

Taming the Pythagorean wolf

4775
GRIFFEL, D.H.

D.H. Griffel

School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

Introduction

Why does a conference on "Science and Imagination" include a paper on musical scales? It is an obscure subject, of interest mainly to musicians, and a small minority of musicians at that. Yet it was once a central problem in science, attracting the attention of Kepler, Galileo, Stevin, Descartes, Huygens, Hooke, Newton, Leibniz, Young, Helmholtz, and many others.

The problem of musical scales is not straightforward. It resembles many difficult problems in the applied and environmental sciences, in that it involves applying quantitative criteria in areas which are difficult to quantify completely. The problem of tuning illustrates (in a context free of difficult value judgements) how wrong one can be.

Musical notes and intervals

The basic physical facts are sketched here very briefly. For more information see refs 1 or 2.

A musical note is a vibration: higher notes correspond to faster vibrations. Most Western music uses only a few of the infinitely many possible notes; they are given names based on the letters A, ...,G.

If one note has vibration frequency twice that of another, we hear them as essentially the same note, in different registers. The two notes are given the same name. Thus the notes of frequency 220 and 440 hertz are both called A (sometimes a and a' respectively).

Two notes of frequencies f and $2f$ are said to be an *octave* apart. It is frequency ratios that determine the musical effect of a pair of notes sounded together. An *interval* is a pair of notes with a given frequency ratio: all intervals with the same ratio sound alike, regardless of the absolute frequencies of the notes.

European music splits an octave into 12 parts, each of which is called a semitone. Thus there are 12 different notes between one A and the next A. They are laid out on a keyboard as shown in Figure 1. Some of them are considered more important, and have white keys and simple names A, B, ...,G. The others have black keys, and names which relate them to the nearest white note. Thus $F^\sharp$ means the black note a semitone above F, and $B^\flat$ means the black note a semitone below B.

The principal intervals and the circle of 5ths

The fundamental interval is the octave, defined above. The next most important are shown in Table 1. Disregard the last column for the present, and think of the intervals as defined by their size in semitones.

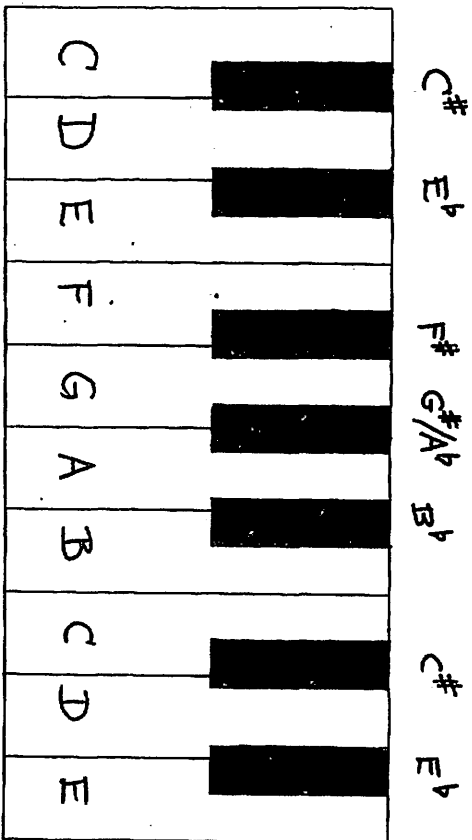


Figure 1 The keyboard.

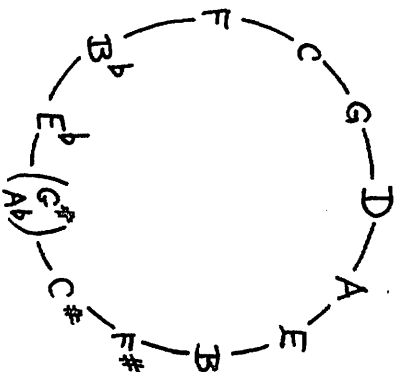


Figure 2 The circle of fifths.

Thus a 5th up from C is G (see Figure 1). It is five white notes up from C, hence the name "5th". But a 5th up from B is the black note F[#], seven semitones away from B. Similarly, a 3rd up from C is E, a step of three white notes. Again, not every step of three white notes is a 3rd: the defining property of an interval is its frequency ratio, given here in terms of semitones.

Going up in steps of a 5th from A gives A, E, B, F[#], C[#], G[#], E^b, B^b, F, C, G, D, A, The sequence runs through all the notes and then returns to its starting point. This is often called the *circle of 5ths*; see Figure 2.

Table 1 Principal intervals

Interval	Size in semitones	Frequency ratio
Octave	12	2/1
5th	7	3/2
4th	5	4/3
3rd	4	5/4

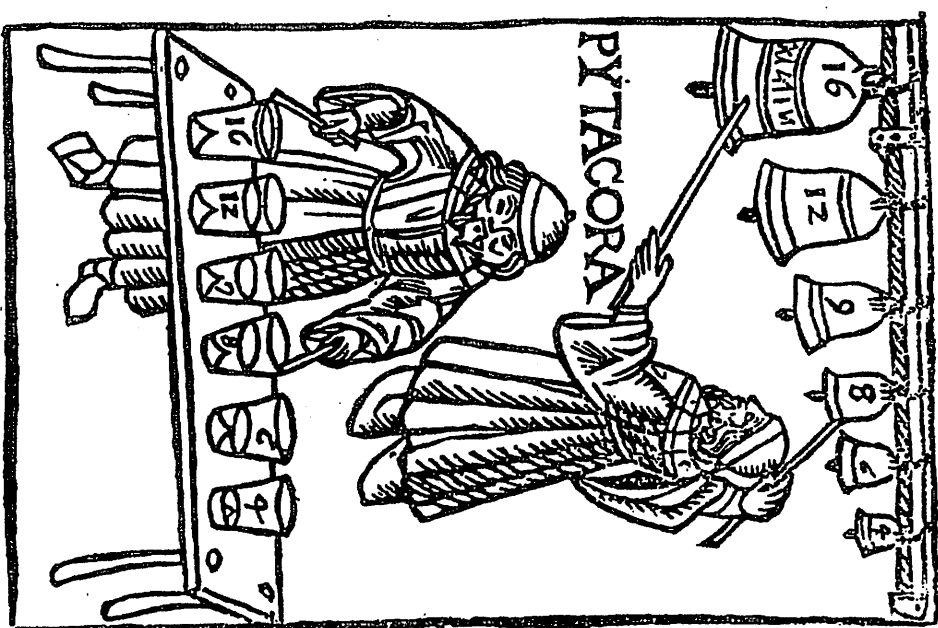


Figure 3 Pythagoras investigating harmonic ratios, according to Gajdus (c. 1500 AD).

Pure intervals

Some intervals and chords sound pleasant, some do not. Experiments (going back to Pythagoras - see Figure 3) reveal a striking and beautiful fact: intervals sound pleasant when their frequency ratios are simple rational numbers, as shown in Table 1. See ref. 3 for the development of an understanding of this; but it is the fact, not the underlying mechanism, that concerns us here.

When an interval is slightly out of tune, an unpleasant quavering sound is heard, called "beats". As the interval is brought into tune, the beats get slower and finally disappear when the interval is pure. See refs 1 or 2 for further information.

The fundamental paradox

Putting intervals together means multiplying their frequency ratios. Proof: if notes X, Y, Z have frequencies x , y , z , then the intervals XY and YZ have ratios y/x and z/y ; combining intervals XY and YZ gives XZ, with ratio $z/x = (y/x)(z/y)$. For example, going up a 5th then a 4th means multiplying the ratios $3/2$ and $4/3$, giving 2. This agrees with the fact that a 5th plus a 4th gives an octave. Inverting an interval (that is, going down rather than up by the given number of semitones) corresponds to inverting the frequency ratio.

We are now in a position to see the fundamental difficulty in tuning a keyboard.

Going up a 5th and then down a 4th is equivalent to going up a tone (two semitones). It follows from the above that the frequency ratio of a tone is $(3/2)/(4/3) = 9/8$. Now, six tones make an octave, which must have frequency ratio 2 - out of tune octaves are musically quite intolerable. But combining six tones gives a frequency ratio $(9/8)^6 = 2.027$. Thus:

It is impossible for all 4ths and 5ths to be perfectly in tune

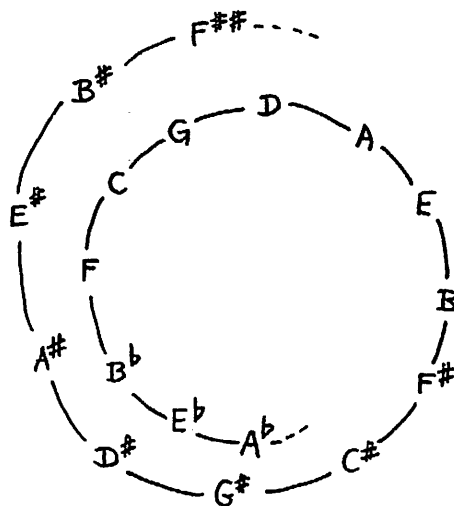


Figure 4 The spiral of 5ths.

The spiral of 5ths

The difficulty can be viewed in terms of the circle of 5ths. Twelve 5ths add up to seven octaves. But 12 pure 5ths give $(3/2)^{12} = 129.7$, whereas seven octaves give $2^7 = 128$. Thus going round a cycle of pure 5ths, starting from C, say, brings one back to a note different from the starting point. It is in fact B[#], which is slightly higher than C when 5ths are tuned pure. The circle of 5ths has become a spiral: see Figure 4.

The Pythagorean wolf

The spiral of 5ths does not fit into a keyboard with only 12 notes to the octave. Some kind of compromise is needed. One solution is to make 11 of the 5ths pure, and distort the last one so as to close the circle of 5ths. The last 5th is then flat by 1.3 percent. It beats about eight times a second (in the tenor register), and sounds really unpleasant. It is called a "wolf 5th", because it howls instead of singing sweetly.

What can one do about this? One answer is to hide the wolf in the forest - among the black notes in Figure 1. Mediaeval and Renaissance music uses only a restricted range of harmonies, so if the wolf 5th is between G[#] and E^b, it will hardly ever be heard.

However, there are other problems.

Pythagorean 3rds

We showed above that when 5ths are pure, a tone has ratio $9/8$. Therefore a 3rd (equal to two tones) has ratio $(9/8)^2 = 1.266$. But this is very out of tune: see Table 1. Thus pure 5ths imply very sour 3rds: they are called Pythagorean 3rds.

The lone wolf 5th could easily be hidden. But with most of the 5ths pure, nearly all the 3rds are badly out of tune. This is a serious problem.

Early mediaeval music was built largely on 4ths and 5ths, and did not use 3rds. Therefore the Pythagorean system was satisfactory. But the development of music using a wider range of harmonies required a system in which 3rds and common chords are tolerably in tune.

Equal temperament

One obvious approach to the problem of tuning is to make all the semitones equal. Since 12 semitones make an octave, each one will then have a ratio of $2^{1/12}$.

It follows that all the 5ths will have frequency ratio $2^{7/12} = 1.498$. The 5ths are said to be "tempered", that is, deliberately made slightly impure. A tuning system with tempered intervals is called a "temperament". The system described here is called "equal temperament" because all 5ths are tempered by the same amount, from 1.5 to 1.498.

The 5ths are only slightly tempered; 1.498 is almost $3/2$. But 3rds in equal temperament have frequency ratio $2^{4/12} = 1.260$, quite badly out of tune. They beat about eight times a second in the tenor register.

On a modern piano the beats are not very noticeable, and we have all grown accustomed to the sound of equal temperament. But on earlier instruments, with a more incisive tone, the beats are much more prominent. So although equal temperament was invented in the sixteenth century, it was not generally used for keyboard instruments until the nineteenth century (when the clear tone of the harpsichord had given way to the blander sound of the piano).

Mean tone

There is no complete solution to the temperament problem. There have been many compromise solutions (see ref. 4). One of the earliest, called "mean tone tuning", tempers the 5ths so as to make the 3rds pure. Because four 5ths make a 3rd (see Figure 2), the error in the Pythagorean 3rd is divided among four 5ths in this system; so each of them needs only a slight tempering, and in practice sounds very nearly pure. This is a good scheme for much Renaissance and early Baroque music. But it has limitations.

All the white notes can be tuned in this way to give pure 3rds and nearly pure 5ths. The black notes can be tuned to give pure 3rds with the white notes. But there are ambiguities. The note marked $G^{\sharp}/A^{\flat}$ in Figure 1 can be tuned to give a pure 3rd with E, or a pure 3rd with C, but not both. If it is a pure 3rd with E, it is called $G^{\sharp}$, and if it is a pure 3rd with C it is called $A^{\flat}$. These two notes are very clearly different; the interval $G^{\sharp}C$ in mean tone has ratio 1.28. This is so far from $5/4$ that $G^{\sharp}C$ in mean tone sounds quite intolerable.

Renaissance and early Baroque music used a limited range of harmonies. Since $G^{\sharp}$ and $A^{\flat}$ were not often needed in the same piece, one could tune to whichever was needed. But as music developed, more flexible temperaments were needed.

Eighteenth century systems

In mean tone, many chords are very nearly pure, but others are practically unusable. Many alternative schemes were devised in the eighteenth century, which even out the inequalities to various extents.

In a typical example (Vallotti's temperament), half of the 5ths are pure and half are slightly impure; five of the 3rds are better than equal temperament and five are worse. The most often used chords are better than in equal temperament, and even the worst are much better than the Pythagorean or mean tone wolves.

Unequal temperaments of this kind have two advantages. Many of the commonest chords are better in tune than in equal temperament. And it follows that chords on different keynotes have different sounds. Therefore music in different keys will sound different. Moving from one key to another is an essential part of the language of classical music. An unequal temperament gives a rougher or sharper sound when the music modulates into remote keys, and the return to the home key near the end of a movement will generally bring a smoother tone-colour. This element of the music is lost in equal temperament.

Bach's "well-tempered clavier"

In 1722, J.S. Bach completed a set of preludes and fugues in all the 24 possible keys. He called it "The well-tempered keyboard", the point being that it needs a temperament in which all chords are acceptably in tune. It used to be thought that "well-tempered" means equally tempered. But looking carefully at the music refutes this.

Bach is rather cautious in using chords which are noticeably impure in typical eighteenth century temperaments. For example, they may be sounded so briefly that there is no time to hear how out of tune they are. Or they may be sounded in one hand while the other hand plays faster-moving notes which distract attention from the impure chord. See ref. 5 for a detailed study, concluding that there is definite evidence that Bach did not intend equal temperament.

Optimisation

The analysis of ref. 5 gives a pattern of "degrees of avoidance" for the various chords; a high degree is assigned to a chord which Bach does not use prominently. One can then look for a temperament for which the pattern of impurities in the various chords matches the pattern of degrees of avoidance. The result may be described as the optimal temperament for playing Bach's well-tempered clavier. It is similar to the eighteenth century unequal temperaments described above.

For more mathematical approaches to optimising temperaments, see refs 6 and 7.

Other kinds of scale?

A completely different line of attack on the temperament problem is to abandon the division of the octave into 12 parts. Equally tempered systems based on dividing the octave into 19, 31, or 53 parts have been studied for several centuries (see ref. 8). But they are still very much a minority taste.

Concluding remarks

The theory should be viewed in the light of practical experience. Instruments tend to be unstable. A harpsichord will not stay perfectly in tune throughout a concert; it responds to the warm and perhaps moist atmosphere of the concert room. So there is little point in calculating temperaments with extreme precision.

But there is another reason for not taking the calculations too seriously. The arguments above are based on the idea that impure chords are a necessary evil, to be avoided as far as possible. But this is an oversimplification. We noted, when discussing eighteenth century systems, that moving between good and bad keys can add to the musical effect of a piece. More generally, a sour chord in an appropriate place can have a powerful emotional effect. Baroque music is all about emotion (as is clear from contemporary writings as well as from the music itself). Unequal temperament is part of the language of this music. A strictly mathematical approach to reducing dissonance misses the musical point.

In many fields, a straight application of quantitative optimisation techniques is likely to go wrong because some important features of the problem are unquantifiable. What is needed is not only mathematical and scientific expertise, but imagination.

References

1. Helmholtz, H., *On the Sensations of Tone*, (Transl.). Longman, London (1885), Dover, New York, (1954).
2. Roederer, J., *Introduction to the Physics and Psychophysics of Music*. Springer, New York (1975).
3. Cohen, H.F., *Quantifying Music*. Reidel, Dordrecht (1984).
4. Lindley, M., "Temperaments" in *The New Grove Dictionary of Music and Musicians*. Macmillan, London (1980).
5. Barnes, J., "Bach's Temperament: Internal Evidence from the Well-tempered Clavier," *Early Music*, 7, 23 (1979).
6. Goldstein, A.A., "Optimal Temperament", *SIAM Review*, 19, 554-562 (1977).
7. Parker, D.B., "Number Harmony", *Mathematical Intelligencer*, 8, 18-21 (1986).
8. Garner, C. and Wilson, R.J., "Musical Block Designs", *Ars Combinatoria*, 16-A, 217-225 (1983).