The Unexpected Number Theory and Algebra of Musical Tuning Systems

or, Several Ways to Compute the Numbers 5,7,12,19,22,31,41,53, and 72

Matthew Hawthorn

"Music is the pleasure the human soul experiences from counting without being aware that it is counting." -Gottfried Wilhelm von Leibniz (1646-1716)

"All musicians are subconsciously mathematicians." -Thelonius Monk (1917-1982)

1 Physics

In order to have music, we must have sound. In order to have sound, we must have something vibrating. Wherever there is something virbrating, there is the wave equation, be it in 1, 2, or more dimensions.

The solutions to the wave equation for any given object (string, reed, metal bar, drumhead, vocal cords, etc.) with given boundary conditions can be expressed as a superposition of discrete *partials*, modes of vibration of which there are generally infinitely many, each with a characteristic frequency.

The partials and their frequencies can be found as eigenvectors, resp. eigenvalues of the Laplace operator acting on the space of displacement functions on the object. Taken together, these frequencies comprise the *spectrum* of the object, and their relative intensities determine what in musical terms we call *timbre*.

Something very nice occurs when our object is roughly one-dimensional (e.g. a string): the partial frequencies become *harmonic*. This is where, aptly, the better part of *harmony* traditionally takes place. For a spectrum to be harmonic means that it is comprised of a *fundamental* frequency, say f, and all whole number multiples of that frequency:

$$f, 2f, 3f, 4f, \ldots$$

It is here also that number theory slips in the back door.

Another way to make partials harmonic is to subject the object to a periodic driving force (e.g. blowing a reed, bowing a violin string). Here I equivocate and use 'partial' to refer to a percieved sinusoidal component, which under forcing is no longer necessarily a natural mode of the object in question. These 'partials' describe the resultant sound rather than the underlying physics, and Fourier analysis is our standard tool for isolating them.

A quick survey will reveal that the great majority of pitched instruments in common usage the world over fall into one or both of the two above categories, with some notable exceptions, and a

few exceptions which prove the rule (bell-casting for instance has been elevated to the level of exact science and high art in an attempt to banish the pesky inharmonic partials).

2   Psychoacoustics

It is a remarkable fact of nature that inside of each of our heads are two little instruments which, to an approximation, perform a real-time Fourier transform of all incoming sound. These would be the Cochlea, the little snail-shell-shaped organs of our inner ear. Every incoming frequency has a resonant node inside the cochlea, a specific place where it tickles little microscopic hairs which then pass a neural signal to the brain. The lower the frequency of an incoming sinusoidal wave, the further into the cochlea is the resonant node.

Now, the cochlea, like any sensing instrument, has a resolution. Frequencies too close together will fail to be distinguished. However, due to the trigonometric identity

$$\cos(\alpha t) + \cos(\beta t) = 2\cos(\tfrac{\alpha+\beta}{2}t)\cos(\tfrac{\alpha-\beta}{2}t),$$

a pair of nearby frequencies sounded simultaneously will be heard as a single frequency at the average (the left term on the right side), modulated periodically in volume by a frequency equal to their difference (the right term; we hear twice the frequency in the parentheses because the sign of the envelope is irrelevant). This modulatory effect is called *beating* by acousticians.

Now, as the superposed frequencies get farther apart and the beat rate increaes, we lose track of the individual beats and experience a sensation that psychoacousticians call "roughness". Most people find this unpleasant, so a worthy goal of designing a tuning system, loosely speaking, is to minimize this effect as much as possible by matching as many partials as closely as we can between small subsets of pitches- *intervals* (pairs of pitches) and *chords* (3 or more pitches) in musical terminology.

3   Number theory

"Mathematics and music, the most sharply contrasted fields of scientific activity which can be found, and yet related, supporting each other, as if to show forth the secret connection which ties together all the activities of our mind, and which leads us to surmise that the manifestations of the artist's genius are but the unconscious expressions of a mysteriously acting rationality."
-19th century German physicist Hermann von Helmholtz

The nice thing about harmonic partials is that the rich multiplicative structure of the natural numbers allows us to match infinitely many of them at once between pairs (triplets, etc.) of pitches. This occurs exactly when $\frac{f_1}{f_2}$ is rational, for a pair of fundamentals f_1, f_2. For example, consider f and $\frac{3f}{2}$. The partials are:
$$f, 2f, 3f, 4f, 5f, 6f \ldots \text{ and}$$
$$\tfrac{3f}{2}, 3f, \tfrac{9f}{2}, 6f, \tfrac{15f}{2}, 9f \ldots$$

Fully half of the partials of the higher frequency are shared with those of the lower frequency, and a third of the partials of the lower are shared with the higher. The musical sensation of all these reinforcing partials is what we call *consonance*. We would be lucky to match more than one or two partials of an inharmonic timbre so well. Below is a table of some harmonic intervals, with

Euler's theoretical measure of dissonance, his "gradus suavitatis" $(GS(\frac{p}{q}) = 1 + \sum_1^k e_i(p_i - 1)$ where $pq = p_1^{e_1} p_2^{e_2} \ldots p_3^{e_3})$, labeled for comparison.

Table 1: Some consonant intervals (frequency ratios)

Ratio	Euler's GS	Musical name
$\frac{2}{1}$	2	Octave
$\frac{3}{2}$	4	Perfect Fifth
$\frac{4}{3}$	5	Perfect Fourth
$\frac{5}{4}$	7	Major Third
$\frac{5}{3}$	7	Major Sixth
$\frac{6}{5}$	8	Minor Third
$\frac{8}{5}$	8	Minor Sixth
$\frac{9}{5}$	9	Minor Seventh
$\frac{7}{4}$	10	Harmonic Seventh
$\frac{7}{6}$	10	Septimal Subminor Third
$\frac{7}{5}$	11	Septimal Tritone (Augmented Fourth)
$\frac{9}{7}$	11	Septimal Supermajor Third (Diminished Fourth)

Now let's talk about the problem of designing a good tuning system.

We define a tuning system as a countable subset G of $\mathbb{R}^+$, corresponding to the frequencies we permit ourselves to use in musical composition. We use G for *gamut* (music) or *group* (math).

Now, for a subset C of rationals we consider consonant (such as table 1 above), we would like it to be the case that

$$\text{for all } f \in G \text{ and } c \in C, \text{ we have } fc \in G \text{ and } fc^{-1} \in G$$

This ensures that we can play the same harmony (or melody) starting from any pitch in G. It's clear then that C can be treated as a set of *generators* for G as a multiplicative group; a suitable gamut can be constructed from a single base pitch, let's say 1 for simplicity, by multiplying by arbitrary integer powers of the generators in C.

Now, a canonical basis for this finitely generated Abelian group is the set of primes dividing the rationals in C. Call this set P. Unique factorization in the integers tells us that this is a *free* basis for G as an Abelian group (or $\mathbb{Z}$-module equivalently if you like); there is no nontrivial dependence between the generators since $p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k} = 1 \Rightarrow e_i = 0 \ \forall i$. We can write

$$G = \{p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k} \mid p_i \in P, e_i \in \mathbb{Z} \ \ \forall i\}$$

I'll call this $\mathbb{Q}|_P$, the rationals whose numerators and denominators factor into primes in P (*p-smooth* rationals is one way to say this in math-speak, and *p-limit* harmony is another way to say it in music theory, where $p = max(P)$). If C is finite then so will be our basis P, and we have G a finitely generated $\mathbb{Z}$-module (Abelian group equivalently). I like the $\mathbb{Z}$-module perspective because it

allows me to think additively, inside of an integer lattice (using the exponents as coordinates), rather than multiplicatively inside of $\mathbb{Q}^+$.

Now, a problem that immediately reveals itself is that, if our basis P has more than two multiplicatively independent generators in it (*which it will if the generators are primes, owing to unique factorization!*) then $G = \mathbb{Q}|_P$ is dense in $\mathbb{R}^+$, in the usual topology (or simply in $\mathbb{R}$ after taking logarithms, if you like). This is a problem because we can't have infinitely many keys on a piano, frets on a guitar, keys on a clarinet, etc. In any given finite span of pitch (an octave, say), we must have only finitely many available pitches (continuously pitched instruments like fretless strings, the trombone, the human voice, etc, don't have this issue, but of course they're only a subset of musically useful instruments).

Thus begins the long historical saga of the conundrum of tuning, which has played out independently in many cultures, each of which has come up with a unique set of solutions. From here we'll explore the uniquely Western European approach to the problem, following basically the modern theoretical framework pioneered by the Dutch physicist Adriaan Fokker in the 1960's (though existing in specific instances of practical usage long before), and then we'll use that framework to explain how specific tuning systems might have arisen in other cultural contexts.

4 The $\mathbb{Z}$-module Homomorphism Perspective

It should be clear to us now that we need to reduce the *rank* (dimension) of $\mathbb{Q}|_P$ somehow. The usual way to do this with a module (or vector space) is with a homomorphism (or linear mapping) into a module (vector space) of lower rank (dimension). Such a mapping is completely determined by the image of the generator set. Let's be concrete here and take our generator set to be the smallest three primes:

$$P = \{2, 3, 5\}.$$

We're now doing 5-limit harmony, or equivalently working in $\mathbb{Q}|_{\{2,3,5\}}$, the 5-smooth rationals. This is the domain of nearly all Western music harmony since the early renaissance, sometime in the 15th century. We're in a free $\mathbb{Z}$-module of rank 3, since we have three generators, and we need to mod out a kernel (think null space) of rank 2 to get an image of rank $3 - 2 = 1$, which will finally be sparse in $\mathbb{Q}^+$. This is the key idea first described by Adriaan Fokker in the mid-20th century (though it had been in use intuitively since the Renaissance). Fokker's language was less mathematical than ours perhaps, but the guts of the idea are there in his work.

A nice way to start building this kernel (nullspace) would be to find a really small (in the sense of being close to 1 in $\mathbb{Q}^+$) rational in $\mathbb{Q}|_P$, say c for *comma* (a musical term for a really small pitch interval), and then fudge the basis primes making up c just enough that c becomes 1 exactly. We want c small because intuitively, the smaller c is, the less we'll have to fudge. This fudging is called *tempering* in music theory, and the result is called a *temperament*. So we find a small (close to 1) 5-smooth rational c and put $\langle c \rangle$ (the submodule/subspace that c generates/spans) in our kernel, which now has rank 1.

What is a good candidate for our comma c? Well, one reasonable criterion would be to choose it to be as 'simple' as possible for its size: roughly speaking, the denominator should be as small as possible, requiring relatively few prime generators to reach. Intuitively, the simpler the kernel basis,

the simpler the relations or $\mathbb{Z}$-dependencies we're introducing, and thus the simpler the musical logic of our tempered tuning system. The simplest rational below a given small size will take the form

$$\frac{n+1}{n},$$

which we call *superparticular*. You can't find a rational closer to 1 than any given superparticular rational without increasing the denominator.

Now, it's a beautiful result in number theory that there are only finitely many superparticular rationals in $\mathbb{Q}|_P$ for a finite set P. Størmer's Theorem (after Norwegian mathematician Carl Størmer) gives a bound on the number of such, and an effective method of computing them all for any given set P using a set of Pell equations. In the 5-limit ($P = \{2, 3, 5\}$), for example, the superparticulars are:

$$\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \frac{6}{5}, \frac{9}{8}, \frac{10}{9}, \frac{16}{15}, \frac{25}{24}, \frac{81}{80}$$

The smallest of these is called the *syntotic comma*, which we can write as

$$\frac{81}{80} = 2^{-4}3^45^{-1} \text{ with module coordinates } (-4, 4, -1). \tag{1}$$

It was the first to be explicitly tempered to unison in the West beginning in the Renaissance, which is when people started accepting 5-limit intervals into the harmonic fold.

The question of the *optimal* tuning for any given temperament (that is, the question of which primes to fudge, and by how much) is beyond the scope of this paper but depends mostly on a choice of weights for the basis primes and a norm on the space spanned by their images in log-frequency space (just $\mathbb{R}^n$ with $n = |P|$), which will give us an optimal 'distance' to the point $(\log(p_1), \log(p_2), \ldots \log(p_k))$, the canonical mapping of the prime basis into log space (the "pure tuning" or simply the identity homomorphism). The Moore-Penrose pseudoinverse from linear algebra can make an appearance here, if we use the L_2 norm, but we'll skip that for the sake of clarity.

For now, let's just assume that we want pure octaves ($\frac{2}{1}$) and pure major thirds ($\frac{5}{4}$). Octaves ($\frac{2}{1}$) have always been sacred and are seldom tempered, owing to the musical phenomenon of *octave equivalence*, the sensation of pitches separated by octaves being in some sense "the same". This principle is employed in nearly every world musical culture. So 2 will stay put, and we'll also keep 5 where it is for simplicity, but we'll fudge 3 a bit. Putting in x for 3 in equations (1) we have

$$2^{-4}x^45^{-1} = 1$$

which yields $x = 3(\frac{81}{80})^{\frac{-1}{4}}$. In other words, we've flattened (lowered) the 3 generator by a quarter of a syntotic comma, multiplicatively speaking. 3 is the *perfect twelfth* in music, or *perfect fifth* when transposed down an octave to $\frac{3}{2}$, so we're tuning our fifths a little flat, in musical terms. We've invented *quarter-comma meantone* (QCM from now on), the keyboard tuning standard of European music from the 16th to the 18th century. Some church organs retained this temperament well into the mid-1800's. It turns out that QCM is the minimax (L_∞) tuning solution on the set of 5-limit intervals in table 1, assuming the syntotic comma is tempered to unison (that is, when we mod out the submodule it generates).

The problem we now face is that we're still rank 2: we have a pure 2 and a flat 3 as generators (5 is no longer needed; $\frac{81}{80} \approx 1 \Rightarrow 5 \approx 3^42^{-4}$). To bring our temperament down to rank 1, we need

another comma in our kernel; at this point we could introduce another $\mathbb{Z}$-dependence between the 2 and 3 generators. Now, any beginning student of music theory will come across the "circle of fifths" pretty quickly. This is a convenient fiction: $2^n = 3^m$ will never be solvable owing to unique factorization; the 'circle' doesn't close in pure tuning. But we're not even working with a pure 3 any more, because we've accepted this QCM flat 3 as a substitute. It too is $\mathbb{Z}$-independent from 2, but we might be able to fudge it just a little bit more to introduce a new dependence. Let x be our flat QCM 3. We want $x^n \approx 2^m$, or $n \log(x) \approx m \log(2)$, or $\frac{log(x)}{log(2)} \approx \frac{m}{n}$ This is a problem of *Diophantine Approximation*, and we can solve it by computing the continued fraction of $\frac{log(x)}{log(2)}$, which is

$$\frac{log(x)}{log(2)} = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \dots}}}}}$$

with succesive approximants

$$\frac{3}{2}, \frac{8}{5}, \frac{11}{7}, \frac{19}{12}, \frac{30}{19}, \frac{49}{31} \dots \tag{2}$$

So if we sharpen or flatten our slightly flat QCM 3 a little more, we can make it close at the octave after 7, 12, 19, or 31 steps respectively. Each of these is equivalent to adding another unique comma to our kernel. At that point we've achieved rank 1, sparse in frequency space, and thus useful on fixed pitch instruments. And we have a cyclic group structure, highly versatile melodically and harmonically.

These numbers are not new. The 19-tone solution was proposed by theorists Guillame Costeley and Francisco de Salinas in the 16th century. 31 was proposed by Lemme Rossi and Christian Huygens in the 17th century, and revived in the 20th century by Adriaan Fokker, whose theory we're now exploring. Fokker's ideas actually spawned a short-lived school of Dutch 31-tone composition. But this is not the course that Western music as a whole has taken, so let's descend back down the cardinalities and consider 12. The 12-per-octave solution was described independently by Chinese mathematician Zhu Zaiju in 1584 and by Flemish mathematician Simon Stevin in 1585. Incidentally, *12-tone equal temperament* is where Western music has finally settled, but it took a long time because many Renaissance theorists were unwilling to accept the impurity of tuning that is implied by adding yet another comma (major thirds are quite sharp in 12-tone equal temperament, for example).

Basically the rank-1 12-tone solution proceeds by adding the *Pythagorean comma* $\frac{3^{12}}{2^{19}}$ to the kernel along with the syntotic comma. Another (mathematically equivalent) way to get to 12 steps per octave is to add the comma $\frac{125}{128} = 2^{-7}5^3$ to our kernel. We can figure out how many pitch classes there are modulo octave equivalence by taking a determinant of the 'vectors' corresponding to our kernel generators, which gives us the volume of the parallelogram they span, and hence the number of lattice points inside it. As expected, we get

$$\begin{vmatrix} 4 & -1 \\ 0 & 3 \end{vmatrix} = 12.$$

Arabic music took another route, probably influenced by ancient Greek texts that Arabic scholars preserved and studied during the middle ages. *Pythagorean* tuning is the standard theoretical starting point there, a theoretically rank-2 system allowing only the primes 2 and 3, and no tempering. Keeping 3 pure now and looking for near-dependencies between 2 and 3–solutions to $\frac{log(3)}{log(2)} \approx \frac{m}{n}$ as

before with the flat QCM 3–we compute the continued fraction expansion of $\frac{\log(3)}{\log(2)}$, and get

$$\frac{\log(3)}{\log(2)} = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{3 + \dots}}}}}$$

with succesive approximants

$$\frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \frac{11}{7}, \frac{19}{12}, \frac{46}{29}, \frac{65}{41}, \frac{84}{53} \cdots \tag{3}$$

Indeed, the Syrian violinist and music theorist Twfiq Al-Sabagh has proposed 53 tones per octave as a tuning standard for Arabic music. Traditional Arabic theory specifies 9 'commas' per whole tone ($\frac{9}{8}$), a property satisfied when we divide the octave into 53 equal parts. Incidentally, 53 pitch-per-octave systems have been described by European (Nicholas Mercator, 17th century, Hermann von Helmholtz, 19th century) and Chinese (Ching Fang, 1st Century) theorists as well.

As an interesting aside, note that the denominators 5 and 7 occurred in both of our lists of approximants, sequences (2) and (3) above. The intervals of the octave and the fifth (or twelfth), corresponding to the prime frequency ratios 2 and 3 respectively, are the most basic and strongest consonances available to our ears, and being low in the harmonic series, will suffer relatively little from inharmonicity (natural physical 'mistuning'). Even so, quite a wide range of (mis)tunings of 'fifths' will yield 5- and 7-tone scales naturally. Thus, we might expect scales of 5 and 7 pitches per octave to be quite common the world over. This is indeed the case. The pentatonic minor and septatonic major, minor, and other modes are known to any trained Western musician. But pentatonic and septatonic scales also reign in Indonesian Gamelan ensemble music ('slendro' and 'pelog' scales, respectively). Pentatonic and septatonic tunings are common in the xylophone-like timbila and the mbira ('thumb piano'), ubiquitous throughout east Africa. Pentatonic scales are the most commonly used in traditional Chinese music, and septatonic scales dominate in the Arabic Maqam and Indian Raga traditions. Amazingly, several bone flutes playing 5- and 7-tone scales and dating to about 6000 BCE have been found in Jiahu in Henan Province, China. The specific tunings vary widely across all these cultures and through history, but the cardinalities are remarkably constant.

Arguably the next frontier for Western harmony lies in the 7-limit. This puts us in a rank-4 $\mathbb{Z}$-module to start, with generator set $\{2, 3, 5, 7\}$, and we will thus need a rank-3 kernel (so 3 $\mathbb{Z}$-independent commas) to get a rank-1 temperament. Following Størmer, there are only a finite number of 7-smooth superparticular rationals. The smallest four of these are

$$\frac{126}{125} = 2^1 3^2 5^{-3} 7^1, \frac{225}{224} = 2^{-5} 3^2 5^2 7^{-1}, \frac{2401}{2400} = 2^{-5} 3^{-1} 5^{-2} 7^4, \frac{4375}{4374} = 2^{-1} 3^{-7} 5^4 7^1,$$

the last two of which are positively microscopic from a musical perspective! Choosing the larger 3 and the smaller 3 respectively and computing the number of equivalence classes modulo octave equivalence (so ignoring the 2 coordinate) with the determinant again gives:

$$\begin{vmatrix} 2 & -3 & 1 \\ 2 & 2 & -1 \\ -1 & 2 & 4 \end{vmatrix} = 31, \text{ and } \begin{vmatrix} 2 & 2 & -1 \\ -1 & 2 & 4 \\ -7 & 4 & 1 \end{vmatrix} = 72.$$

So we might surmise that 31-tone-per-octave equal temperament should do quite well in the 7 limit (and it does!). This is one reason why Fokker was advocating for it; he wanted to see the 7th

harmonic accepted as musically useful in Western classical music. 12-tone equal temperament by comparison does a pretty bad job at approximating 7-limit consonances, since $2^{\frac{34}{12}}$ is almost a third of a generator (*semitone* in Western music terms) away from 7.

Incidentally 72 pitches per octave has been adopted as a tuning standard by some Turkish Qanun builders (a zither-like plucked string instrument). Surely they came to it from a much different direction, but it's food for thought. It is a nice coincidence that 72 is a multiple of 12; modern Western notation and instruments could theoretically adapt to it without too much fuss. This is as far as we will go here. As if that were not enough mathematics, in the sequel, the Riemann Zeta function makes an appearance. For now, just note that (following the work of mathematician and now self-styled music theorist Gene Ward Smith), if we compute the integral of the absolute value of zeta between successive nontrivial zeroes, singling out the pairs of zeroes which yield an increasing sequence of values of this integral, then multiply their imaginary parts by $\frac{\log(2)}{2\pi}$ and choose the unique integer lying between these, we get the sequence

$$2, 5, 7, 12, 19, 31, 41, 53, 72 \ldots,$$

all of which we have since met (barring the trivial 2) by other means.

5 Further Reading

You can read Adriaan Fokker's original 1969 paper in English at

http://www.huygens-fokker.org/docs/fokkerpb.html

and many more of his papers and those of others (though not generally in English) at

http://www.huygens-fokker.org/microtonaliteit/literatuur.html

For a great deal of more modern exposition on the mathematical theory explained here (and much more), check out

http://xenharmonic.wikispaces.com/Regular+Temperaments
and   http://x31eq.com/tuning.htm

Xenharmonic is an amazing wiki to just poke around in. Being a wiki, it is written by all kinds of people, so the quality, tone, and technicality vary widely. A lot of the mathematically deep stuff there is coming from Gene Ward Smith, and you can read more about his Zeta function musings at

http://xenharmonic.wikispaces.com/The+Riemann+Zeta+Function+and+Tuning

Incidentally, Peter Buch came across this Riemann Zeta business independently, and approaches it from a slightly different and more elementary angle in

http://xenharmonic.wikispaces.com/file/view/Zetamusic5.pdf

If you're interested in delving into the psychoacoustics/physics foundations or the multitude of other places mathematics can enter music theory beyond tuning I highly recommend the book *Music: A Mathematical Offering*, which the author Dave Benson of the University of Aberdeen shares graciously at

http://homepages.abdn.ac.uk/mth192/pages/html/music.pdf