GRESHAM

COLLEGE

FAGGOT'S FRETFUL FIASCO:

THE UNSUNG GEOMETRY OF MUSICAL SCALES

A Lecture by

PROFESSOR IAN STEWART MA PhD FIMA CMath
Gresham Professor of Geometry

4 October 1995

Faggot's Fretful Fiasco

Ian Stewart

If you look at a guitar, mandoline, or lute — any stringed instrument with frets — you'll see that the frets get closer and closer together as the note gets higher. This is a consequence of the physics of vibrating strings. Today's Western music is based upon a scale of notes, generally referred to by the letter A-G, together with symbols # (sharp) and ♭ (flat). Starting from C, for example, successive notes are

	C#		D#			F#		G#		A#	
C		D		E	F		G		A		B
	D♭		E♭			G♭		A♭		B♭	

and then it all repeats with C, but one octave higher. On a piano the white keys are C D E F G A B, and the black keys are the sharps and flats.

This system is a compromise between confliciting requirements, all of which trace back to the Pythagorean cult of ancient Greece. The Pythagoreans discovered that the intervals between harmonious musical notes can be represented by whole number ratios. They demonstrated this experimentally using a rather clumsy device known as a *canon*, a sort of one-string guitar. The most basic such interval is the *octave*: on a piano it is a gap of eight white notes. On a canon, it is the interval between the note played by a full string and that played by one of exactly half the length. Thus the ratio of the length of string that produces a given note, to the length that produces its octave, is 2/1. This is true independently of the pitch of the original note. Other whole number ratios produce harmonious intervals as well. The main ones are the *fourth*, a ratio of 4/3 , and the *fifth*, a ratio of 3/2. Starting at a base note of C these are

C	D	E	F	G	A	B	C
base			fourth	fifth			octave
1	2	3	4	5	6	7	8

and the numbers underneath show where the names came from. Other intervals are formed by combining these building-blocks.

You can find these ratios on a guitar. Place your left forefinger *very* lightly on the string, and move it slowly along while plucking the string with the right hand. Do *not* depress the string so that it hits any frets. In some positions you'll hear a much louder note. The easiest to find is the octave: place your finger at the middle of the string. The other two places are one third and one quarter along the string.

All guitarists recognise the basic intervals octave, fourth, and fifth. In combination with the fundamental they form the common major chord. A standard 12-bar blues, in the key of C, employs the chord sequence

C/// C/// C/// C/// F/// F/// C/// C/// G/// F/// C/// G///

or a near variant (often with seventh chords instead of major ones in the fourth and final bars).

It is thought that, in order to create a harmonious scale, the Pythagoreans began at a base note and ascended in fifths. This yields a series of notes played by strings whose lengths have the ratios

$$1 \quad \left(\frac{3}{2}\right) \quad \left(\frac{3}{2}\right)^2 \quad \left(\frac{3}{2}\right)^3 \quad \left(\frac{3}{2}\right)^4 \quad \left(\frac{3}{2}\right)^5$$

or

$$1 \quad \frac{3}{2} \quad \frac{9}{4} \quad \frac{27}{8} \quad \frac{81}{16} \quad \frac{243}{32}\,.$$

Most of these notes lie outside a single octave, that is, the ratios are greater than 2/1. But we can descend from them in octaves (dividing successively by 2) until the ratios lie between 1/1 and 2/1. Then we rearrange the ratios in numerical order, to get

$$1 \quad \frac{9}{8} \quad \frac{81}{64} \quad \frac{3}{2} \quad \frac{27}{16} \quad \frac{243}{128}\,.$$

On a piano, these correspond approximately to the notes
CDEGAB.
As the notation suggests, something is missing! The gap between 81/64 and 3/2 sounds 'bigger' than the others. We can plug the gap neatly by adding in the fourth, a ratio of 4/3, which is F on the piano. In fact, we could have incorporated it from the start if we had *descended* from the base note by a fifth, adding the ratio 2/3 to the front of the sequence, and then ascended by an octave to get 2×(2/3)= (4/3)

The resulting scale corresponds approximately to the white notes on the piano, and is shown in **(Fig.1)**.

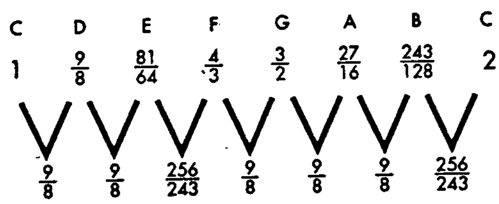

Fig.1

The last line shows the intervals between successive notes, also expressed as ratios. There are exactly two different ratios: the *tone* 9/8 and the *semitone* 256/243. An interval of two semitones is $(256/243)^2$, or 65536/59049, which is approximately 1.11. A tone is a ratio of 9/8 = 1.125. These are not quite the same, but nevertheless two semitones pretty much make a tone. Thus there are gaps in the scale: each tone must be divided up into two intervals, each close to a semitone.

There are various schemes for doing this. The *chromatic scale* starts with the fractions $(\frac{3}{2})^n$ for n = -6, -5, ... , 5, 6. It reduces them to the same octave by repeatedly multiplying or dividing by two, and then places them in order. The result is shown in **(Fig.2)**. Each sharp bears a ratio 2187/2048 to the note below it, and from which it takes its name; each flat bears a ratio 2048/2187 to the note above. There's a glitch in the middle: two notes, F$^\#$ and G$^\flat$, are trying to occupy the same slot, but differ very slightly from each other.

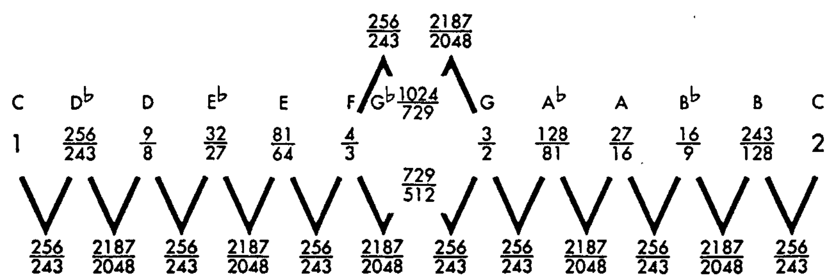

Fig.2.

There are many other schemes, also leading to distinctions between sharps and flats, but they all involve a 12-note scale that is very close to that formed by the white *and* black notes of the piano.

The reason for the glitch in the chromatic scale, and the reason that there are many different schemes for building scales, is that no 'perfect' 12-note scale can be based on the Pythagorean ratios of 3/2 and 4/3. By a perfect scale I mean one where the ratios are all the same, so we get

$$1 \quad r \quad r^2 \quad r^3 \quad r^4 \quad \dots \quad r^{12} = 2$$

for a fixed number r. The Pythagorean ratios involve only the primes 2 and 3: every ratio is of the form $2^a 3^b$ for various integers a and b. For instance $243/128 = 2^{-7}3^5$. Suppose that $r = 2^a 3^b$ and $r^{12} = 2$. Then $2^{12a}3^{12b} = 2$, so $2^{12a-1} = 3^{-12b}$. But an integer power of 2 cannot equal an integer power of 3, by uniqueness of prime factorization. Similar arguments show that *no* fixed integer ratio can work.

This mathematical fact puts paid to any musical scale based on Pythagorean principles of the harmony of whole numbers; but it doesn't mean we can't find a suitable number r. The equation $r^{12} = 2$ has a unique positive solution, namely

$$r = \sqrt[12]{2} = 1.059463094\dots.$$

The resulting scale is said to be *equitempered.*

If you start playing a Pythagorean scale somewhere in the middle — a change of *key* — then the sequence of intervals changes slightly. Equitempered scales don't have this problem, so they are useful if you want to play the same instrument in different keys. Musical instruments that must play fixed intervals, such as pianos and guitars, generally use the equitempered scale. The Pythagorean semitone interval is $256/243 = 1.05349\dots$, which is close to $\sqrt[12]{2}$, so the name 'semitone' is used for the basic interval of the equitempered scale.

How does this lead to the positions of the frets on a guitar? Think about the first fret along, corresponding to an increase in pitch of one semitone. The length of string that is allowed to vibrate has to be $1/r$ times the length of the complete string. So the distance to the first fret is $1 - 1/r$ times the length of the complete string. To get the next distance, you just observe that everything has shrunk by a factor of r, so the spaces between successive frets are in the proportions

$$1 \quad 1/r \quad 1/r^2 \quad 1/r^3$$

and so on. Now r is bigger than 1, so $1/r$ is less than 1, and that means that the distances between successive frets are *smaller* (**Fig.3**).

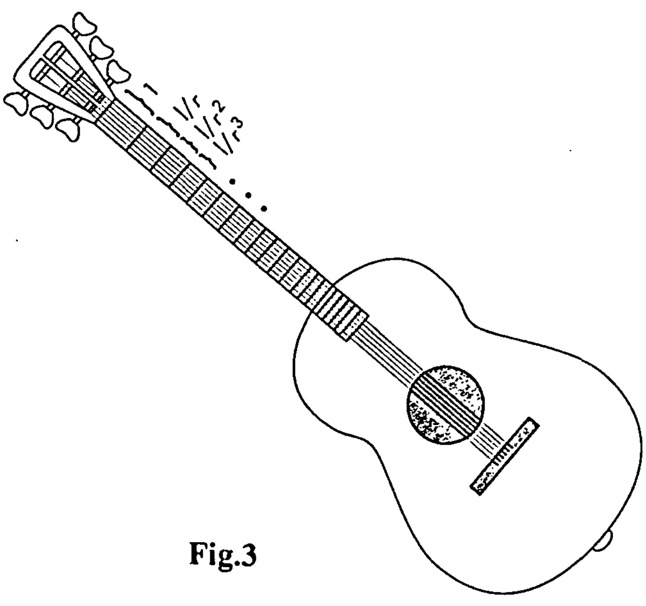

Fig.3

When the Greeks were faced with numbers such as $^{12}\sqrt{2}$ that cannot be written as exact fractions — which they called *irrational* numbers — they usually resorted to geometry. According to tradition, Greek geometry placed considerable emphasis on those lengths that can be constructed using only a ruler and a pair of compasses. For example, squares and square roots can be so constructed. However, it can be proved that there is no ruler-and-compass construction for $^{12}\sqrt{2}$.

The equitempered scale is a compromise, an approximation. The true fourth sounds more harmonious than the equitempered fourth, and singers find it more natural. Since the equitempered scale is a compromise, we may ask whether there is some *approximate* geometric construction that tells you where to put the frets on a guitar. Not only is there an approximate construction, but it has a very curious history. The story illustrates the deep elegance of mathematics, but it is also a humbling tale: an outstanding triumph of a practical man nullified by a professional mathematican's carelessness.

In the 16th and 17th centuries, finding geometrical methods for placing frets upon musical instruments - lute and viol rather than guitar - was a serious practical question. In 1581 Vincenzo Galilei, the father of the great Galileo Galilei, advocated the approximation

$$18/17 = 1.05882\ldots .$$

This led to a perfectly practical method, in common use for several centuries. In 1636 Marin Mersenne, a monk better known for his prime numbers of the form 2^p-1, approximated an interval of four semitones by the ratio $2/(3-\sqrt{2})$. Taking square roots twice, he could then obtain a better approximation to the interval for one semitone:

$$\sqrt{\sqrt{(2/(3-\sqrt{2}))}} = 1.05973\ldots$$

which is certainly close enough for practical purposes. The formula involves only square roots, and thus can be constructed geometrically as in Fig.6. However, it is difficult to implement this construction in practice, because errors tend to build up. Something more accurate than Galilei's approximation, but easier to use than Mersenne's, was needed.

In 1743 Daniel Strähle, a craftsman with no mathematical training, published an article in the *Proceedings of the Swedish Academy* presenting a simple and practical construction (**Fig.4**). Let QR be 12 units long, divided into 12 equal intervals of length 1. Find O such that OQ = OR = 24. Join O to the equally spaced points along QR. Let P lie on OQ with PQ 7 units long. Draw RP and extend it to M so that PM = RP. If RM is the fundamental pitch and PM its octave, then the points of intersection of RP with the 11 successive rays from O are successive semitones within the octave, that is, the positions of the 11 frets between R and M.

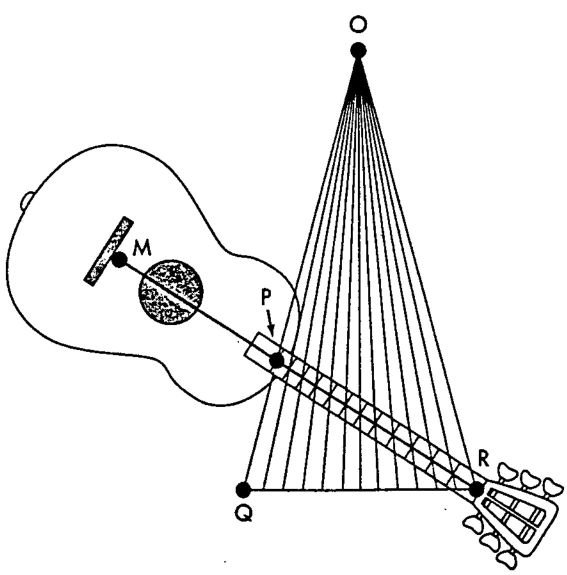

Fig.4

You might like to try it out, and compare with measurements from an actual instrument. But how accurate is it? The famous Swedish geometer and economist Jacob Faggot performed a trigonometric calculation to find out, and appended it to Strähle's article, concluding that the maxmum error is 1.7%. This is about five times more than a musician would consider acceptable.

Faggot was a founder member of the Swedish Academy, served for three years as its secretary, and published eighteen articles in its *Proceedings*. In 1776 he was ranked as number four in the Academy: Carl Linnaeus, the botanist who set up the basic principles for classifying animals and plants into families and genera, was just ahead of him in second place. So when Faggot declared that Strähle's method was inaccurate, that was that. For example, F.W.Marpurg's *Treatise on Musical Temperament* of 1776 lists Faggot's conclusion without describing Strähle's method.

It was not until 1957 that J.M.Barbour of Michigan State University discovered that Faggot had made a mistake.

Faggot began by finding the base angle ∠OPQ of the main triangle: it is 75°31'. From this he could find the length RP and the angle ∠PRQ. Each of the eleven angles formed at the top of the main triangle by the rays from the base could also be calculated without difficulty: it was then simple enough to find the lengths cut off along the line RPM.

However, Faggot had computed ∠PRQ as 40°14', when in fact it is 33°32'. This error, as Barbour puts it, 'was fatal, since ∠PRQ was used in the solution of each of the other triangles, and exerted its baleful influence impartially upon them all.' The maximum error reduces from 1.7% to 0.15%, which is perfectly acceptable. Thus far the story puts mathematicians, if not mathematics itself, in a bad light. If only Faggot had bothered to *measure* ∠PRQ. But Barbour went further, asking *why* Strähle's method is so accurate. He found a beautiful illustration of mathematics's ability to lay bare the reasons behind apparent coincidences. There is no suggestion that Strähle himself adopted a similar line of reasoning: as far as anyone knows his method was based upon the intuition of the craftsman rather than any specific mathematical principles.

The spacing of the nth fret along the line MPR can be represented on a graph **(Fig.5)**. We take the x-axis of the graph to be the line QR in Fig.10, with Q at the origin and R at 1. We move MPR so that it forms the y-axis of the graph, with M at the origin, P at 1, and R at 2. The successive frets are placed along the y-axis at the points 1, r, r^2,..., r^{11}, $r^{12} = 2$. (Note that this differs from the ratios $1/r$, $1/r^2$, ... mentioned above, because we are working from the opposite end of the string.)

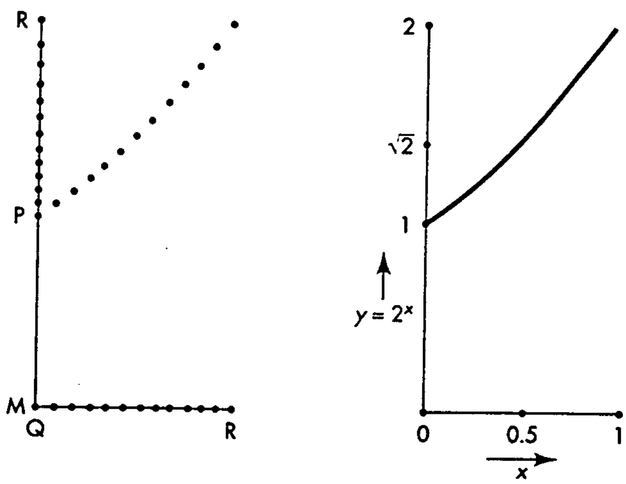

Fig.5

A mathematician would call Strähle's construction a *projection* with *centre* O from a set of equally spaced points along QR to the desired points along MPR. It can be shown that such a projection always has the algebraic form

$$y = (ax+b)/(cx+d) \tag{1}$$

where a,b,c,d are constants. This is called a *fractional linear function.*

For Strähle's method, you can check that the constants in (1) are $a = 10$, $b = 24$, $c = -7$, $d = 24$, so the projection takes a given point x on QR to the point $y = (10x+24)/(-7x+24)$ on MPR. I'll call this formula *Strähle's function.* Strähle didn't derive it: it's just an algebraic version of his geometric construction. However, it is the key to the problem.

If the construction were exact, we would have $y = 2^x$. Then the thirteen equally spaced points $x = n/12$ on QR, where $n = 0,1,2,...,12$, would be transformed to the points $2^{n/12} = (2^{1/12})^n = r^n$ on MPR, as desired for exact equal temperament. But it's *not* exact, even though Barbour's calculations show that it's very accurate. Why? The clue is to find the *best possible approximation* to 2^x, valid in the range $0 \le x \le 1$, and of the form $(ax+b)/(cx+d)$.

One way to do this is to require the two expressions to agree when $x = 0, \frac{1}{2}$, and 1. That gives three equations to solve for a, b, c, d; namely

$$b/d = 1 \tag{2}$$

$$(\tfrac{1}{2}a+b)/(\tfrac{1}{2}c+d) = (\sqrt[12]{2})^{12/2} = (\sqrt[12]{2})^6 = \sqrt{2} \tag{3}$$

$$(a+b)/(c+d) = 2. \tag{4}$$

At first sight we seem to need one more equation to find four unknowns, but really we only need the *ratios* b/a, c/a, and d/a, so three equations are enough. You should really try to solve the equations yourself at this juncture; but here's one method. We may fix the value of d to be anything nonzero, and we decide to set $d = \sqrt{2}$. Then (2) implies $b = \sqrt{2}$ as well. Equation (3) becomes

$$(\tfrac{1}{2}a+\sqrt{2}) = \sqrt{2}(\tfrac{1}{2}c+\sqrt{2})$$

or

$$a + 2\sqrt{2} = c\sqrt{2} + 4 \tag{5}$$

and (4) becomes

$$a + \sqrt{2} = 2(c + \sqrt{2})$$

or

$$a = 2c + \sqrt{2}. \tag{6}$$

Eliminating a from (5,6) we get

$$2c+3\sqrt{2} = c\sqrt{2} + 4$$

so

$$c(2-\sqrt{2}) = 4-3\sqrt{2},$$

and multiplying by $2+\sqrt{2}$ we get

$$2c = (4-3\sqrt{2})(2+\sqrt{2}) = 8-6\sqrt{2}+4\sqrt{2}-6 = 2-2\sqrt{2},$$

whence

$$c = 1-\sqrt{2}.$$

Finally we solve for a from (6), to get $a = 2 - \sqrt{2}$. Thus this approach leads to the values

$$a = 2 - \sqrt{2}$$
$$b = \sqrt{2}$$
$$c = 1 - \sqrt{2}$$
$$d = \sqrt{2}$$

and the best possible approximation (in our chosen sense) to 2^x by a fractional linear function takes the form

$$y = \frac{(2 - \sqrt{2})x + \sqrt{2}}{(1 - \sqrt{2})x + \sqrt{2}}. \tag{7}$$

That doesn't look much like Strähle's function, but now comes a final bit of nifty footwork. Barbour estimated the *error* in terms of the approximation 58/41 to $\sqrt{2}$, and derived Strähle's formula that way. Isaac Schoenberg did the same in 1982. If you just substitute 58/41 for $\sqrt{2}$ in (7) then you get $(24x+58)/(-17x+58)$, which is different from Strähle's function.

Nevertheless, the most natural thing to do is change $\sqrt{2}$ to some approximation — but not 58/41. Here's how. There is a series of rational numbers that approximate $\sqrt{2}$. One way to get them is to start from the equation $p/q = \sqrt{2}$ and square to get $p^2=2q^2$. Because $\sqrt{2}$ is irrational, you can't find integers p and q that satisfy this equation (or, more accurately, because you can't find integers p and q that satisfy this equation, $\sqrt{2}$ must be irrational). But you can come close by looking for integers p and q such that p^2 is close to $2q^2$. The best approximations are those for which the eror is smallest; that is, solutions of the equation $p^2 = 2q^2 \pm 1$. For example, $3^2 = 2.2^2+1$, and $3/2 = 1.5$ is moderately close to $\sqrt{2}$. The next case is $7^2 = 2.5^2-1$, leading to $7/5 = 1.4$, which is closer. Next comes $17^2 = 2.12^2+1$, yielding the approximation $17/12 = 1.4166...$, closer still. You can go on forever. To see how, consider the *continued fraction* for $\sqrt{2}$.

A *continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}}$$

which we abbreviate to $[a_0;a_1,a_2,a_3,\dots]$.

The *continued fraction* for $\sqrt{2}$ is obtained as follows. Start with the identity

$$\sqrt{2} = 1 + \frac{1}{1+\sqrt{2}}$$

and then substitute the right-hand side into itself in place of $\sqrt{2}$ to get

$$\sqrt{2} = 1 + \cfrac{1}{1 + 1 + \cfrac{1}{1+\sqrt{2}}}$$

$$=1 + \cfrac{1}{2 + \cfrac{1}{1+\sqrt{2}}}.$$

Repeating the process, we see that (in standard notation for continued fractions)

$\sqrt{2} = [1;2,2,2,2,\dots.]$.

If we truncate the continued fraction at some finite position, we get a rational approximation to $\sqrt{2}$. The theory of continued fractions tells us that this must be the best possible rational approximation (with a given size of denominator), and not surprisingly we get a rational p/q with $p^2 = 2q^2 \pm 1$. For example,

$[1;2] = 3/2$
$[1;2,2] = 7/5$
$[1;2,2,2] = 17/12$
$[1;2,2,2,2] = 41/29$

and so on. We recognise the first three approximations; and for the fourth we find that

$41^2 = 2.29^2 -1$.

Indeed, if we write

$[1;2,\dots (n \text{ copies}) \dots 2] = p_n/q_n$

then

$$p_n/q_n = 1 + \frac{1}{1 + [1;2,\ldots\text{ (n-1 copies) }\ldots 2]}$$

$$= 1 + \frac{1}{1 + p_{n-1}/q_{n-1}}$$

$$= \frac{2q_{n-1} + p_{n-1}}{q_{n-1} + p_{n-1}} .$$

Comparing numerators and denominators we obtain a pair of recurrence relations

$$p_n = 2q_{n-1} + p_{n-1}$$
$$q_n = q_{n-1} + p_{n-1}.$$

For example, from $p_3 = 17$, $q_3 = 12$ we generate

$$p_4 = 2.12+17 = 41$$
$$q_4 = 12+17 = 29.$$

Continuing this process we get a table of approximations:

n	p_n	q_n
1	3	2
2	7	5
3	17	12
4	41	29
5	99	70
6	239	169
7	577	408
8	1393	985
9	3363	2378
10	8119	5741

each successive q_n is the sum of the two numbers in the row above; each p_n is twice the second plus the first number in the row above. So we have a quick and efficient way to generate rational approximations to $\sqrt{2}$, and incidentally we have proved that the Diophantine equation $p^2 = 2q^2 \pm 1$ has infinitely many solutions. Pursuing these ideas leads to a beautiful theory of the so-called *Pell equation*

$$p^2 = kq^2 \pm 1.$$

In fact it was Lord William Brouncker, and not John Pell, who developed the theory: the ideas were erroneously attributed to Pell by Leonhard Euler.

At any rate, we have lots of approximations to $\sqrt{2}$, among them being 17/12. Now back to Strähle's equation. Divide the numerator and denominator of formula (1) by 2 and rewrite it as the equivalent formula

$$\frac{x + \frac{1}{\sqrt{2}}(1-x)}{\frac{x}{2} + \frac{1}{\sqrt{2}}(1-x)} .$$

Then replace $\sqrt{2}$ by the approximation 17/12, so that $1/\sqrt{2}$ becomes 12/17. This gives

$$\frac{x + \frac{12}{17}(1-x)}{\frac{x}{2} + \frac{12}{17}(1-x)}$$

This simplifies to give

$$\frac{10x + 24}{-7x + 24}$$

which is *precisely* Strähle's formula!

So Strähle's construction is very accurate because it effectively combines *two* good approximations:

- The best fractional linear approximation to 2^x is formula (7) above.
- Strähle's function is obtained from formula (7) by replacing $\sqrt{2}$ by the excellent approximation 17/12. (In fact David Fowler has pointed out to me that while 12/17 is not a convergent of the continued fraction for $\sqrt{2}$, it is a so-called *intermediate convergent.*)

Thanks to the mathematico-historical detective work of Barbour, we now know not only that Strähle's method is extremely accurate: we also have a very good idea of *why* it's so accurate. It's related to basic ideas in approximation theory and in number theory. This leaves just one question unanswered — and, barring a miracle or time travel, unanswerable. *How on earth did Strähle think of his construction to begin with?*

FURTHER READING

J.M.Barbour, A geometrical approximation to the roots of numbers, *American Mathematical Monthly* **64** (1957) 1-9.

Jacob Faggot, Trigonometrisk Uträkning, på den nya temperaturen för theonernes stämming å claveret, *Proceedings of the Swedish Academy* 1743 IV 286-291.

David Fowler, *The Mathematics of Plato's Academy: a New Reconstruction*, Clarendon Press, Oxford 1987.

Ottó Károlyi, *Introducing Music*, Penguin, Harmondsworth 1965.

Isaac J. Schoenberg, On the location of the frets on a guitar, *American Mathematical Monthly* **83** (1976) 550-552.

Isaac J. Schoenberg, *Mathematical Time Exposures*, Mathematical Association of America, 1982.

Ian Stewart, *Another Fine Math You've Got Me Into...*, W.H.Freeman, New York 1992.

Daniel P. Strähle, Nytt påfund, til at finna temperaturen i stämningen för thonerne på claveret ock dylika instrumenter, *Proceedings of the Swedish Academy* 1743 IV 281-286.

GRESHAM COLLEGE

Policy & Objectives

An independently funded educational institution, Gresham College exists

- to continue the free public lectures which have been given for 400 years, and to reinterpret the 'new learning' of Sir Thomas Gresham's day in contemporary terms;

- to engage in study, teaching and research, particularly in those disciplines represented by the Gresham Professors;

- to foster academic consideration of contemporary problems;

- to challenge those who live or work in the City of London to engage in intellectual debate on those subjects in which the City has a proper concern; and to provide a window on the City for learned societies, both national and international.

Founded 1597

Gresham College, Barnard's Inn Hall, Holborn, London EC1N 2HH
Tel: 020 7831 0575 Fax: 020 7831 5208
e-mail: enquiries@gresham.ac.uk
Web site : www.gresham.ac.uk