Continued Fractions and Hofstadter's INT Function

Standard Continued Fractions

We are so used to seeing numbers written in decimal format that it is easy to forget that this is not the only or even necessarily the best way to write numbers. For a start, there is nothing special about the number 10 and numbers written in octal or hexadecimal look very different.

There is, in fact, a much more natural way to write non-integer numbers. Take the number π for example (= 3.141592654...). To a first approximation, π = 3 (plus a little bit).

Now if you want to do a bit better you have got to concentrate on the little bit which actually equals 0.14159... Of course the little bit is always less than 1 so if we want to express the size of this little bit in terms of an integer it is natural to examine the reciprocal of 0.14159 which is just a little over 7. So now we can say that π is approximately equal to 3 plus the reciprocal of 7. (3 + 1/7 is, of course equal to 22/7 – the well known schoolboy approximation to π.)

But why stop there? The reciprocal of 0.14159... is actually 7.062513... so lets knock off the 7 and examine the reciprocal of 0.062513.... This is 15.99659... We can go on doing this for ever, always knocking off the integer part and taking the reciprocal of the remainder. The sequence which this generates for π is [3; 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, …] and if we want to write down the sequence of successive fractions which are generated they look like this:

$$[3;7] = 3 + \frac{1}{7} = \frac{22}{7} \tag{1}$$

$$[3;7,15] = 3 + \cfrac{1}{7 + \cfrac{1}{15}} = 3 + \frac{15}{106} = \frac{333}{106} \tag{2}$$

$$[3;7,15,1] = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1}}} = 3 + \cfrac{1}{7+\cfrac{1}{16}} = 3 + \frac{16}{113} = \frac{355}{113} \tag{3}$$

all of which are better and better approximations to π. Now the next number in the sequence is 292 which is quite a large number so when we take the reciprocal of this number, adding it in is not going to make a great deal of difference. This is why the fraction 355/113 is such a good approximation. In fact it differs from π only in the seventh decimal place!

Notation

There are many ways of notating continued fractions but according to Wikipedia, the standard way is by listing the successive components in square brackets, separated by commas, except the first pair which are separated by a semi-colon thus: [3; 7, 15, 1, 292, 1 ...]. As we shall see, this notation is not sufficiently flexible for my purposes so I shall introduce a new symbol – the double slash // which means 'divided by everything which follows'. i.e. $x \,//\, y + z$ is equivalent to $x/(y+z)$. Using this useful symbol we can write the complicated fraction in equation (3) much more succinctly as follows:

$$\pi \approx 3 + 1//7 + 1//15 + 1//1 = 355/113 = 3.14159292... \tag{4}$$

The numbers 3, 7, 15, 1 etc. are called coefficients but for reasons which will become clear I shall refer to the 3 as the *zeroth* coefficient, the 7 as the *first* coefficient, the 15 as the *second* and so on.

My notation also allows for the possibility of writing continued fractions with numerators which are not equal to 1. For example it can be shown (amazingly) that

$$\pi = 3 + 1^2//6 + 3^2//6 + 5^2//6 + 7^2//6 + ... \qquad (5)$$

but I shall not be needing this facility as I shall be talking exclusively about continued fractions with unit numerators from now on. For this reason I shall also drop the figure 1 as well and write equation (4) as $\pi \approx 3 + //7 + //15 + //1$.

I must add a word of warning, however. When evaluating continued fractions, you must always work from the *end to the beginning*. 0 + //3 + //2 equals 0 + //3½ = 2/7. It does NOT equal (0 + 1/3) + //2. This would equal 1/3 + 1/2 = 5/6.

Nearest Integer Continued Fractions

If you look back at the way we calculated the standard continued fraction for the number π you may recall that at one stage we got to the number 15.99650... which generated the integer 15 and remainder 0.99759... Now it may have occurred to you that if we want to find the best approximation to π, it would have been better to take the *nearest integer* i.e. 16 as the next term of the continued fraction. Good idea. But if we do this we must accept that the remainder is negative i.e. -0.00341.... The reciprocal of this is -293.255...

What this means is that we can also express π even more accurately as

$$\pi \approx 3 + //7 + //16 - //293 = 3.141592653 \qquad (6)$$

which, while having the same number of terms in it as example (4), is now accurate to nine decimal places! This expansion is known as the *nearest integer continued fraction* or NICF.

You can now see another reason why my notation is better than the standard one – it allows for negative terms as well as positive ones. (It is worth commenting here that, even if there are negative terms in a NICF, the actual coefficients are *always positive*. We never have to write expressions like 3 + //–7 + //15 even though such expressions are not ill-formed.)

Properties of Continued Fractions

The method of generating a continued fraction is closely related to Euclid's method for finding the Greatest Common Denominator (GCD) of two numbers. For example, to find the GCD of 2130 and 678 we proceed as follows:

1. Subtract 678 from 2130 repeatedly until the remainder is too small. This can be done 3 times and the remainder is 96.
2. Subtract 96 from 678 repeatedly until the remainder is too small. This can be done 7 times and the remainder is 6.
3. Subtract 6 from 96 repeatedly until the remainder is too small. This can be done 16 times and the remainder is 0.
4. The GCD of 2130 and 678 is therefore 6

It should not come as much of a surprise to the see the coefficients of the NICF expansion of π appearing (3, 7 and 16) because 2130/678 reduces to 355/113 – the best approximation to π.

Another interesting fact about continued fractions is that some irrational numbers generate repeating sequences of coefficients. For example:

$$\sqrt{2} = 1 + //2 + //2 + //2 + \ldots \qquad (7)$$

It is not difficult to see why. If we write

$$X = 1 + //2 + //2 + //2 + \ldots \qquad (8)$$

then $$X - 1 = //2 + //2 + //2 + \dots \quad (9)$$
so $$X - 1 = //2 + (X - 1) = 1/(2 + X - 1) \quad (10)$$

which is a quadratic equation one of whose roots is $\sqrt{2}$. Numbers with square roots in are called quadratic numbers and all quadratic numbers generate repeating continued fractions and vice versa.

Probably the most interesting property,however, of a continued fraction is that it gives us a way of giving meaning to the idea that some numbers (whether rational or irrational) are *more rational* than others.

Firstly we should note that the continued fraction (whether standard or nearest integer) of a rational number is finite in length. This is because Euclid's algorithm as described above is bound to terminate. (Contrast this with the decimal expansion of 22/7 which is infinite in length). In the case of rational numbers we might claim that the fewer coefficients there are in the (nearest integer) continued fraction of a number, the more rational it is. For example: 6/13 = 0 + //2 + //6 but 5/13 is 0 + //3 – //3 – //2 so on my basis, 6/13 is slightly more rational than 5/13.

This is probably stretching the point a bit but there is no doubt that the coefficients in the NICF tell us a lot more about the internal structure of a number than do the digits in its decimal expansion.

When it comes to irrational numbers the continued fraction is infinitely long but, as we have seen, occasionally you come across a large coefficient such as 293 in the NICF of π. This means that at this point, π is very close to a rational number. We could define the rationality of an irrational number as being related in some way to the *smallness* of its coefficients. This raises an interesting question. What is the least rational irrational number? The answer must be

$$\varphi = 1 + //1 + //1 + //1 + \dots \quad (11)$$

Now it is obvious (especially when written in my notation) that

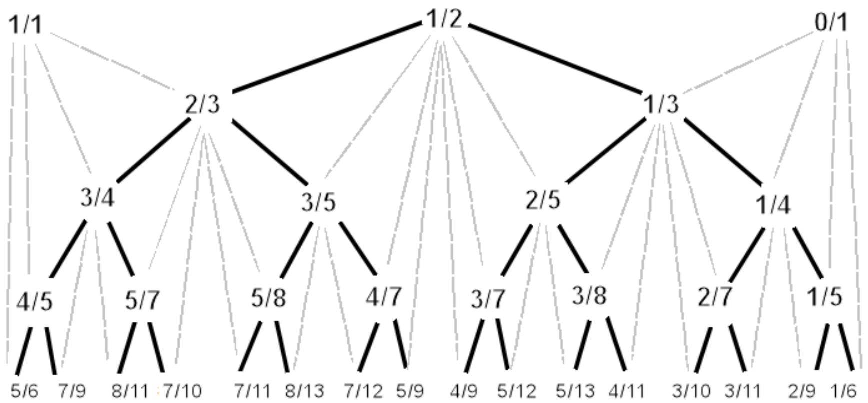
$$\varphi = 1 + //\varphi \quad (12)$$

from which it follows that φ is equal to the golden ratio 1.618...

Properties of Nearest Integer Continued Fractions

Standard continued fractions have been widely studied but Nearest Integer Continued Fractions, in spite of their more fundamental nature, have not received nearly so much attention. In particular I am not aware that anyone has pointed out the close relation which exists between NICF's and the Stern-Brocot Tree.

The Stern-Brocot tree is a way of organising the rational fractions is such a way that each fraction is the Farey sum or mediant of two fractions above it. (The Farey sum of a/c and b/d is $(a + b)/(c + d)$.) It looks like this:



The Stern-Brocot tree

For example, 3/5 is the Farey sum of 1/2 and 2/3; 5/12 is the Farey sum of 3/7 and 2/5. It is a remarkable – and not easily proved – fact that *every rational fraction* appears somewhere in the tree and *always in its lowest possible terms*. (You will not find 3/9 of 21/14 in the tree.)

I wish to propose a different way of constructing the tree from which the the above statements can easily be proved.

The first thing to point out is that all the coefficients in a NICF are greater than or equal to 2. This is because, unlike standard continued fraction, the magnitude of the 'remainder' in a NICF is always <= ½.

Any fraction whose NICF ends in … *k* + //2 can also be written as … (*k*+1) – //2. For example: 2/5 = 0 + //2 + //2 or 0 + //3 – //2 and from each of these starting points it is possible to generate an infinite sequence of unique NICF's by incrementing the last coefficient.

Any fraction whose NICF ends in ± //*k* where k is not equal to 2 can also be written as a NICF of the next higher order as ± //(*k*–1) + //1. For example 1/4 = 0 + //4 can be written as a second order NICF 0 + //3 + //1. Again, this can be used to generate another infinite sequence of NICF's by incrementing the last coefficient.

In fact any fraction, whether or not its last coefficient is 2, can spawn two sequences of fractions one getting smaller and smaller, the other larger and larger.

Here is a table of the first four generations of NICF's generated in this way:

<table>
<tr><td rowspan="16">0 + //2
1 – //2</td><td rowspan="8">0 + //3</td><td rowspan="4">0 + //4
(0 + //3 + //1)</td><td rowspan="2">0 + //5</td><td>0 + //6</td></tr>
<tr><td>0 + //4 + //2</td></tr>
<tr><td rowspan="2">0 + //4 – //2
0 + //3 + //2</td><td>0 + //4 – //3</td></tr>
<tr><td>0 + //3 + //3</td></tr>
<tr><td rowspan="4">0 + //3 – //2
0 + //2 + //2</td><td rowspan="2">0 + //3 – //3</td><td>0 + //3 – //4</td></tr>
<tr><td>0 + //3 – //2 + //2</td></tr>
<tr><td rowspan="2">0 + //2 + //3</td><td>0 + //2 + //2 + //2</td></tr>
<tr><td>0 + //2 + //4</td></tr>
<tr><td rowspan="8">1 – //3</td><td rowspan="4">1 – //2 + //2
1 – //3 – //2</td><td rowspan="2">1 – //2 + //3</td><td>1 – //2 + //4</td></tr>
<tr><td>1 – //2 + //2 + //2</td></tr>
<tr><td rowspan="2">1 – //3 – //3</td><td>1 – //3 – //2 + //2</td></tr>
<tr><td>1 – //3 – //4</td></tr>
<tr><td rowspan="4">1 – //4</td><td rowspan="2">1 – //3 + //2
1 – //4 – //2</td><td>1 – //3 + //3</td></tr>
<tr><td>1 – //4 – //3</td></tr>
<tr><td rowspan="2">1 – //5</td><td>1 – //5 – //2</td></tr>
<tr><td>1 – //6</td></tr>
</table>

And here is a table of the fractions which they generate:

<table>
<tr><td rowspan="16">1/2</td><td rowspan="8">1/3</td><td rowspan="4">1/4</td><td rowspan="2">1/5</td><td>1/6</td></tr>
<tr><td>2/9</td></tr>
<tr><td rowspan="2">2/7</td><td>3/11</td></tr>
<tr><td>3/10</td></tr>
<tr><td rowspan="4">2/5</td><td rowspan="2">3/8</td><td>4/11</td></tr>
<tr><td>5/13</td></tr>
<tr><td rowspan="2">3/7</td><td>5/12</td></tr>
<tr><td>4/9</td></tr>
<tr><td rowspan="8">2/3</td><td rowspan="4">3/5</td><td rowspan="2">4/7</td><td>5/9</td></tr>
<tr><td>7/12</td></tr>
<tr><td rowspan="2">5/8</td><td>8/13</td></tr>
<tr><td>7/11</td></tr>
<tr><td rowspan="4">3/4</td><td rowspan="2">5/7</td><td>7/10</td></tr>
<tr><td>8/11</td></tr>
<tr><td rowspan="2">4/5</td><td>7/9</td></tr>
<tr><td>5/6</td></tr>
</table>

It is immediately clear that the table is identical to the Stern-Brocot tree.

Now the Stern-Brocot tree was generated using Farey addition. In order to prove that the two tables are identical, all we have to show is that sequences obtained by incrementing the last coefficient of the NICF obey Farey addition. Take the sequence 1 – //3 – //k where $k >= 2$ (which is 3/5, 5/8, 7/11 …) Each one is the Farey sum of the previous member of the sequence and the fraction towards which the sequence is heading – in this case 2/3. It is easy to see why this must be the case. Suppose an NICF ends with ± //j ± //k. Evaluating the last coefficient gives us

$$\ldots \pm // j \pm // k = \ldots \pm // \frac{jk \pm 1}{k} = \ldots \pm \frac{k}{jk \pm 1}$$

As we evaluate the NICF back to the beginning, the numerator and denominator get more complicated but they will always remain a linear function of k. The result, when all the coefficients have been evaluated will be a fraction of the form $\frac{pk \pm r}{qk \pm s}$. The start of the sequence will be when $k = 1$ or $k = 2$. The fraction towards which the sequence is heading will be when k tends to ∞. This will be $\frac{p}{q}$. Two adjacent members of the sequence will be $\frac{pk \pm r}{qk \pm s}$ and $\frac{p(k+1) \pm r}{q(k+1) \pm s}$ which is, of course, the Farey sum of $\frac{pk \pm r}{qk \pm s}$ and $\frac{p}{q}$. Q.E.D.

Having shown that the Stern-Brocot tree can be generated using NICF's, it is almost self-evident that every rational number will appear once and once only in the tree but the simple proof is as follows:

Once you have decided whether the nearest integer to ½ is zero or one, every rational fraction has a unique NICF. Suppose there exists a rational fraction whose NICF which is not to be found in the tree. Suppose that the order of this NICF is n and that the last coefficient is k i.e. the NICF ends in //j ± //k. Since the Stern-Brocot tree is made up of sequences in which the last coefficient takes on all possible values >= 2, none of these can be in the tree either; nor can the fraction whose NICF ends in //j ± //1. This particular NICF would be written as a NICF of order n – 1 ending in //j ± 1.

In other words if there exists a rational fraction of order *n* which is not in the tree, then there must exist a rational fraction of order $n-1$ which is not in the tree either. Taking this to its logical conclusion, it follows that there must be a rational fraction of order 1 which is not in the tree. But this is not the case, therefore our theorem is proved.

Even more remarkable is the fact that the Stern-Brocot tree only contains fractions in their lowest possible terms. This is due to fact that the evaluation of any continued fraction only generates such fractions for the following reason:

During the evaluation of any continued fraction you are continually turning an expression of the form $//a \pm b/c$ (where *b* and *c* have no common factor) into an ordinary fraction: $c/(ac \pm b)$. We wish to prove that if this is the case, then *c* and $(ac \pm b)$ have no common factor either. If they did we would be able to write $c = np$ and $(ac \pm b) = mp$ or $anp \pm b = mp$. But this would imply that $b = \pm p(m - an)$ which contradicts our assumption the *b* and *c* have no common factor. Now every continued fraction ends with the fraction $1/k$ and since 1 and *k* have no common factors, then the result of every calculation during the evaluation, including the last which generates the rational fraction which the CF represents, will have no common factors either.

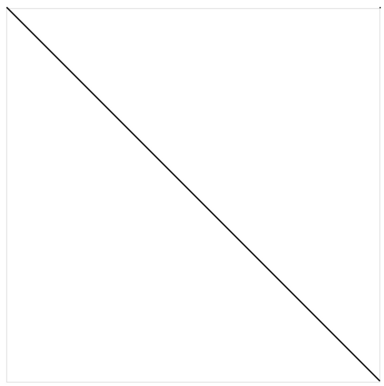
Complementary Numbers

Every rational fraction p/q has what I call a *first order complement* equal to $1 - p/q$. In the Stern-Brocot tree the complements occupy places which are a mirror image of each other about the centre line. Looking at the NICF's of a pair of first order complements you will see that they have the following form $0 + //a \pm //b \pm \dots$ and $1 - //a \pm //b \pm \dots$

Now look at just the upper half of the tree. All the fractions have complementary reflections about the centre line such that $0 + //a + //b \pm \dots$ becomes $0 + //a{+}1 - //b \pm \dots$ and $0 + //a - //b \pm \dots$ becomes $0 + //a{-}1 + //b \pm \dots$ I call this a *second order complement*.

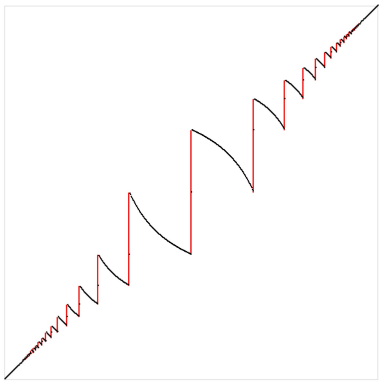
The rule should be pretty obvious. To generate the n^{th} order complement of a continued fraction you change the sign before the n^{th} coefficient (not counting the zeroth coefficient of course) and if you changed it from + to – you *increase* the $n{-}1^{th}$ coefficient by 1; and if the sign was originally – and you changed it to + you *decrease* it by 1 . Doing this procedure twice will, of course, return to the original number as expected.

Here is a graph which was generated by calculating the first order complement of the numbers between 0 and 1. Not surprisingly, the graph is just a graph of $y = 1 - x$.



First complement

Here is a graph of the second order complements between 0 and 1.



Second complement

It is worth studying this graph in some detail. The graph consists of a series of arcs interspersed by discontinuous jumps (indicated in red). These jumps occur at 1/2, 1/3 and 2/3, 1/4 and 3/4, 1/5 and 4/5 etc. Lets examine why there is such a jump at $x = 1/4 = 0 + //4$

The NICF of a number which is slightly less than 1/4 is $0 + //4 + e$ while the NICF of a number which is slightly greater then 1/4 is $0 + //4 – e$. If we apply the second order complement to these numbers we get $0 + //5 – e$ and $0 + //3 + e$. These are obviously very different.

In general the size of the jump at any fraction of the form $1/n$ (n > 2)will be

$$\frac{1}{n-1} - \frac{1}{n+1} = \frac{2}{n^2-1} \tag{13}$$

so for example the length of the red line at $n = 4$ is 2/15.

(When $n = 2$ the situation is slightly different. The NICF of a number slightly smaller than 1/2 is $0 + //2 + e$ while the NICF of a number slightly larger is not $0 + //2 – e$. but $1 – //2 + e$. The second order complements of these are $0 + //3 – e$ and $1 – //3 + e$ and the difference between these is not 2/3 in accordance with equation (13) but only 1/3.)

Another question that might occur to you is this. Why are the segments curved and not straight?

Consider the section between $x = 1/3$ and $x = 1/2$. In particular, consider a number slightly larger than 1/3 e.g. 0.35 or 7/20. The NICF of 7/20 is $0 + //2 + //1 + //6$. The second complement of this number is $0 + //3 – //1 + //6 = 7/15$, In general, the NCIF of all the numbers slightly larger than 1/3 have the form $x = 0 + //2 + e$ (where e is a number between 0 and 1) $= 1/(2 + e)$ and the second complement of x has the form $y = 0 + //3 – e = 1/(3 – e)$. If we substitute $e = 1/x – 2$ into this expression we obtain $y = x/(5x – 1)$ which is a rectangular hyperbola – hence the curves.

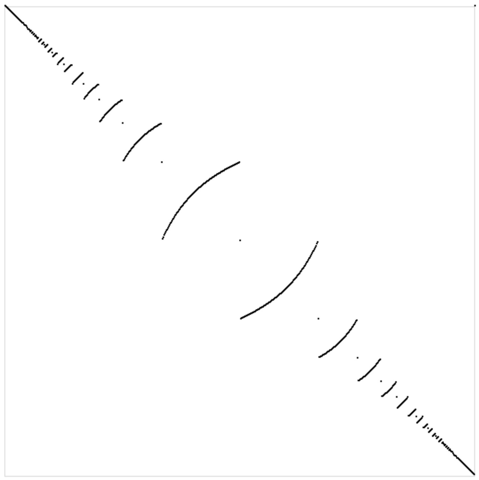
Hofstadter's INT function

At last!

Well, not quite yet. We have seen how we can apply the complement process to the first and second coefficients separately. What happens if we apply both at the same time? And does it matter in which order we apply the complements?

We can answer the second question straight away. No the order doesn't matter. Consider the number $a + //b + //c$. Applying complements to the first and second coefficients in that order gives us first $a + 1 – //b + //c$ and then $a + 1 – //b + 1 – //c$. Applying the second complement first gives us $a + //b + 1 – //c$ and then $a + 1 – //b + 1 – //c$ which is exactly the same.

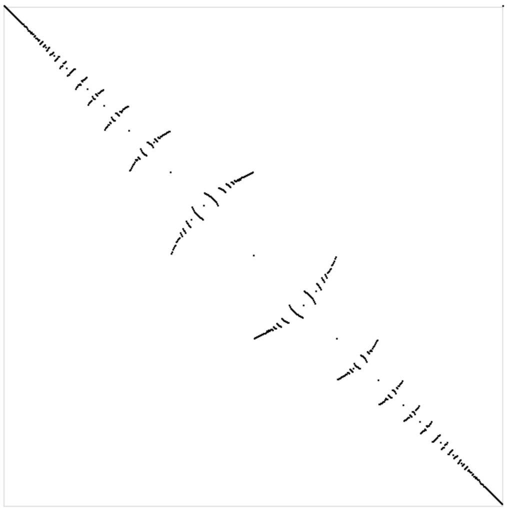
This is the result of applying both complements:



First and second complements

No surprises there.

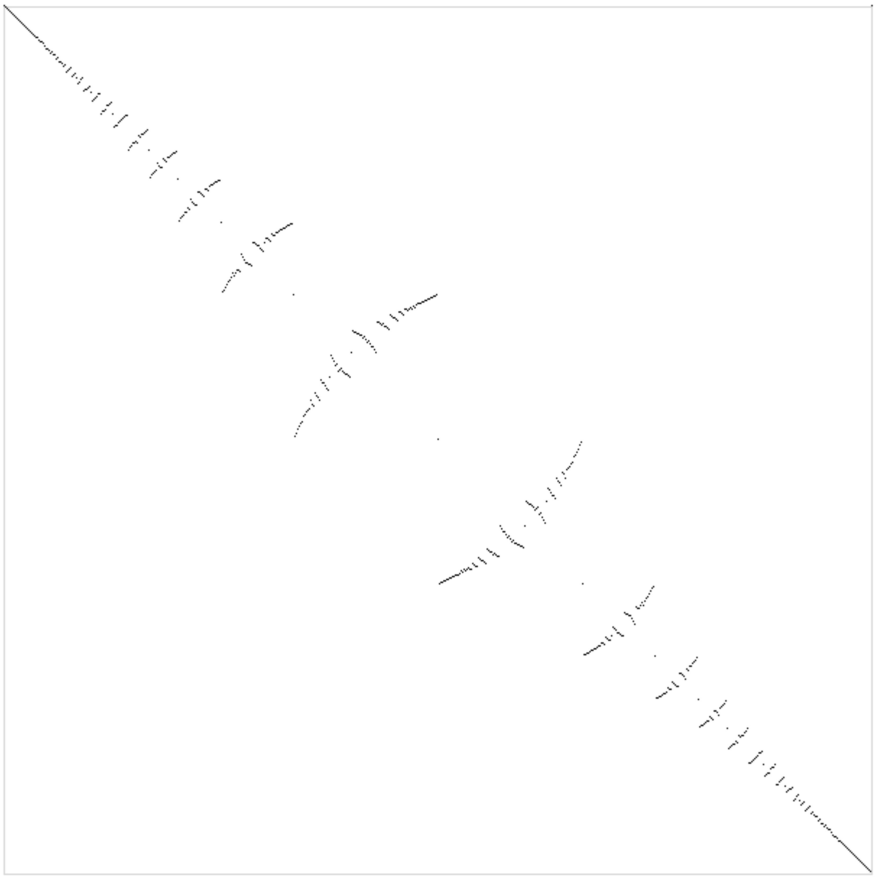
But what if we apply the third complement as well?



First and second and third complements

Wow! Each arc has been divided into an infinite number of smaller arcs in a self similar pattern.

But why stop at three? In general any number between 0 and 1 can have any number of coefficients so we must complement *all* of them. The result is Hofstadter's famous INT function which he (rather cryptically) described on page 140 in his brilliant book 'Godel, Escher, Bach' published in 1979. Every arc is composed of tiny arcs each of which are composed of even tinier arcs and so on in infinite regress.

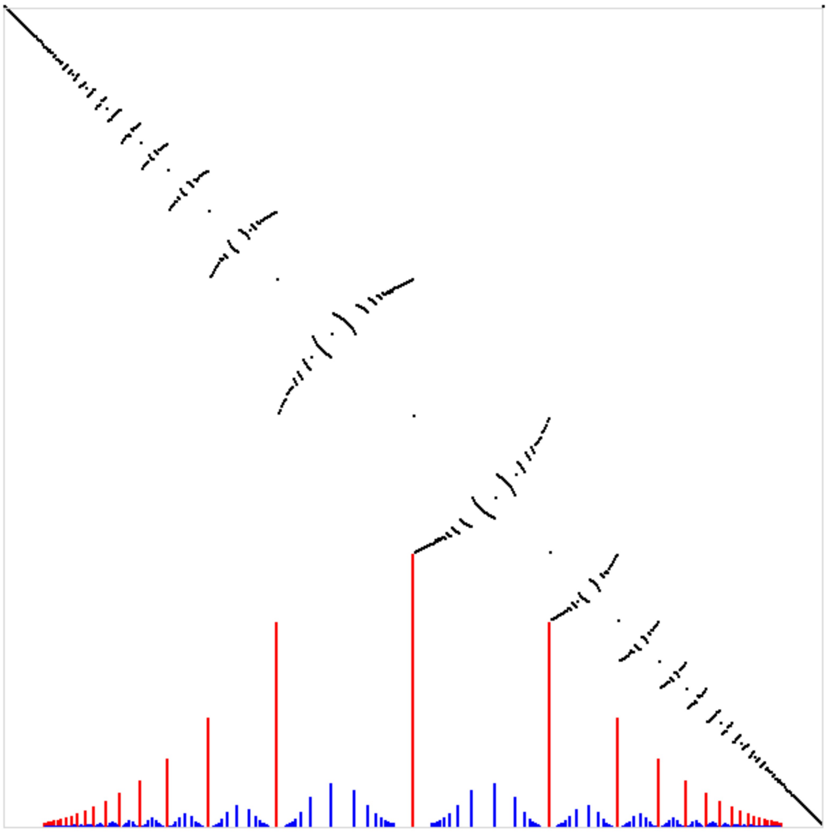


Hofstadter's INT function

The structure of the number line

Previously I suggested a rather crude method of calculating the 'rationality' of a rational number. Hofstadter's INT function shows us a better way.

We have seen that at every rational fraction, the INT function has a discontinuity and that the discontinuity is characterised by a jump of finite length. We can define the 'rationality' of a number as the length of this line. For numbers of the form $1/n$ (what we might call *first* order numbers) this is $2/(n^2 - 1)$ (equation 13). Now it is fairly easy to derive an expression for the 'rationality' $\boldsymbol{R}$ of *second order* numbers of the form $0 + //n \pm //m$ and a plot of $\boldsymbol{R}$ against x will show a series of spikes of varying height at every rational value of x. Naturally the rationality of the irrational numbers will be zero and the rationality of the integers is 1.



First and second and third complements with rationality graph showing the rationality of first order numbers in red and the rationality of second order numbers in blue

I like to think that the rationality graph gives us a vivid image of all the numbers on the number line. If we were to include the rationalities of the third order numbers in green, each gap between the blue lines would be filled with an infinite number of green lines between each of which would be an infinite number of even smaller lines corresponding to the rationality of the fourth order numbers and so on and so on.

And of course, even if we filled the whole graph with coloured lines at every rational number there would still be an uncountable number of black dots between all of them representing the irrational numbers.

Carr Bank: August 2020