sound becomes lower in pitch.

The study of sound goes back thousands of years. Legend has it that Pythagoras, passing a blacksmith's shop one day, heard the cacophony of several blacksmiths' hammers striking their anvils. Curiously, sometimes when two hammers struck at the same time, the combined sound would be pleasant, but at other times it would be horribly discordant. Intrigued by this phenomenon, Pythagoras entered the shop to observe. Eventually he realised that the times when the sounds blended together harmoniously were when the hammerheads happened to be related in exact ratios like 2:1 and 3:2, and he rushed home to try to replicate the effect – not with different weights of hammer but with different length strings, on a simple instrument called a monochord. Alas, the blacksmith shop story is apocryphal – doubling the size of a hammer head does not have the same effect on pitch as doubling the length of a string. But no matter: the mathematically important part of the story is the experiments with stringed instruments.

What the Pythagoreans found was that if you pluck a string, say on your lyre, and then halve the length and pluck it again, the two sounds produced are very similar, and sound pleasant together. Strings two-thirds or threequarters as long as the original also make pleasing harmonies with the original string. From now on we will talk about relative frequencies of these tones, rather than string lengths. Frequency is inversely proportional to string length, so halving the string length results in strings in the ratio 2:1 in length, and has the effect of doubling the frequency. The two other important intervals are obtained from ratios 3:2 (the shorter string is two-thirds the length and the frequency increases by a factor of 3/2) and 4:3 (the shorter string is three-quarters the length, raising the frequency by a factor of 4/3).

The halving of the length of a string produces a sound so similar that we almost think of it as the same note. Nowadays we call the interval between the two notes an octave. As an example, unison singing in a mixed group does not usually involve everyone singing precisely the same note. The women would tend to sing "the same note" an octave higher than the men.

The Pythagoreans prioritised octaves (2:1), perfect fifths (3:2) and perfect fourths (4:3). If we follow a perfect fifth with a perfect fourth, the net effect is to multiply these ratios. So, if our initial frequency is f, then a perfect fifth above this would be frequency $\frac{3}{2}f$, and then a fourth above that would be $\frac{4}{3}\times\frac{3}{2}\times f = 2f$, an octave higher.

Incidentally, the first two notes of the song *Somewhere over the rainbow* form an octave interval; the first two notes of *Twinkle Twinkle Little Star* form a perfect fifth (or for a more intellectual example, the first two notes of *Also Sprach Zarathustra* by Richard Strauss); for a perfect fourth you can sing *Here Comes the Bride* (also known as the Bridal Chorus from Wagner's Lohengrin) or *Amazing Grace*.

<u>The trouble with Intervals</u>

On a modern piano, we can obtain all twelve different pitch classes using intervals of a fifth. Starting at a low C, we go up a fifth to G, then to D, then A, E, B, and so on. After twelve iterations of this, we return to a C – it's seven octaves above where we started.

Legend has it during his experiments with a monochord (a simple instrument with one string), Pythagoras made observations about progressions of perfect fifths. He set up two monochords; the first starting with a long string that produced a low note. He then successively halved the length of the string seven times, to end up with a tone seven octaves higher. On the second monochord, he started with the same low note, but this time at each step he shortened the string length by two-thirds each time (increasing the frequency by a factor of 3/2), so as to raise the pitch by a perfect fifth. After twelve steps, the final note was sounded together with the final note on the first monochord. We might expect these to be the same. But each time however accurately the measurements were done, the two notes were just slightly, disturbingly, out of tune with each other.

We can see the reason for this: if we start with a note at frequency f, and raise the pitch by seven octaves, the new frequency is $2^7 f$, or $128f$. But if instead we raise the pitch by twelve perfect fifths, the new frequency is $\left(\frac{3}{2}\right)^{12} f$ or approximately $129.7f$.

But when we try this on a modern keyboard, we find no such problem. So, the “fifths” on a modern keyboard must be different! We are going to explore this, and since we will be dealing with many things calling themselves perfect fifths, we will, where necessary, distinguish the interval made by the ratio 3:2 by calling it a “pure” perfect fifth, or just a pure fifth. Similarly, 4:3 is a pure fourth; later we will encounter other pure intervals.

Comparing Scales

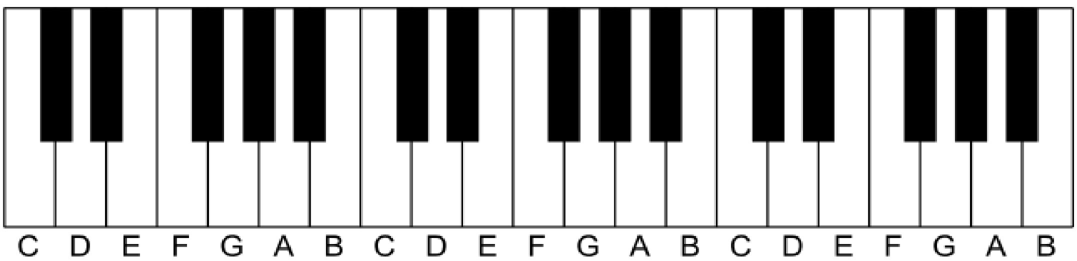
Alexander John Ellis (1814-1890), invented, for the purpose of technical analysis of different tunings, a minute subdivision of the octave known as the *cent*. This gives a way of comparing frequency ratios using a logarithmic measure in which one octave is divided into 1200 cents. One advantage of this is that it converts multiplication to addition (like log tables). The definition is that the ratio $\frac{f_2}{f_1}$ of two frequencies is equal to c cents, where $\frac{f_2}{f_1} = 2^{c/1200}$.

A frequency ratio r therefore corresponds to 1200 $\log_2 r$ cents. If we want to follow a ratio r_1 that is a_1 cents, with a ratio r_2 that is a_2 cents, then the ratio of the outcome is $r_1 r_2$, but the new cent value is

$$1200 \log_2 r_1 r_2 = 1200(\log_2 r_1 + \log_2 r_2) = 1200 \log_2 r_1 + 1200 \log_2 r_2 = a_1 + a_2.$$

A major advantage of this notation is that the same interval at different frequencies comes out at the same number of cents, and this tallies with how the human ear differentiates frequencies. It has been found that the smallest discernible difference in pitch is about 5¢, and, unless you are a professional musician, differences of a few cents more than this minimum, while perhaps detectable, are tolerable. The difference between 7 octaves and 12 perfect fifths is certainly enough to trouble the human ear, at about 23.5¢.

At this point it’s useful to recap a bit of basic music theory and notation. Look at the piano keyboard below.[1]



[1] Picture credit: https://commons.wikimedia.org/wiki/File:Klaviatur-3-en.svg

The white notes on a piano are labelled with the letters A to G. These represent the notes in the scale of C major. A frequency ratio of 3:2 corresponds to the interval from C to G, the fifth note, so it's called a perfect 5th. From G up to C is a fourth, which corresponds to 4:3. Pairs of notes a whole number of octaves apart have the same letter name. The black note between two white notes is the sharp (♯) of the note below, and the flat (♭) of the note above. So, the note between C and D can be referred to as C$^{\sharp}$ or D$^{\flat}$. The smallest gap is a semitone, such as that between E and F, or F and F$^{\sharp}$. Two semitones make a tone. An octave is made up of twelve semitones.

Every scale has the same arrangement of notes and intervals between them. C major, for instance, has C, D, E, F, G, A, B, C – if we include the final C this is eight notes, hence the word "octave". The intervals between E and F, and between B and C, are semitones; the remaining intervals are tones. This pattern (tone, tone, semitone, tone, tone, tone, semitone) is replicated whatever the starting, or *tonic* note (G major has G, A, B, C, D, E, F$^{\sharp}$, G).

The Pythagorean Tuning

Let's begin by looking at the Pythagorean method of constructing the scale. This makes use only of the pure intervals of octave, perfect fifth, and perfect fourth. Remember, the ratio of frequencies required to go up a (pure) perfect fifth is 3:2; to go down a perfect fifth is 2:3. Up a perfect fourth is 4:3; down is 3:4.

It goes as follows (you could start from any note, of course, but for simplicity we will use C). Starting from C, go up a perfect fifth (3:2 ratio) – this gives you G, with F being a perfect fourth (4:3) up from C. If you go down a perfect fourth (3:4) from G, you get D (net effect 9:8). A perfect fifth up from D is A. Going down a perfect fourth from A is E, going up a perfect fifth from E is B, and we have found all the notes of the scale of C major. You could keep going with this to add in the missing semitones – the black notes – if you wanted; for example, F$^{\sharp}$ is a perfect fifth up from B.

The final result of these Pythagorean calculations is the following, where the fractions given represent the relative frequencies of the notes compared to the initial tonic note (in this case C).

C	D	E	F	G	A	B	C
1	$\frac{9}{8}$	$\frac{81}{64}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{27}{16}$	$\frac{243}{128}$	2

There are some issues $\frac{3}{2}$ with this tuning. We have noticed that if you extend over many octaves using a combination of fourths and fifths, you don't actually get back to a whole number of octaves at any point. After twelve different pitch classes have been determined, the next one will be almost, but not quite, something you already have, so that is where you stop, and this is probably the reason why our musical tradition has twelve distinct notes in an octave. There are, in fact, other combinations of perfect fifths that come closer to reaching a whole number of octaves. For example, 53 perfect fifths approximate 31 octaves very closely, with an undetectable 3.6¢ discrepancy. The Pythagoreans did explore a 53-note scale. But the practical difficulties are too great. It would represent too overwhelming a choice to a composer and be extraordinarily difficult to do in practice for performers and instrument-makers. And no matter how far you go, a perfect concordance can never be achieved, because there is no positive integer power of $\frac{3}{2}$ that equals a positive integer power of 2.

$$\left(\frac{3}{2}\right)^m = 2^n,$$

We can see this because then we would have $3^m = 2^{m+n}$. But over a small range, and where you only have to be in tune with yourself singing, or playing the lyre, the twelve note “circle of fifths”, though in truth an infinite spiral, is a reasonable compromise.

Notice what happens in the Pythagorean scale when we compare intervals of consecutive notes. Below we see the relative frequency of each note compared to the one immediately before it.

$$\frac{9}{8}, \quad \frac{9}{8}, \quad \frac{256}{243}, \quad \frac{9}{8}, \quad \frac{9}{8}, \quad \frac{9}{8}, \quad \frac{256}{243}$$

So, we find that in the Pythagorean scale, an interval of a tone is given by a 9:8 ratio, and a semitone by 256:243. But shouldn’t two semitones make a tone? In the Pythagorean scale, two semitones correspond to a ratio of about 1.109:1, but a tone is 1.125:1. Later, the mediaeval writer Boethius suggested a division of the 9/8 tone into two ‘semitones’ of 18:17 and 17:16. It seemed to be impossible to divide even a semitone, let alone an octave, into exactly equal parts.

Licentious Modulations

Up until the middle ages, it was widely held that the ‘ideal’ music would be based on these pure fourths, fifths and octaves. Early church music featured plainsong chants. Gradually singers started to add fifths or fourths above the original line, creating a sound called organum. Variations crept in, and things really started to take off when musical notation was developed (by Guido de Arezzo among others, inventor of the stave and of the do-re-mi notation for the scale) that allowed compositions to be written down. This allowed singers to read and perform more complicated melodies, adding embellishments and harmonies. This pollution of the pure sound had its detractors: as early as 1132, a decree from the Cistercian order was quite firm that choristers should stop singing “in a womanish manner with tinkling […] as if imitating the wantonness of minstrels”. Happily, the wantonness continued, and music, both sacred and profane, developed apace with multi-part harmonies and other inventions added to the richness of the sound. Every innovation seems to have come with its critics. Centuries later, Claudio Monteverdi, composer of some of the most transcendently beautiful music of the baroque period, was slated for his “licentious modulations” and “mountainous collections of cacophonies”.

Along with the increasing complexity of choral music, in the secular world we start to see the development of new keyboard instruments. A letter from King John I of Aragon, from 1367, refers to a new instrument “similar to the organ, but sounded by means of strings”. By 1400 an instrument called the clavicembalum had been invented, or as we call it now, the harpsichord, which produces sound by plucking strings as the keys are pressed. In spite of Sir Thomas Beecham’s famous remark (many centuries later) that a harpsichord playing sounds like two skeletons copulating on a tin roof, this new instrument was a spectacular innovation that was wildly popular for centuries (and many of us still think it makes a beautiful sound). One feature of these new keyboard instruments is that because fixed tones are produced by the keys, it is difficult to finesse away any slight mismatches caused by the choice of tuning, as might be possible in a stringed instrument where finger position determines the note.

Another development which was bad news for Pythagorean tuning was the new harmonies that became fashionable in England in the fourteenth century – thirds and sixths. The cheerful ‘major third’ (as in the first two notes of “*When the Saints come Marching in*”) was obtained by having two notes in the ratio 5:4, while the melancholy ‘minor third’ (think the first phrase of Beethoven’s fifth symphony, or of Chopin’s funeral march in his Sonata in B-flat) corresponds to 6:5. Thirds are complemented by sixths – a major third followed by a minor sixth (or minor third and major sixth) creates an octave[2]. These edgy new thirds and sixths made their way over to the rest of Europe in the early fifteenth century, due in large part to the influence of English composer John Dunstable.

It became increasingly important to work out how to tune instruments so that they could play not only pure fourths and fifths, but also pure thirds and sixths. The problem is that in the Pythagorean tuning, a “major third” sounds horrible – if we look the major third from C to E in our scale above, it has the ratio 81:64 (or 408¢) rather than the desired 5:4 (or 386¢) of the pure major third. This is an unacceptable difference of 22¢. It’s the same story with minor thirds.

Solving the Tuning Problem

One way to address the problem of the new intervals is to replace the thirds and sixths in the Pythagorean tuning with pure thirds and sixths. A system that achieves both pure fifths and pure thirds in a given scale is called *justintonation*. One such system was suggested by the Spanish music theorist Bartolomeo Ramos de Pareja in 1482. Starting, for example, from C, we have the pure perfect fourth F and the perfect fifth to G as usual, with D being a perfect fourth below G. But now we set E to be the pure major third (5:4) above C, with A being a major third above F, and B a major third above G.

In other words, we replace $1, \frac{9}{8}, \frac{81}{64}, \frac{4}{3}, \frac{3}{2}, \frac{27}{16}, \frac{243}{128}, 2$ with $1, \frac{9}{8}, \frac{5}{4}, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, \frac{15}{8}, 2.$

This is all very well if we stick to C major. The thirds, fourths, fifths and sixths now sound lovely. But we have now broken some of the fifths! If we try to play the interval from D to A, which should be a pure perfect fifth and have ratio $\frac{3}{2}$, we find that in fact we have a ratio of , $\frac{5}{3} \div \frac{9}{8}$ which is $\frac{40}{27}$, over 20¢ away from a pure fifth. The big stumbling block, then, is that just-intonation cannot be achieved for all scales at the same time, meaning that on a keyboard we cannot transpose melodies into different keys without changing how they sound, and furthermore, the sanctity of the pure perfect fifth is destroyed.

The legitimacy of just-intonation was a matter of heated debate. Among the opponents was Franchinus Gaffurius (1451-1522), a well-known composer, teacher and friend of Leonardo da Vinci, who found the loss of pure fifths unconscionable. As often happens, practitioners were just getting on with trying to find pragmatic solutions to the problem. Gaffurius reports on one of these solutions in his influential book *Practica Musicae* (1486), which was published during his time as maestro di cappella at Milan cathedral. Gaffurius was a working musician, so although he favoured the Pythagorean tuning in principle, he was very well aware that in the real world sometimes we must deviate from our perfectionist ideals. He reports in *Practica Musicae* that some organists were adjusting the lengths of their organ pipes in order to very slightly modify, or temper, the ratios of their fifths. This was called *temperament*, or *participata.*

The process was this: we know that an interval of twelve pure perfect fifths is close to seven octaves. The first four of these intervals, starting from C, go C, G, D, A, E, with that high E being a frequency $\left(\frac{3}{2}\right)^4$ times the initial frequency.

[2] There is a major sixth at the start of *My Bonnie lies over the ocean;* there are plenty of minor sixths in the main theme of Scott Joplin’s *The Entertainer*, though not absolutely the first two notes. The first two notes of Chopin’s Waltz op 64 no. 2 are a minor sixth, as is the beginning of the guitar intro to *In My Life* by the Beatles, but it is quite rare to begin a melody with this interval.

If we bring this down two octaves, we get the $\frac{81}{64} \approx 1.266$ of the Pythagorean tuning, compared to the pure major third given by $\frac{5}{4}$ = 1.25.
The gap between these two tones is called *syntonic comma*, or simply *comma*, and it is 22¢. The solution the organists found to this was to make each of the fifths that we combined just a tiny bit smaller, each by a quarter of this comma. Making this high E a pure major third causes it to have exactly five times the frequency of the initial C. So, we need to set G, for instance to be $\sqrt[4]{5}$ times that frequency. The organists were not saying this explicitly but were simply reducing their fifths by "a bit". The resulting G is around 697¢, compared to the pure fifth of about 702¢. The difference, then, at 5¢, is acceptable to almost all listeners. The difference in the D and A are a bit worse, but still tolerable.

This tuning is an example of what is called *mean-tone temperament,* so named because it involves finding the geometric mean, or average, of intervals. Mean-tone temperament still has some problems though. As an example, suppose we tune our organ with respect to D major. The tonic triad (1st, major third, fifth) then will be D, $F^{\#}$, A. But now imagine trying to play a piece in the key of $D^{\flat}$ on that same keyboard. The scale goes $D^{\flat}$, $E^{\flat}$, F, $G^{\flat}$ $A^{\flat}$. Both $F^{\#}$ and $G^{\flat}$ are trying to occupy the semitone slot between F and G. On our piano they are the same note. You can tune this note to be a pure major third ($F^{\#}$) in relation to D, or a pure perfect fourth in relation to $D^{\flat}$, but not at the same time.

This problem was dealt with by some instrument makers by providing extra keys – for example, a different key for $F^{\#}$ and $G^{\flat}$. There are extant copies of contracts for cathedral organs from as early as 1480 that specify such expedients. Meanwhile harpsichords were often supplied with a choice of several keyboards. The composer Nicola Vicentino (1511-1575/6) designed an instrument called an archicembalo which was a modified harpsichord with two keyboards, each of which had three rows of keys, producing up to 36 notes per octave! Mersenne, whom we will mention again later, included many diagrams of keyboards with 17, 19 and other numbers of keys to the octave, in his 1636 book *Harmonie Universelle*. Such instruments were of course extremely challenging, both to make, and to learn to play. From the 15th to the 17th century dozens, even hundreds, of variations on mean-tone temperament were suggested, and gradually people started to talk seriously about the possibility of making all twelve of the tones in the octave equally spaced.

There was a conceptual problem with doing this. Early theorists such as Boethius had balked at the suggestion, because, taking our tone as 9:8, if we want to split this exactly in half, we need to find the square root of 9/8. But there is no fraction whose square is 9/8 – the square root of 9/8 is $3\sqrt{2}/4$, an irrational number. Boethius concluded that we cannot split the tone exactly in half – one of the intervals must be slightly greater than the other (in the Pythagorean system the full tone is 204¢, the 'semitone' is 90¢, so one interval is 90¢ and the other is 114¢). Boethius suggested that the steps be made in ratios of 18:17 and 17:16, corresponding to 99¢ and 105¢. These is a much more even split.

However, although irrational numbers were mistrusted, a sheen of respectability was given to those that are obtainable as the geometric mean of two rational numbers by the fact that Euclid gives a geometric solution to doing this. It allows one to find the geometric mean between two numbers. One could repeat this to split an interval into four, eight, even sixteen parts, so the mean-tone temperament splitting the major third equally into four is not problematic. But splitting it into twelve is harder. However, techniques were eventually developed to do this, and eventually the psychological challenge of accepting irrational numbers passed away too.

Equal Temperament

The Dutch scientist Simon Stevin (1548-1620) argued strongly for *equal temperament* – the subdivision of the octave into twelve exactly equal intervals. He had advocated throughout his career for all numbers (negative, irrational and so on) to be treated the same when doing things like solving equations, so the use of irrational numbers to divide the octave felt perfectly natural to him. In fact, he felt that equal temperament was the only natural tuning, because all the others contain so many flaws, annoying commas, and ultimately irreconcilable inconsistencies. The idea that the 3:2 ratio is the pure “natural” fifth, he says, is nonsense – it is clearly just an approximation, and this error was all the fault of the Ancient Greeks, who stubbornly persisted in this idea even though they had themselves shown it to be fatally flawed by observing that a succession of twelve such intervals ends with a note out of tune with the first. Stevin explained that since Dutch is the only true language of science, it is understandable that Greek speakers were not able to see the true state of things.

To produce equal temperament, each semitone must of course be exactly be $\sqrt[12]{2}$ apart. This corresponds to exactly 100¢ per semitone. The new tuning was different enough from some of the pure intervals to be detectable, particularly for major thirds (400¢ for equal temperament, 386¢ for the pure major third; whereas fifths were OK – 700¢ compared to 702¢), so it took a while for people to get used to it. But the advantages outweighed the disadvantages, as musical composition became increasingly complex and adventurous, and all twelve tones were routinely needed.

As the theoretical concept of equal temperament became more popular, the issue of how to achieve it in practice grew ever more pressing. How do you actually construct, for example, the frets on a lute in such a way as to produce all those $\sqrt[12]{2}$ intervals? Vincenzo Galilei (1520-1591), a music teacher and lutenist, came up with one pragmatic solution. He thought about the old suggestion of subdividing the tone into intervals of 18:17 and 17:16, and suggested that 18:17 be used as a ratio for the positioning of lute frets. The correct $\sqrt[12]{2}$ interval is 100¢, whereas 18:17 is 99¢, so the difference per semitone is undetectable. After seven of these semitones we have 693, which is flatter than the truly equal temperament 700 would give, but a little tweaking of the tension, finger pressure and so on can finesse that away. The method was easy to implement, and widely adopted.

If we look at the fretboard of a six-course lute (a course is a pair of strings played together, except for the first course, which is a single string called the chanterelle), we can see this in action. The strings go from the bridge, along the neck to the nut at the top. To get the first semitone, the fret must divide the total string length in a ratio of 18:17. Thus, the distance from the nut to the first fret is 1/18th of the total length of the string. The distance from the first fret to the second fret is 1/18th of the remaining 17/18th distance from the first fret to the bridge, so we get $\frac{17}{324}$ of the total distance, slightly less than $\frac{18}{324}=\frac{1}{18}$ the between the nut and the first fret.

This gap decreases further between subsequent frets. Nowadays we work with better approximations to the $\sqrt[12]{2}$ interval, but this still results in the same phenomenon, of frets getting closer together, in the frets of the guitar, mandolin and other fretted stringed instruments (the violin family do not have frets).

I should mention here one other tuning that found some favour – the *well-temperament* proposed in 1681 by Andreas Werckmeister for keyboard instruments. Some intervals were purer than others, but no key was unacceptably out of tune. The effect was that changing the key in which a piece was played would have subtle effects on its sound. Some keys were more melancholy, others were brighter and more cheerful. Any pianist hearing the phrase will immediately ask whether Bach (1685-1750) preferred welltemperament or equal temperament. His *Well-Tempered Clavier*, published nearly 300 years ago in 1722, contains a prelude and fugue in each of the 12 major and 12 minor keys. It was written “for the needs and use of musical youth, as well as those already experienced in this study for the passing of time”. But the title could simply refer to something being tempered in a good way, so that all keys are playable, which could include equal temperament, rather than specifically referring to Werckmeister’s tuning. There are strong opinions in both directions. The book could have been intended to show the different character of the keys, or it could have been intended to show that pieces could have the same flavour in any key. Bach transposed material from one key to another on many occasions, seemingly without worrying that this would change the essential quality of the harmony, but there are plenty who argue the opposite case. We will probably never know.

It is a curious circumstance that the person given credit for the first system of equal temperament tuning was someone from an entirely different musical tradition, one with none of the issues of thirds and fifths that was compelling the development of compromises in temperament. The person in question was Chu Tsai-Yü. Born in 1536 in Ho-nei, China, he was a Chinese prince, descended from the fourth Ming Emperor. His interest in division of the octave into twelve equal parts seems to have arisen from an ancient Taoist ceremony marking the passing of the twelve months of the year. Each month has its own defined musical pitch, determined by a bamboo ‘pitch pipe’ of a given length, and the ideal is to have these twelve pitches exactly evenly distributed. By Chu’s era, this ancient ritual was not itself very important, but the mathematical puzzle had taken on a life of its own. In Chu’s 1584 treatise *A New Account of the Science of the Pitch Pipes*, he proposed a method of equal division which used the ratio $\frac{749}{500}$ for a perfect fifth – at 699.65¢ this is an extremely good approximation to the 700¢ of equal temperament.

It took a very long time for equal temperament to be adopted. People still hankered for the lost purity of major thirds, and it became a bit of an intellectual hobby to dream up new variations on the idea. Christiaan Huygens, for example, used logarithms to calculate an equal division of the octave into thirty-one part-tones, so as to give a wider range of possibilities. He also designed a keyboard that could be fitted onto an ordinary harpsichord in such a way that the keyboard could be positioned as desired to make the required tones available. With this division, each part is about 39¢; an interval of ten parts is 387¢, which is just 1¢ away from a pure major third (386¢). An interval of eighteen parts is 697¢, compared to the 702¢ of a pure fifth. But 31 is just too many notes to really be practical.

The Mathematics of Frequency

It wasn’t until the 17th century that rules about the frequency produced by vibrating strings were clearly formulated. Vincenzo Galilei had discovered that the pitch of a vibrating string is proportional to the square root of the tension at which it is held – this is a contender for the earliest known discovery of a square root law. Vincenzo’s son was of course Galileo Galilei (1564-1642). He contributed to the theory as well. In his last book *Dialogues Concerning Two New Sciences* (1638), there is a long discussion of vibrations and sounds. Galileo says that the frequency of a vibrating string is inversely proportional to its length, directly proportional to the square root of the tension, and inversely proportional to the square root of the string’s weight. Anyone who has played a guitar or other stringed instrument has experienced the fact that you can, for example, raise the pitch (increase the frequency) either by shortening the string (by pressing your finger against a fret), by tightening the screw on the string to increase the tension, or by using a lower density string (compare the strings for the lowest and highest notes).

At almost the same moment, Marin Mersenne (1588-1648) was independently stating essentially the same formulae for vibrating strings, which, moreover, he had verified experimentally. They are now known as Mersenne's laws. Though perhaps better known in mathematical circles for his work on what are now called Mersenne primes (prime numbers of the form $2^p - 1$, where p is prime), he published two important books on music: *Harmonicorum Libri* (1635) and *Harmonie Universelle* (1636). Mersenne was a French monk whose extensive network of correspondents – including Descartes, Pascal and Galileo, was hugely important for the development of scientific and mathematical ideas in Europe at that time.

Nowadays we encapsulate these laws in the following equation.

$$f = \frac{1}{2l}\sqrt{\frac{T}{\mu}}$$

Here, f is the frequency in Hertz (Hz), l is the length of the string (in metres), T is the tension (in Newtons), and μ is the linear density (mass per unit length, in kg per metre).

Mersenne actually measured frequencies, the first person to do so. He did this by making use of the rule that the frequency halves when the length doubles. He set up a monochord and adjusted the length of the string until it produced a very high but recognisable pitch. Then he doubled the length repeatedly (equivalent to lowering the pitch by one octave each time) until he could physically see the vibrations and be able to count them. To find the frequency of the original note, it was then just a question of doubling the measured frequency the required number of times. From this the frequency of all other notes of the scale could be found, based on the comparative lengths of the strings used to produce them. There is scope for huge inaccuracy in this, because any error in the initial measurement doubles with every octave.

Harmonies and Harmonics

We have seen how, over many centuries, the problem of division of the musical scale was worked on until the solution of equal temperament became standard. But we have not addressed one key (pun absolutely intended) question. Why is it that these intervals are so favoured?

Galileo Galilei wrote about this in his *Discourses and Mathematical Demonstrations Relating to Two New Sciences*. He made an analogy with pendulums. Galileo's first biographer, Viviani, tells a story about how Galileo came to study pendulums. In 1583, at the tender age of nineteen, Galileo was attending Mass one day at the cathedral in Pisa, when he started watching a lantern hanging from a long chain in the dome swinging backwards and forwards in the breeze. He decided to see how long it took for each oscillation to complete, timing them using his pulse, and was surprised to discover that the period was always the same irrespective of how high up the lamp started – in other words the amplitude of the motion. This led him, later, to suggest using pendulums in time-keeping devices, because you could start them swinging and each swing would take the same amount of time, even as the pendulum gradually lost energy and the swings decreased in amplitude. This pendulum clock origin story is somewhat problematic, because records show that the hanging lantern in question was not actually installed in Pisa cathedral until 1587, some four years after Galileo was supposed to have watched it swinging. Galileo doesn't write about pendulums having a constant period of oscillation until 1598, and, although he started using them as time-keeping devices from 1602, he didn't come up with the idea of making a pendulum clock until 1637 – and it wouldn't be until 1656 that a working pendulum clock would be built by Christiaan Huygens.

Anyway, Galileo had perhaps helped his father with some of his musical experiments with weighted strings, so this may have been in his mind. What Galileo does in the *Discourses* is to ask us to imagine two pendulums, one swinging twice as fast as the other. This would make a pleasing visual image, he says, because it would take exactly two swings of the first for one swing of the other, and then they would be in the same position again – and that's why an octave (2:1 ratio) sounds nice. Similarly, if one pendulum completes three swings for every two of the second, again they are perfectly aligned after those swings. This corresponds to a perfect fifth (3:2). But something like 9:8 (a tone) would take a large number of swings for alignment, so would look quite chaotic. This wouldn't sound nice. And even worse, if you take something like a ratio of $\sqrt{2}:1$, then the pendulums would *never* align, so this combination of notes would sound horrible. This corresponds to an interval called a *tritone* – three full tones, or half an octave – for example C to F♯. This is not something you see being used in musical composition very much, at least not until the 20th century. Although this sounds plausible to begin with, it doesn't really answer the question. After all, an 8:5 ratio sounds nice (a minor 6th), but a 7:5 ratio doesn't.

Many tried to shed light on this question. Even Samuel Pepys wanted to get in on the act. On 20th March 1668 he was "at my chamber all the evening pricking down some things, and trying some conclusions upon my viall, in order to the inventing a better theory of musique than hath yet been abroad; and I think verily I shall do it". On 2nd April 1668, he went "with Lord Brouncker and several [other members of the Royal Society] to the King's Head Taverne by Chancery Lane, and there did drink and eat and talk, and, above the rest, I did hear of Mr. Hooke and my Lord an account of the reason of concords and discords in musique, which they say is from the equality of vibrations; but I am not satisfied in it, but will at my leisure think of it more, and see how far that do go to explain it". (William Brouncker was first President of the Royal Society; Robert Hooke was the then Gresham Professor of Geometry.) The very next day he went "by coach to Duck Lane, to look out for Marsanne, in French, a man that has wrote well of musique, but it is not to be had, but I have given order for its being sent for over, and I did here buy Des Cartes his little treatise of musique, and so home". By January 1669 he is "in the right way of unfolding the mystery of this matter, better than ever yet". Sadly, that is the last we hear from him on the subject. It seems the problem may have been harder than he thought!

So, what's really going on? The key observation was probably made first by Joseph Saveur (1653-1716). Although he was born with severe hearing problems, he was fascinated by the science of sound – he coined the term *acoustics*, from the Greek *akoustikos*, meaning audible. (He also found a very accurate way to measure frequencies, using what are known as *beats*, that is still used today[3].) Studying vibrating strings, he placed small pieces of paper at various places on the strings. He noticed that there were places where they moved much less, in particular, in the places where the string was divided exactly in half, into thirds, into quarters and to a lesser extent greater fractions. We now know this is because these are the stationary points when whole numbers of waves fit into the length of the string. Thus, the frequencies that result are precisely whole number multiples of the original fundamental frequency of the full-length string. In addition to the fundamental frequency, we also hear, increasingly faintly, the *overtones* or *harmonics*. Below are shown the (approximate) notes corresponding to the first ten harmonics of a low C note.

[3] When two similar frequencies play together, their waves interfere with each other, resulting in peaks and troughs of sound, known as beats. Measuring the beat rate allows you to work out the relative frequencies. If the two frequencies are a (the higher) and b, then a little trigonometry tells us that the beats occur $a - b$ times per second.

The first harmonic is the note itself, with frequency f (in this diagram we have used C two octaves below middle C). The second corresponds to a string half the length, so we get double the frequency: the same note C but one octave higher, $2f$). The third is a string one-third the length, giving a sound at $3f$. Remembering that multiplying a frequency by 3/2 corresponds to an interval of a pure perfect fifth, we can rewrite this as $\frac{3}{2} \times (2f)$, so this is a perfect fifth above the second harmony: a G.

The fourth harmony is frequency $4f$ – corresponding to middle C in the example. The fifth is $5f$, which is $\frac{5}{4} \times (4f)$, a pure major third above 4f. And so on.

Subsequent harmonies 4 are extremely faint and most people can't even hear them, so they are less and less "expected" parts of the sound. It's not until the 9th harmony that you get what corresponds to a tone.

Why Do Different Musical Instruments Sound Different?

When a musical instrument is played, the sound we hear is based on the initial "transient" part of the sound, and then the set of resonant frequencies that make up the main part of the sound. In a stringed instrument, such as a violin, we set one (or more) strings vibrating by either plucking, or using a bow to move, the string. The transient sound is critical to our experience of it, but very tricky to analyse mathematically. However, the main part of the sound is somewhat more amenable to analysis. For instance, in a string fixed at both ends we have something called the wave equation. If we have a string fixed at both ends, and we disturb it in some way at time t = 0, then the displacement y of the string at point x after time t is a function of both x and t. The equation it satisfies therefore depends on both x and t. The exact relationship is given by the wave equation, as shown – but feel free to ignore it if you don't understand the symbols!

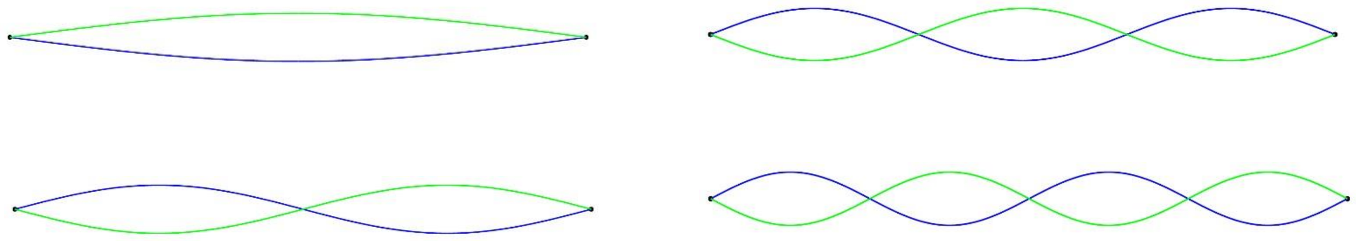
$$\frac{\partial^2 y}{\partial t^2} = \frac{T}{\mu} \times \frac{\partial^2 y}{\partial x^2}$$

Here, T is the tension in the string, and μ is the linear density. Essentially it is saying that the way in which the displacement y varies as a function of time is very similar to the way in which it varies as a function of the position along the length of the string. The French mathematician Jean-le-Rond d'Alembert (1717-1783) found a method to solve this equation. The solution is in the form of a function or wave A travelling forwards, added to a wave B travelling backwards. When a wave travelling to the left hits the end of the string, it "bounces back" and returns inverted as a wave travelling to the right, then bounces back again at the other end of the string. This means that $A(t + 2l) = A(t)$, where l is the length of the string, so the function is periodic (repeating) with period $2l$. Also $B(x) = -A(L - x)$, so B is also periodic with period $2l$. Now we can use a brilliant bit of mathematics due to Joseph Fourier (1768-1830). In the course of research into the conduction of heat, he managed to show, amazingly, that any periodic function can be broken up as a combination of sine waves. I'll be talking a lot more about the sine curve in one of my lectures next year, so I won't elaborate on this now, but it's possible to work out exactly which sine curves are used, and this in turn allows us to replicate sounds electronically.

It turns out that each solution to the wave equation for a string fixed at both ends is a linear combination of sine curves of period $2l$. The smallest of these would correspond to the function $\sin\left(\frac{\pi z}{l}\right)$ writing z for an arbitrary variable.

But we also have $\sin\left(\frac{2\pi z}{l}\right)$, $\sin\left(\frac{3\pi z}{l}\right)$, $\sin\left(\frac{4\pi z}{l}\right)$ and so on. The first of these gives a pure sound with frequency, say, f. The second would then have frequency $2f$, and so on. These are precisely the harmonics that Saveur found. The way the string is set in motion (plucking, bowing etc), along with the shape and size of the instrument, and what it is made of, all play a role in determining the relative amplitudes of the harmonics that make up the sound. It is this which makes different instruments sound different.

The sine wave corresponding to the fundamental frequency (the first harmony) and the subsequent three harmonies, are shown below – we show the highest and lowest positions of the string in blue and green.



These calculations are based on a stringed instrument. For a woodwind instrument, we have a situation where the air is being moved inside a tube. The *acoustic pressure* P is the difference between the pressure at a given position and time in the tube, and the ambient air pressure. Initially we have equal air pressure everywhere, given by whatever the ambient air pressure is. So $P = 0$ at time 0. Then air enters the tube; if we have both ends of the tube open, as in a flute, then the boundary condition is that the ends of the tube remain at acoustic pressure $P = 0$, and the acoustic pressure inside fluctuates. So, this is like the wave equation again; we get all the possible multiples of the initial sine wave as solutions. If the tube is closed at one end, such as in a clarinet, then the wave of pressure reflects off that closed end, as it has nowhere to go, so we get maximum pressure at the closed end. This means we see odd multiples of the fundamental frequency. This is one of the reasons why flutes and clarinets don't sound the same.

Every instrument produces different combinations of the fundamental frequency and its multiples. Now, if we think of the second harmonic, an octave higher, with frequency $2f$. Its harmonics will be all the multiples of $2f$. So, all its harmonics are also harmonics of a note at f. That's why they go so nicely together. Similarly, for any number of octaves apart, they will share almost all of their harmonics. A note at $3f/2$ is not a harmonic of a note at f, but it lies an octave below f's harmonic at $3f$. So, every harmonic of $3f$ is also a harmonic of $3f/2$. That is, f and $3f/2$ share half of their harmonics! And all the harmonics of both f and of $3f/2$ are harmonics of $f/2$. Although they don't sound similar enough for us to give them the same name, as we do with f and $2f$, they are certainly similar enough for us to enjoy hearing them together. This is a convincing argument for why that 3:2 ratio is one that we enjoy hearing. Similarly, if we think about 5:4, our major third, although a note with frequency 5f/4 is not a harmonic of a note at f, both f and $5f/4$, and all their harmonics, are themselves harmonics of $f/4$, the note two octaves below f. So, again we would expect the 5:4 ratio to result in a pleasing combination. The further through the list of harmonics we go, though, the quieter the harmonics are, and the less able our ears are to detect them. Certainly, it's not the case that all pitch classes derivable from harmonics sound "nice" with the original note.

By the way, as these harmonics most naturally arise in stringed instruments, in cultures where the instruments are different, our musical scale would not necessarily be the most "natural". David Benson, in his book on mathematics and music, gives the example of Indonesian gamelan music, where the instruments are all percussive. Such instruments do not produce exact integer multiples of a fundamental frequency, and so the Western musical scale does not work with gamelan music.

And Finally, A Musical Illusion

Now that we understand how sounds are made up, in theory we can produce any sound by simply adding together the desired component pure tones. This simplifies things slightly if we genuinely want to synthesise a real musical instrument, because if we don't want a very artificial sound, we have to also mimic the transient sound at the start of the sound. It also allows us to play around and create some strange aural illusions, in the manner of optical illusions. We will hear one of these in the lecture; it was invented by the American cognitive scientist Roger Shepard (b. 1929) and is a repeating scale that appears to be perpetually rising. This feat is by creating each tone out of ten harmonics spaced one octave apart, in such a way that the middle ones are loudest, getting quieter at the top and bottom ends, in the same configuration each time. Shepard scales have been used in several pop songs (most recently the Franz Ferdinand track *Always Ascending* (2018), whose accompanying video was shot in a way that the camera also appeared to be always rising). A Shepard glissando, in other words a continuous sound rather than a discrete set of tones, was also used in Christopher Nolan's 2008 film *The Dark Knight* for the sound of Batman's motorbike (the Batpod) to give the impression that it was constantly accelerating, without abrupt changes in the sound. There are many other such illusions. For example, Diana Deutsch discovered in 1975 that if two Shepard tones are separated by exactly half an octave (the *tritone* mentioned earlier), then you might expect that people would have trouble working out which note is higher. But, actually, the vast majority of people are quite sure in each case that one note is definitely higher. It's just that in 50% of the cases they say the first note is higher, and in 50% of cases the second note, when in fact neither is really true.

Further Reading

Benson, D., 2006. *Music: A Mathematical Offering.* Cambridge: Cambridge University Press.

Fauvel, J., Flood, R. & Wilson, R., 2003. *Music and Mathematics.* Oxford: Oxford University Press.

Harkleroad, L., 2006. *The Math Behind the Music.* Cambridge: Cambridge University Press.

Isacoff, S., 2003. *Temperament: How Music Became a Battleground for the Great Minds of Western Civilization*, Vintage. Maor, E., 2018. *Music by the Numbers: From Pythagoras to Schoenberg.* Princeton University Press.

Further listening

- To play with different tunings such as Pythagorean, Just Intonation, Mean Temperament and so on, there is the Microtonal Synthesiser from Offtonic at https://offtonic.com/synth/index.html.
- The Rational Keyboard is fun. The author describes it as "a web app demonstrating a consonance model in just intonation. It's kind of like a piano with infinitely many keys (one for each rational number) that move around and resize based on what sounds good." http://fritzo.org/keys/#style=piano
- For a Shepard glissando: Baysan, U. and Macpherson, F. (July 2017), "Shepard Tone Illusion" in F.
  Macpherson (ed.), The Illusions Index.
  Retrieved from https://www.illusionsindex.org/i/shepard-scales.
- For the Shepard Scale from the lecture: https://www.youtube.com/watch?v=PCs1lckF5vI, created by Howard Freeland.

- There is an example of a Diana Deutsch illusion – the octave illusion – at https://www.youtube.com/watch?reload=9&v=IMMsK9rjBWo.