The search for a mathematically perfect scale

Aurora Samuels

1.1 Introduction

It is often claimed that 'music is maths', however to mathematicians, this relationship is often seen as more of a beautiful tragedy rather than perfect harmony. The foundation of Western music is built upon a fundamental impossibility: the ability to balance the physics of a vibrating string with the laws of prime factorisation. While musicians perceive the circle of fifths as a closed loop, mathematically it is an infinite spiral that never returns to its origin.

In this essay, I will explore the musical scale as an increasingly sophisticated rational approximation of the irrational. If you have ever played a musical instrument or taken any musical theory, you will recognise the number 12: the number of notes within an octave (or the number of notes within the **chromatic** scale.) However, today we'll explore the mathematical advantages of different musical scales, such as a 53-note system.

Figure 1: a chromatic scale starting on D. While the stave has 13 notes on it, it represents the 12 semitones within a scale: the first and last note both represent a D.

1.2 Preface

Throughout this essay, I have put some words in **bold**, these are musical terms, which I have defined at the end in case you are not familiar with music theory.

Throughout, I will often refer to a major fifth and an octave.

A major fifth is a musical interval – a step from one note to another. Most well known is the opening 2 notes to the Star Wars theme tune.

An octave is also a musical interval, most commonly between the notes in the word "some-where" in the song "somewhere over the rainbow" from the Wizard of Oz.

2. The Harmonic series

Although to mathematicians, the Harmonic series may most famously connote $\sum \frac{1}{n}$ today we will be exploring another type of harmonic series: the sequence of musical notes whose frequency is an integer value of a fundamental frequency.

First, we will explore the mathematical properties of a string fixed at both ends, essential to the mechanisms of many instruments, like a guitar string, or a pianos steel cord, using the 1-dimensional wave equation, derived from the tension, length, and mass of a string:

$$\frac{\partial^2 y}{\partial t^2} = v^2 \frac{\partial^2 y}{\partial x^2}$$

If we let the solution y(x,t) be a product of a spatial function X(x) and temporal function T(t):

$y(x,t) = X(x)T(t)$

We can substitute this into the wave equation:

$$\frac{\partial^2}{\partial t^2}(X(x)T(t)) = v^2 \frac{\partial^2}{\partial x^2}(X(x)T(t))$$

$$X(x)\frac{\partial^2}{\partial t^2}(T(t)) = v^2 T(t)\frac{\partial^2}{\partial x^2}(X(x))$$

$$X(x)\frac{\partial^2 T}{\partial t^2} = v^2 T(t)\frac{\partial^2 X}{\partial x^2}$$

From this rearrangement, we can see that partial derivation with respect to t only affects T, and partial derivation with respect to x only affects X.

To get all the x's on one side, and all the t's on the other, divide both sides by $v^2 X(x)T(t)$

$$\frac{X(x)\frac{\partial^2 T}{\partial t^2}}{v^2 X(x)T(t)} = \frac{v^2 T(t)\frac{\partial^2 X}{\partial x^2}}{v^2 X(x)T(t)}$$

$$\frac{1}{v^2 T}\frac{\partial^2 T}{\partial t^2} = \frac{1}{X}\frac{\partial^2 X}{\partial x^2}$$

We can see that the LHS only changes if t changes, and the RHS only changes if x changes.

Therefore, for the LHS=RHS for all values of x and t, they must be equal to a constant.

This is the separation constant, $-k^2$

(A negative constant is used so the solutions oscillate, a positive constant would give exponential growth/ decay which wouldn't happen in this model of a stable musical string)

If we focus solely on the spatial ordinary differential equation, we can see that:

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} = -k^2$$

$$\frac{\partial^2 X}{\partial x^2} = -k^2 X$$

$$\frac{\partial^2 X}{\partial x^2} + k^2 X = 0$$

Let $X(x) = e^{rx} \rightarrow \frac{\delta^2 X}{\partial x^2} = r^2 e^{rx}$

$$r^2 e^{rx} + k^2 e^{rx} = 0$$

$$e^{rx}(r^2 + k^2) = 0$$

e^{rx} can never be zero, so we can divide through by this

$$r^2 + k^2 = 0$$

Giving us $r = \pm ik$

$$X(x) = e^{rx} = C_1 e^{ikx} + C_2 e^{-ikx}$$

Where C_1 *and* C_2 are constants

We know from Euler's formula that:

$$e^{i\theta} = \cos\theta + \mathrm{i}\sin\theta$$

$$X(x) = C_1(\cos(kx) + \mathrm{i}\sin(kx)) + C_2(\cos(-kx) + \mathrm{i}\sin(-kx))$$

$$X(x) = C_1(\cos(kx) + \mathrm{i}\sin(kx)) + C_2(\cos(kx) - \mathrm{i}\sin(kx))$$

$$X(x) = (C_1 + C_2)\cos(kx) + i(C_1 - C_2)\sin(kx)$$

Let $A = i(C_1 - C_2)$

Let $B = (C_1 + C_2)$

$$X(x) = A\sin(kx) + B\cos(kx)$$

The string's tied at both ends, so apply physical boundary conditions:

At $x = 0, X(0) = 0$

$$X(0) = 0 = A\sin(0) + B\cos(0)$$

Therefore $B = 0$

At $x = L, X(L) = 0$

$$X(L) = 0 = A\sin(kL)$$

$$\sin(kL) = 0$$

This occurs when kL is an integer multiple of π

$$kL = n\pi$$

$$k = \frac{n\pi}{L}\ \{n \in \mathbb{Z}^+\}$$

Since we can calculate angular frequency, ω, using $\omega = kv$

$$\omega_n = \frac{n\pi v}{L}$$

And it can be calculated by $\omega = 2\pi f$ which we can substitute in:

$$2\pi f_n = \frac{n\pi v}{L}$$

$$f_n = \frac{nv}{2L}$$

Since n must be an integer, the string can only vibrate at frequencies that are whole number multiples of its fundamental frequency, f_1

$$f_n = nf_1$$

Therefore, n=2 gives its second harmonic, $2f_1$

And n=3 gives its third harmonic, $3f_1$

Musical intervals are simply the ratios between the frequencies. An octave is the ratio between the 2nd and 1st harmonic, 2:1, e.g., if the frequency of a string is 440Hz (A), then one octave above it would vibrate at 880Hz. A perfect 5th is the ratio between the 3rd and 2nd harmonic, 3:2, $\frac{f_3}{f_2} = \frac{3f_1}{2f_1} = \frac{3}{2}$

For example, if a string vibrates at 256Hz (C), then $256 \times \frac{3}{2} = 384$ Hz, which is a G (5 notes above a C, within the C major scale, i.e. its perfect 5th)

As a general rule, Hermann von Helmholtz discovered that combinations of pitches whose harmonic series overlap tend to sound more consonant than those who do not.

3.1 The prime number conflict

Pythagoras discovered that it was possible to make a musical scale by continuing through the circle of fifths. Say we start at C, which has a frequency of 256Hz. We obtain the frequency of its perfect fifth (G) by multiplying by $\frac{3}{2}$. Then repeat, finding the frequency of the perfect fifth of G, which is D. His results were as follows:

$$C\ 256\ \rightarrow G\ 384\ \rightarrow D\ 576\ \rightarrow A\ 864\ \rightarrow E\ 1296\ \rightarrow B\ 1944\ \rightarrow F\#\ 2916\ \rightarrow C\#\ 4374\ \rightarrow G\#\ 6561\ \rightarrow D\#\ 9841.5\ \rightarrow A\#\ 14762.25\ \rightarrow F\ 22143.375\ \rightarrow C\ 33215.0625$$

Using what we know about the 2:1 ratio of octaves, we should be able to keep dividing 33215.0625 by 2 until we reach our original frequency of C (256Hz).

Instead, dividing by 2 seven times (or simply dividing by 2^7) results in a frequency of 259.5 Hz (to 4 significant figures), rather than 256Hz associated with C.

We can explore this by setting the problem up as an equation:

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

$$3^n = 2^{m+n}$$

However, by the fundamental theorem of arithmetic, every positive integer greater than 1 has a unique prime factorisation.

The number 3^n has a prime factorisation consisting of the prime factor 3 multiplied by itself n times. Similarly, the number 2^{m+n} has a prime factorisation consisting of the prime factor 2 multiplied by itself $(m + n)$ times.

Since 2 and 3 are distinct primes, these 2 factorisations can never be equal for any positive integer n and m.

This means that no finite stack of fifths will ever land exactly on an octave.

Instead, the Pythagorean system is based on the fact that $2^{19} \approx 3^{12}$, or likewise $(\frac{3}{2})^{12} \approx 2^7$, which means that going up 12 fifths and then down 7 octaves brings you *almost* back to where you started. (we will explore why 7 and 12 are chosen as approximations for n and m in 3.3).

The small discrepancy between these two values is called the Pythagorean Comma and is defined as the ratio between 3^{12}: 2^{19}.

Note	C	C#	D	D#	E	F	F#	G	G#	A	A#	B	C
Interval	Unison	Minor second	Major second	Minor third	Major third	Perfect fourth	Augmented fourth	Perfect fifth	Minor sixth	Major sixth	Minor seventh	Major seventh	Octave
Ratio	1:1	256:243	9:8	32:27	81:64	4:3	729:512	3:2	128:81	27:16	16:9	243:128	2:1

Figure 2: The ratio of frequencies (relative to the frequency of a C) for notes in Pythagorean tuning, derived from the simple 3:2 and 2:1 ratios between perfect fifths and octaves.

We can also see from figure 2 that the ratio between every other note (what musicians call a whole step,W) is 9:8, and the ratio between every adjacent note (a half step,H) is 256:243.

Two half steps do not equal one whole step, $H^2 \neq W$

And similarly, if we do $\frac{W}{H^2} = \frac{9/8}{(\frac{256}{243})^2} = \frac{3^{12}}{2^{19}}$ which you will recognise as the Pythagorean Comma.

This means that going from C to C# and then going from C# to D (two half steps) is different from going from C to D (one whole step).

So, although the Pythagorean tuning system produces pure sounding fifths and octaves, it has significant drawbacks:

- two half steps do not equal one whole step.
- the circle of fifths does not close, leading to inconsistencies when **enharmonic** equivalents are no longer equivalent

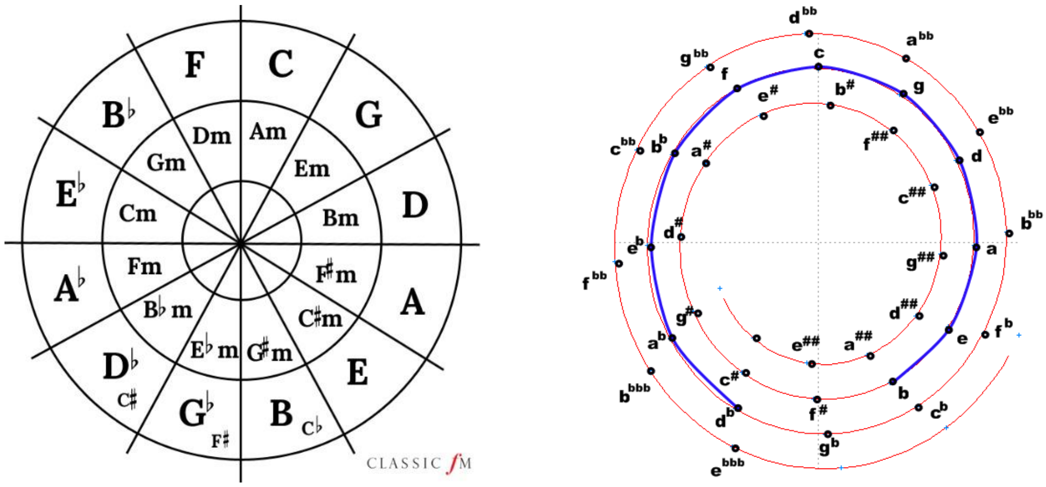

Figure 3: musicians often refer to the 'circle of fifths', when in reality it acts like more of a spiral with Pythagorean tuning.

3.2 Equal temperament

Equal temperament is the modern solution to the Pythagorean comma; rather than stacking perfect fifths, each equal temperament divides the octave into twelve perfectly even semitones.

This would mean each half step would increase by a factor of $\sqrt[12]{2}$ so that $H^2 = W$ and an octave retains the 2:1 ratio. However, this compromises the 3:2 ratio of a perfect fifth, instead it would be a multiple of the fundamental frequency by $2^{\frac{7}{12}} \approx 1.49831 \ldots \neq 1.5$

3.3 Using continued fractions

We have shown there are no integer values of n and m that will satisfy the equation, so instead we can find integer values which give approximately the right results.

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

$$m = n \log_2 1.5$$

$$\frac{m}{n} = \log_2 1.5 \approx 0.5849625 \ldots$$

We can do this using continued fractions to generate a sequence of convergents, which represent the optimal ratios of m:n.

$$x_0 \approx 0.5849625 \ldots$$

The integer part, a_0: $\lfloor 0.5849625 \ldots \rfloor = 0$, remainder: $0.5849625 \ldots - 0 = 0.5849625 \ldots$

First inversion: $x_1 = \frac{1}{0.5849625\ldots} \approx 1.7095113 \ldots$

Integer part, a_1: $\lfloor 1.7095113 \ldots \rfloor = 1$, remainder: $1.7095113 \ldots - 1 = 0.7095113 \ldots$

Second inversion: $x_2 = \frac{1}{0.7095113\ldots} \approx 1.409421 \ldots$

Integer part, a_2: $\lfloor 1.409421 \ldots \rfloor = 1$, remainder: $1.409421 \ldots - 1 = 0.409421 \ldots$

Third inversion: $x_3 = \frac{1}{0.409421\ldots} \approx 2.4424726 \ldots$

Integer part, a_3: $\lfloor 2.4424726 \ldots \rfloor = 2$, remainder: $2.4424726 \ldots - 2 = 0.4424726 \ldots$

Continuing this process gives us $a_4 = 2, a_5 = 3, a_6 = 1, a_7 = 5$

We can write this continued fraction as [0;1,1,2,2,3,1,5,...]

$$C_n = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ldots + a_n}}}}$$

$C_1 = 0 + \frac{1}{1} = 1$ (Trivial solution)

$C_2 = 0 + \frac{1}{1+\frac{1}{1}} = \frac{1}{2}$ (A 2-note scale)

$C_3 = 0 + \frac{1}{1+\frac{1}{1+\frac{1}{2}}} = \frac{3}{5}$ (A 5-note scale, known as the pentatonic scale)

$C_4 = \frac{7}{12}$ (The standard Western 12-note chromatic scale)

$C_4 = \frac{24}{41}$

$C_5 = \frac{31}{53}$ (A mathematically superior scale)

Where the numerator of each fraction represents the position of the perfect fifth, and the denominator represents the number of equally spaced notes in the scale.

But what does any of this mean? The fifth convergent, our 'mathematically superior scale' would consist of dividing our 2:1 octave into 53 evenly spaced notes(if using equal temperament). Because it utilizes equal temperament, it closes the circle of fifths, and, as we will shows in 3.4, produces much purer sounding intervals.

However, a scale of this size has downsides: logistically it would be a nightmare to play - imagine a keyboard requiring 53 keys for every 12 keys of a typical keyboard we use today. Additionally, with so many notes in an octave, notation can get confusing.

3.4 Cents

To quantify the error in different scales, we use the *cents* system, first introduced in 1875 by Alexander Ellis, which divides the octave into 1200 logarithmic units called 'cents'. The size of an interval in cents between 2 frequencies f_1 and f_2 is given by:

$$n = 1200 \times \log_2(\frac{f_2}{f_1})$$

Each whole tone is 200 cents, while each semitone is 100 cents.

A mathematically perfect fifth is $1200 \times \log_2\left(\frac{3}{2}\right) = 701.955 \ldots$ cents.

However, using the modern equal temperament, a perfect fifth is $1200 \times \log_2\left(2^{\frac{7}{12}}\right) = 700$ cents. This is the cost of using equal temperament, each fifth is slightly flattened by 1.955 cents (albeit only roughly one fiftieth of a semitone)

This produces a $\frac{701.955-700}{701.955} \times 100 = 0.2785\%$ difference compared to our mathematically perfect fifth.

Instead, if we use the 53-note scale, a perfect fifth is $1200 \times \log_2(2^{\frac{31}{53}}) = 701.887$ cents, producing a $\frac{701.955-701.887}{701.955} \times 100 = 0.009687\%$ difference, which is almost 30x smaller than the difference produced by equal temperament. This means that its perfect fifth would vibrate at a much more similar frequency to the mathematical perfect fifth, creating a 'purer' sound.

4.Conclusion

Ultimately, the beautiful tragedy of the Pythagorean comma should not be seen as a flaw to be corrected, but a ground from which musical diversity can grow. If the laws of prime factorisation allowed for a perfect circle of fifths, our musical system would be closed, instead the mathematical impossibility of a perfect scale creates a stepping stone into a new world of creativity. We have briefly explored just a few of the many tuning systems and how mathematics influences them, but importantly, how it inspires new systems, like the 53-note scale derived from increasingly accurate approximations. And when someone argues that 'music is maths', I believe that despite the undeniable hidden structures linking them, the greatest link is the creativity they provoke.

Musical terms glossary:

Chromatic:a musical scale consisting of all 12 pitches within an octave.

Enharmonic:Notes with the same pitch (same frequency), but different names (e.g F# and G♭)

#:This symbol means ‘sharp’, a note one half step up from the original note

♭:This symbol means ‘flat’, a note one half step down from the original note

Word count: 1976

Bibliography:

Anon., n.d. [Online]
Available at: https://mathcs.holycross.edu/~groberts/Courses/MA110/Lectures/PythScale-web.pdf

Anon., n.d. [Online]
Available at: https://www.jjensen.org/spiral5ths/Spiral5ths.html

Benson, D., 2008. *Music: A Mathematical Offering.* Aberdeen : s.n.

Krishnan, G. G., n.d. *Continued Fractions.* [Online]
Available at: https://pi.math.cornell.edu/~gautam/ContinuedFractions.pdf