Mathematical theory of scales

Joshua Ruiter

October 6, 2020

Abstract

Musical scales are fundamental for the theorical study of music, and also practically important for musicians to understand how to play music that sounds in tune. Despite the fact that most modern musicians are only aware of one tuning system (equal temperament), there have been many historic solutions to the problem of scale construction, each with good motivations. We model such scales mathematically, and describe the fundamental limitations of scale construction.

Contents

1 Scales

1.1 Frequencies and intervals

Definition 1.1. A **pitch** or **fequency** is a positive real number. We think of it in units of vibrations per second, perhaps vibrations of a vibrating string on a musical instrument. So the set of pitches is $\mathbb{R}_{>0}$.

Example 1.2. The standard tuning note for orchestras (professional and amateur alike) is 440 Hz. This is the A above middle C. That is to say, when the concertmaster stands up to tune the orchestra, and they play their open A string, that string is vibrating 440 times per second. Alternately, when the oboe plays a note to give everyone a reference pitch, the column of air inside the oboe is vibrating 440 times per second.

Larger frequency numbers correspond to "higher" pitches. So for example, 880 Hz is a "higher" pitch than the tuning note A. To be more precise, it is one octave higher. On a piano, "higher" pitch means farther to the right of the keyboard. On a string instrument (e.g. violin, cello, guitar), higher pitch means shortening the string more (using a finger to stop it at a point and shorten the length which can vibrate freely).

Definition 1.3. An **interval** or **frequency ratio** is a ratio of two frequencies. So if a, b are two frequencies, we obtain an interval $\frac{a}{b}$. We think of this as the "distance" between the pitches a and b. So an interval is also a positive real number, so the set of intervals is also $\mathbb{R}_{>0}$.

If we have two intervals, we can combine them to get a new interval. We start at some pitch x, go up by an interval y, to get a new pitch xy. Then we go up by a new interval z, to get a new pitch xyz. The interval from x to xyz is $\frac{xyz}{x} = yz$. That is to say, combining intervals musically corresponds mathematically to multiplication in $\mathbb{R}_{>0}$.

$$\begin{tikzcd} x \arrow[r, dash, "y"] \arrow[rr, dash, bend right, "yz"'] & xy \arrow[r, dash, "z"] & xyz \end{tikzcd}$$

Assumption 1.4. Pitch is only meaningful contextually. A single frequency is never "in tune" or "out of tune;" only an interval can be "in tune" or "out of tune."

This assumption is based on how our ears perceive pitch. Perhaps there are some cultural assumptions behind this, and perhaps someone with perfect pitch would have a different experience, but we're going to stick with this assumption. But this leads us to a nice compact definition of what a scale is.

Definition 1.5. A **scale** is a set of intervals containing 1. That is, a scale is just a subset $S \subset \mathbb{R}_{>0}$ containing the identity.

Once we have picked a scale, we can say that anything a musician plays is either in tune or out of tune, with respect to that scale. They play some series of notes, which basically means they play some series of pitches. We then compare any two pitches by taking the frequency ratio, finding the interval between them. If that interval is in our scale, it's in tune. If any interval between any two notes they played is not in our scale, they played out of tune. The insistence that a scale contains 1 is just saying that any note is in tune with itself. If that wasn't the case, then nothing would ever be in tune, so we need to assume this to have anything meaningful.

Definition 1.6. The interval 1 is called **unison**.

So we're assuming that playing in unison is always in tune.

Example 1.7. The set $\left\{1, \frac{3}{2}\right\}$ is a finite scale. This would only allow perfect fifths as "in tune" intervals.

Example 1.8. An extreme example is $S = \mathbb{R}_{>0}$. This scale says that everything goes. Every possible interval is in tune, so it's impossible to play out of tune, with respect to this scale. Obviously this has some practical deficiencies, since clearly there are some combinations of pitches which are not so pleasant to hear.

Definition 1.9. The **Pythagorean scale** is the set of ratios

$$P = \left\{ \left(\frac{3}{2}\right)^n : n = 0, 1, \ldots, 11 \right\}$$

Definition 1.10. Let $n \in \mathbb{Z}$. The set $S_n = \left\{2^{m/n} : m \in \mathbb{Z}\right\}$ is a scale. This is called **n-tone equal temperament**. In particular, the case $n = 12$, the 12-tone equal temperament scale, has been pretty much the universal standard in both popular and classical music for the past century.

1.2 Octave equivalence

In music, there is a very common phenomenon, which most people know about even if they aren't trained musicians. This phenomenon is called octave equivalence.

Definition 1.11. An **octave** is the interval 2. Two pitches/frequencies/notes are **an octave apart** if the interval between them is an octave.

Fact 1.12. Two notes which are an octave apart sound "the same."

This is not to say that we cannot distinguish between them. It is not hard to tell an octave and a unison apart when hearing them, even for the untrained ear. They are clearly different in some way. However, notes which are an octave apart play "the same" role structurally in music.

For example, if you change the only bass line of your favorite pop song by replacing it with something an octave higher or an octave lower, it will still sound like basically the same song. As a common example, in a male/female vocal duet, frequently the two voices will sing "the same thing" one octave apart, at the same time. To our ears, it sounds like they are singing "the same" melody, even though it differs by an octave.

What are the mathematical consequences of octave equivalence? Mathematically, we are saying that an octave and a unison are the same interval. We are taking the set $\mathbb{R}_{>0}$

and saying that $x = 2x$. We know how to represent this, we just take the quotient group $G = \mathbb{R}_{>0}/2\mathbb{R}_{>0}$. That is, we quotient by the subgroup $\{2^n : n \in \mathbb{Z}\}$. We will now stop thinking of intervals as elements of $\mathbb{R}_{>0}$, we will think of intervals as elements of the quotient G. Similarly, we will now think of scales a subsets of G containing the identity.

Example 1.13. Because of octave equivalence, $\frac{3}{2}$ (perfect fifth) and $\frac{3}{4}$ (perfect fourth) are the same interval.

Remark 1.14. The functions e^x and log are group isomorphisms

$$(\mathbb{R}, +) \underset{\log}{\overset{\exp}{\rightleftarrows}} (\mathbb{R}_{>0}, \times)$$

So the quotients are also isomorphic.

$$(\mathbb{R}/\log 2, +) \underset{\log}{\overset{\exp}{\rightleftarrows}} (\mathbb{R}_{>0}/2, \times)$$

With some scaling, $\mathbb{R}/\log 2 \cong \mathbb{R}/\mathbb{Z}$. As yet another way to think about it, $\mathbb{R}/\mathbb{Z}$ is isomorphic to the unit circle in the complex plane.

$$\mathbb{R}/\mathbb{Z} \xrightarrow{\cong} S^1 \qquad x \mapsto e^{2\pi i x}$$

1.3 Overtones

Fact 1.15. Suppose some pitch x is played on an instrument. Call x the **fundamental frequency**. When the instrument is a string or wind instrument, other pitches also sound at the same time, though less clearly. These pitches are called **overtones**, and they only occur at positive integer multiples of x. Their strength decays rapidly as they get farther from x, so human ears only meaningfully detect the first few overtones, say $< 10x$.

Note that this is not some property of the universe or of sound itself. This is really a property of how strings and columns of air vibrate and how instruments are constructed. There are instruments which do not have overtones in this way, for example percussion instruments like drums do not create overtones.

Corollary 1.16. *Small integer multiples of a pitch x should be considered "in tune" with x.*

In terms of intervals, this means the intervals $3, 4, 5, 6, \ldots, n$ should be in a "good" scale. Because of octave equivalence, we can normalize these so that they occur between 1 and 2, that is, we can choose coset representatives in the interval $[1, 2)$.

$$3 = \frac{3}{2} \qquad 4 = 1 \qquad 5 = \frac{5}{4} \qquad 6 = 1 \qquad 7 = \frac{7}{4} \qquad \cdots$$

This gives some explanation for why ratios of small integers show up as common intervals. These have names. In particular $\frac{3}{2}$ is called a **perfect fifth**, and $\frac{5}{4}$ is called a **major third**.

Example 1.17. If you're having a hard time connecting this to the music you know, a perfect fifth is the first interval in the "alphabet song," between "B" and "C." Or maybe you know that tune better as "Twinkle, Twinkle, Little Star."

Remark 1.18. If you have some musical training, what you've been taught to think is a "perfect fifth" or "major third" has more to do with 12-tone equal temperament than these small integer ratios. Why is this? The 12-tone equal tempered scale doesn't contain these exact ratios. In 12-tone equal temperament, we approximate fifths and thirds.

Name	Just interval	12-tone ET approximation
Major third	$5/4 = 1.25$	$2^{4/12} \approx 1.259921$
Perfect fourth	$4/3 = 1.\overline{3}$	$2^{5/12} \approx 1.334839$
Perfect fifth	$3/2 = 1.5$	$2^{7/12} \approx 1.498307$

These are pretty good approximations, but imperfect.

Remark 1.19. The Pythagorean scale is a logical conclusion from, we should just build our scale from small integer ratios. We start with 1, then add $\frac{3}{2}$, then start "stacking perfect fifths" on top of each other. Mathematically, we're just adding higher integer powers of $\frac{3}{2}$. Why stop after 11 then? Well,

$$\left(\frac{3}{2}\right)^{12} \approx 129.746\ldots \qquad 2^7 = 128$$

Musically, this says that 12 perfect fifths is very close to 7 octaves. The fifths are a bit too big, though. But rather than add in this new interval, which would be very very close to unison but not quite, we stop and just say, this is about the same. What are the negative consequences of this? Well, our last fifth is forced to be smaller than we want it to. Instead of that last interval being $\frac{3}{2}$, it's

$$\frac{2^7}{\left(\frac{3}{2}\right)^{11}} \approx 1.47981\ldots$$

We've put off all of the problems till the last possible moment, then squeezed the last "fifth" to fit into some very tight shoes. Even for those not so musically inclined, I think this last fifth is noticeably worse sounding than a true $\frac{3}{2}$ perfect fifth.

But this "bad fifth" causes more problems for the Pythagorean scale than you might think. The real problem comes around when you start transposing. Let's say you decide to base your Pythagorean scale on the note A = 440 Hz, so 440 Hz is 1, you tune a piano just going by fifths and octaves from there. This is not merely theoretical, it's something you could actually do.

Once you have this piano tuned, you try to play something. Let's say you play in a key which is far away from this bad fifth. Mathematically speaking, you pick a base frequency which is many factors of $\frac{3}{2}$ away from the bad one, so somewhere in the middle. Let's say $\left(\frac{3}{2}\right)^5$, five fifths up from A, which would be the key of G#. When you play in this key, things sound pretty good, because that bad fifth from D to A doesn't come up very often in this key.

If you go up one more fifth to D#, things still sound decent. But as you keep changing keys and playing in keys close to the bad fifth, that bad fifth comes up more and more in the music, and it sounds very bad. In the worst possible scenario, if you play in the key of

D major, then the very first fifth up from D to A will be the bad fifth. The major triad D, F#, A will always have this bad fifth, which will sound noticeably out of tune to our ears. And this is just unacceptable, because you can't compositionally avoid this major triad.

So basically, on this Pythagorean piano, you can't really play music composed in keys which are near the bad fifth, without sounding bad. This motivates our next discussion, which captures this problem mathematically, and discuss mathematically how to avoid it.

1.4 Transposition

A slightly more sophisticated musical concept than octave equivalence is transposition. Musical transposition happens all the time in all kinds of music, both classical and pop.

Perhaps another name you've heard of for this is a "key change." Say you're going along in some piece of music, and there's definitely a note which is "home base," it's the 1 in our scale. This is the note that the bass line plays the most, and especially the bass line ends up on this note at the end of almost every musical phrase. Then, something strange happens, and everything shifts a bit, and there's a new "home" note. The music underwent a key change.

In classical music and music theory, this is called **tranposing**. The whole frame of reference for notes has shifted from one reference point to a new one. Mathematically, all this means is that our scale shifts from one set of intervals to a new one which is scaled by the interval between the original "home" note and the new "home" note.

Definition 1.20. Let $S \subset \mathbb{R}_{>0}/2$ be a set of pitches, and let $x \in \mathbb{R}_{>0}/2$ be an interval. The **transposition** of S by x is the set $xS = \{xs : s \in S\}$.

Fact 1.21. A piece of music (a.k.a. a set of pitches) sounds in tune if and only if any transposition of it sounds in tune.

Definition 1.22. A scale S is **transposition invariant** if $S = xS$ for any $x \in S$.

What does it mean musically for a scale to be transposition invariant? It means that you can transpose in the middle of a piece of music by any interval within your scale.

Example 1.23. The scale $S = \left\{1, \frac{3}{2}\right\}$ is NOT transposition invariant. Why is this? Because we can't transpose by $\frac{3}{2}$. If we do we get

$$\frac{3}{2}S = \left\{\frac{3}{2}, \frac{9}{4} = \frac{9}{8}\right\} \neq S$$

Example 1.24. The Pythagorean scale $P = \left\{\left(\frac{3}{2}\right)^n : n = 0, \ldots, 11\right\}$ is not transposition invariant. If we multiply all of S by $\frac{3}{2}$, we get a new interval $\left(\frac{3}{2}\right)^{12}$ which is very close to $2^7 = 1$, but not equal even under octave equivalence.

Example 1.25. The scale $S_2 = \left\{\left(\frac{3}{2}\right)^n : n \in \mathbb{Z}\right\}$ is scale invariant. It contains $1 = \left(\frac{3}{2}\right)^0$, and multiplying by any $\left(\frac{3}{2}\right)^n$ does not change S.

Group theoretically, this is the cyclic subgroup generated by $\frac{3}{2}$. It is infinite. Because $2^k = 1$, we can also write it as

$$S_2 = \left\{3^n 2^k : n \in \mathbb{Z}\right\}$$

This scale is not quite as bad as using all of $\mathbb{R}_{>0}/2$ as a scale, but it's still pretty bad, because S_2 "fills up" the interval $[1, 2)$ pretty quickly. Practically speaking, this means that this scale also says that almost anything goes, since we can approximate any real number in $[1, 2)$

Example 1.26. Equal tempered scales are transposition invariant. $S = \left\{2^{m/n} : k \in \mathbb{Z}\right\}$ is scale invariant. Group theoretically, this is the cyclic subgroup generated by $2^{1/n}$, and it is finite of order n.

Lemma 1.27. *A scale is transposition invariant if and only if it is a subgroup of G.*

Proof. Let $S \subset G$ be a scale, and suppose it is a subgroup. Then $S = xS$ for any $x \in S$ clearly so S is transposition invariant.

Conversely, suppose S is translation invariant. We know $1 \in S$, and we know $S = xS$ for any $x \in S$, so S is closed under multiplication. We just need to verify S is also closed under inversion. Let $x \in S$. Since $S = xS$, there exists $y \in S$ so that $xy = 1$, so $y = x^{-1} \in S$. $\square$

1.5 Subgroups of S^1

Recall $G = \mathbb{R}_{>0}/2$.

Proposition 1.28. *The torsion subgroup of G is*

$$T = \left\{2^{n/m} : n \in \mathbb{Z}_{\geq 0}, m \in \mathbb{Z}_{>0}, \gcd(n, m) = 1\right\}$$

In general, the order of $2^{n/m}$ is $\frac{m}{\gcd(n,m)}$, so assuming $\gcd(n, m) = 1$, the order is just m.

Proof. All such elements are torsion, because $(2^{n/m})^m = 2^n = 1$. If an abritrary element $x \in G$ is torsion, then $x^n = 2^m$ for some $n, m \in \mathbb{Z}_{\geq 0}$. If $m = 0$, then x is a root of unity, but the only root of unity in $\mathbb{R}_{>0}$ is 1, so $x = 1$. If $m \neq 0$, then $x = 2^{m/n}$. We may assume $\gcd(n, m) = 1$ since otherwise we just reduce n/m to n'/m' with $\gcd(n', m') = 1$, and $2^{n/m} = 2^{n'/m'}$.

Now for the statement regarding order. Clearly, $\left(2^{n/m}\right)^{m/\gcd(n,m)} = 2^{n/\gcd(n,m)} = 1$. Since $\gcd(n, m)$ is by definition the largest integer dividing both n and m, we cannot decrease $\frac{m}{\gcd(n,m)}$ while having $\frac{n}{\gcd(n,m)}$ be an integer. Thus $\frac{m}{\gcd(n,m)}$ is the order. $\square$

Proposition 1.29. *Let $n \in \mathbb{Z}_{>0}$. The only cyclic subgroup of G of order n is $\mathbb{Z}/n\mathbb{Z}$ generated by $2^{1/n}$. That is to say, G has a unique cyclic subgroup of each order.*

Proof. Suppose $H \cong \mathbb{Z}/n\mathbb{Z}$ is a cyclic subgroup of G. H must be generated by an element $x = 2^{k/m}$ of order n, so $m = n$. Any k coprime to n makes x a generator, but all of these generate the same subgroup. $\square$

Remark 1.30. In terms of our interpretation of G as $\mathbb{R}/\mathbb{Z}$, the unique subgroup of order n is generated by $\frac{1}{n}$ (remember $\mathbb{R}/\mathbb{Z}$ is an additive group). In terms of the interpretation of G as the circle, the unique subgroup of order n is the group of nth roots of unity.

Lemma 1.31. *Let $a, b \in \mathbb{Z}$. If $\gcd(a, b) = 1$, then $\gcd(a + b, ab) = 1$.*

Proof. Suppose $\gcd(a, b) = 1$ and $\gcd(a + b, ab) \neq 1$. Then there is a prime $p > 1$ dividing $a + b$ and ab. Since $p|ab$, it must divide at least one of a, b. WLOG assume $p|a$. Since $\gcd(a, b) = 1$, $p \nmid b$. But then it is impossible for p to divide $a + b$, since it divides a but not b. $\square$

Proposition 1.32. *Every finite subgroup of G is cyclic.*

Proof. Let $T \subset G$ be the torsion subgroup, and let $H \subset T$ be a finite subgroup. By the classification of finite abelian groups, H can be written as a direct sum of cyclic groups with orders which are pairwise relatively prime. So we write

$$H \cong \mathbb{Z}/a_1\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/a_i\mathbb{Z}$$

where $a_1, \ldots, a_i$ are all pairwise relatively prime. By induction it suffices to prove H is cyclic (of order relatively prime to $a_3, \ldots, a_i$) in the case $i = 2$.

Suppose H is a direct sum of two cyclic subgroups, $H \cong (\mathbb{Z}/a_1\mathbb{Z}) \oplus (\mathbb{Z}/a_2\mathbb{Z})$. By Proposition 1.29, these two cyclic subgroups are generated by $2^{1/a_1}$ and $2^{1/a_2}$ respectively. Then H contains the element

$$x = 2^{\frac{1}{a_1}} 2^{\frac{1}{a_2}} = 2^{\frac{1}{a_2} + \frac{1}{a_2}} = 2^{\frac{a_1+a_2}{a_1 a_2}}$$

By lemma 1.31, $\gcd(a_1 + a_2, a_1 a_2) = 1$. Thus x generates a cyclic subgroup of H of order $a_1 a_2$, which is the order of H, so $H = \langle x \rangle$. As previously noted, the result now follows by induction. $\square$

Corollary 1.33. *If a scale is transposition invariant and finite, it is an equal tempered scale.*

Remark 1.34. This gives some mathematical backing to the idea of using an equal tempered scale. It says that either you need to give up transposition invariance, or have infinitely many allowable intervals, if you want to do anything else with your scale.

Proposition 1.35. *No finite transposition invariant scale contains any rational interval.*

Proof. We know what all the finite transposition invariant scales look like, they are generated by $2^{1/n}$. But $2^{m/n}$ is never rational unless it is an integer multiple of 2, in which case it is just equal to 1, as an interval. $\square$

2 Equal temperaments compared

If we have decided that we really want tranposition invariance and finiteness, we're still left with the problem of how to choose an equal tempered scale. There are a whole countable infinite of them to choose from, so how can we pick?

We return to the conclusion of the discussion around overtones - if we're going to play this music on string and wind instruments, intervals like $\frac{3}{2}$ and $\frac{5}{4}$ are desirable. We know that no equal temperament system actually includes these intervals, but we can try to get approximations. Let's just focus on $\frac{3}{2}$ and trying to approximate that, since approximating multiple at once seems a lot harder. So we have the question, For which n is there an m so that $2^{m/n}$ is a good approximation for $\frac{3}{2}$?

Let's do some naive arithmetic manipulations first.

$$2^{m/n} \approx \frac{3}{2} \iff \frac{m}{n} \approx \log_2 \frac{3}{2} = \log_2 3 - \log_2 2 = \log_2 3 - 1 \iff \frac{m+n}{n} \approx \log_2 3$$

So really, our question is just,

What are good rational approximations of $\log_2 3$?

Well, what does "good" mean here? Let's make a definition.

Definition 2.1. Let $x \in \mathbb{R}$. A (reduced) rational number $\frac{c}{d}$ is a **best approximation of the first kind** for x if

$$\left|x - \frac{c}{d}\right| < \left|x - \frac{p}{q}\right|$$

for any fraction $\frac{p}{q}$ with $q < d$. In other words,

$$\left|x - \frac{c}{d}\right|$$

is minimal among such expressions involving rational numbers with denominator less than d. That is, $\frac{c}{d}$ is the best rational approximation of x except for fractions with larger (or equal) denominator.

Definition 2.2. Let $x \in \mathbb{R}$. A (reduced) rational number $\frac{c}{d}$ is a **best approximation of the second kind** for x if

$$|dx - c| < |qx - p|$$

for any fraction $\frac{p}{q}$ with $q < d$.

Remark 2.3. The inequality for best approximations of the second kind is strictly stronger. If $q < d$ and $|dx - c| < |qx - p|$, then we can divide the left side by d and the right side by q to get

$$\left|x - \frac{c}{d}\right| < \left|x - \frac{p}{q}\right|$$

It's a bit subtle, but best approximations of the second kind are actually quite a bit more stringent.

2.1 Continued fractions

Definition 2.4. A **continued fraction** is an expression of the form

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}$$

We notate it more succinctly as

$$[a_0; a_1, a_2, a_3, \ldots]$$

Theorem 2.5. *Every real number x can be written as a continued fraction. If x is irrational, then the continued fraction representation is unique.*

Remark 2.6. Rather than give a proof, I'll describe a recursive algorithm for calculating a continued fraction representation. Let $x \in \mathbb{R}$. Recall that the floor of x is written $\lfloor x \rfloor$ and is the largest integer less than or equal to x.

1. Set $x_0 = x$, and set $a_0 = \lfloor x_0 \rfloor = \lfloor x \rfloor$ to be the floor of x. This is the first digit in the continued fraction representation of x.

2. Set $x_{i+1} = \frac{1}{x_i - a_i}$, and set $a_{i+1} = \lfloor x_{i+1} \rfloor$. (If at any point $x_i = a_i$ then the algorithm terminates and x is rational.)

Example 2.7. Let's compute the first few digits of the continued fraction of $\log_2 3$.

$$\begin{aligned} x_0 &= \log_2 3 \approx 1.58\ldots & a_0 &= \lfloor x_0 \rfloor = 1 \\ x_1 &= \frac{1}{x_0 - a_0} = \frac{1}{\log_2 3 - 1} \approx 1.70\ldots & a_1 &= \lfloor x_1 \rfloor = 1 \\ x_2 &= \frac{1}{x_1 - a_1} = \frac{1}{\frac{1}{\log_2 3 - 1} - 1} \approx 1.40\ldots & a_2 &= \lfloor x_2 \rfloor = 1 \end{aligned}$$

Let's stop there and I'll just tell you how things continue.

$$\log_2 3 = [1; 1, 1, 2, 2, 3, 1, 5, 2, 23, 2, 2, 1, 1, 55, \ldots] = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2 + \ldots}}}}$$

Theorem 2.8 (Other interesting properties of continued fractions)**.**

1. *The continued fraction terminates if and only if x is rational.*

2. *Every rational can be represented in exactly two ways, which only differ in the "last two" entries.*
$$[a_0; a_1, \ldots, a_{n-1}, a_k] = [a_0; a_1, \ldots, a_{n-1}, a_k - 1, 1]$$

3. *Suppose x is irrational. The continued fraction for x is eventually repeating if and only if x is a root of a quadratic.*

Definition 2.9. Let $x = [a_0; a_1, a_2, \ldots]$ be a continued fraction. The kth **convergent** of x is the rational number $c_k(x) = [a_0; a_1, \ldots, a_k]$. These are successively better rational approximations for x.

Example 2.10. Let's compute the first few convergents of $\log_2 3$.

$$\begin{aligned}
c_0 &= 1 \\
c_1 &= 1 + \frac{1}{1} = 2 \\
c_2 &= 1 + \frac{1}{1 + \frac{1}{1}} = \frac{3}{2} \\
c_3 &= 1 + \frac{1}{1 + \frac{1}{1+\frac{1}{2}}} = 1 + \frac{1}{1 + \frac{2}{3}} = 1 + \frac{3}{5} = \frac{8}{5} \\
c_4 &= 1 + \frac{1}{1 + \frac{1}{1+\frac{1}{2+\frac{1}{2}}}} = \frac{19}{12} \\
c_5 &= 1 + \frac{1}{1 + \frac{1}{1+\frac{1}{2+\frac{1}{2+\frac{1}{3}}}}} = \frac{65}{41}
\end{aligned}$$

Notice how the even convergents approach x from below and form an increasing sequence, and the odd convergents approach x form above and form a decreasing sequence. This is a general fact about convergents.

Theorem 2.11. *A fraction is a best rational approximation of the second kind for x if and only if it is a convergent of x.*

Example 2.12. Since a best approximation of the second kind is also of the first kind, the theorem tells us that $\frac{65}{41}$ is the best rational approximation to $\log_2 3$ among all fractions with denominator ≤ 41.

2.2 Returning to equal temperaments

Now that we know something about good rational approximations for $\log_2 3$, let's return to why we wanted to know about such things, which had to do with scales. Recall we wanted to find m, n so that

$$2^{m/n} \approx \frac{3}{2}$$

which is equivalent to

$$\frac{m+n}{n} \approx \log_2 3$$

In the expression $2^{m/n}$, n represents the number of even divisions of the octave n-tone equal tempered scale, and m represents the number of those intervals with which we represent our perfect fifth interval $\frac{3}{2}$. From our calculations above, we have the following good approximations for $\log_2 3$.

k	$c_k(\log_2 3) = \frac{n+m}{n}$	n	m
2	$3/2$	2	1
3	$8/5$	5	3
4	$19/12$	12	7
5	$65/41$	41	24

What are the consequences of this table for music theory? It says that if you want an n-tone equal tempered scale to have a "best approximation" to a perfect fifth in the mathematical sense which I've described, you have to use $n = 2, n = 5, n = 12, n = 41$ or something bigger. Let's think about these.

When $n = 2$, the 2-tone equal tempered scale isn't really what normally think of as a scale, musically speaking. It meets the mathematical definition I gave, but it's a very boring scale. It has only two notes, and that one note is exactly half (multiplicatively) of an octave. It's

$$2^{1/2} \approx 1.415\ldots$$

which is a very poor approximation of a perfect fifth, and would sound very bad to the ears. So it's not a very useful scale.

When $n = 5$, we have the 5-tone equal tempered scale. So there are 5 notes, 5 equal divisions of the octave. Our approximation for the perfect fifth $\frac{3}{2}$ is

$$2^{3/5} \approx 1.5157\ldots$$

which isn't bad. This is actually a somewhat interesting scale. But it still doesn't really have enough different notes for all the things composers want to do.

When $n = 12$, we get the 12-tone equal tempered scale that we're all so used to. The approximation for the perfect fifth $\frac{3}{2}$ is

$$2^{7/12} \approx 1.4983\ldots$$

This interval is oftened called a **tempered fifth**. It is every so slightly smaller than a true fifth. This is a difference that a trained ear can hear, but it's not so bad.

When $n = 41$, suddenly we have 41 notes in our scale, more than three times as many as 12TET. On the other hand, the approximate fifth is very good.

$$2^{24/41} \approx 1.500419\ldots$$

The main problem with this scale is that the basic smallest interval $2^{1/41}$ is just very small, so small that people would have a hard time telling apart $2^{1/41}$ and $2^{2/41}$.

This is the end of the story, for today. We have some sort of mathematical justification for why we use 12TET. It is the best choice of ET which has a good approximation to the perfect fifth, except for ones which have way too many and too small of notes.