Tilling the Collatz Tree: A new approach

by Fernando Franco Félix

Abstract

We find three tiles that can be used to construct the entire Collatz Tree and we study their properties. Then we find an algorithm that can find any sequence of tiles in the Collatz Tree, this is the link to download a python code that performs that algorithm: github.com/Frigorifico9/ConstellationFinder.

Oddly enough in some cases the algorithm requires rational numbers to find some constellations, from there we generalize the Collatz Tree to all real numbers. Finally we discuss how to generalize this approach and if it could have some applications studying chaotic systems.

But this wouldn't be a paper about the Collatz Conjecture if we didn't also try to prove it. In doing that we find an equation whose integer solutions correspond to all the possible cycles, and we seem to find that an infinite sequence cannot exist.

1 Introduction

In this work I will go rather slowly and I'll try to provide proofs and examples for everything, but I also write this in such a way that it should be easy to jump ahead if you've already understood the main idea. For example, if you don't need me to prove that the Collatz Tree always branches in numbers of the form $6n+4$ you can continue to "Node taxonomy" without any problem.

When we apply the Collatz rules to any number we get a sequence. These rules are: if n is even the next number is $n/2$, and if n is odd the next number is $3n+1$. If two sequences share one number in common they become identical form that point on, so we can arrange them into a tree. The "trunk" of this tree is the sequence 16-8-4-2-1, and from 16 it splits into the many branches above. The conjecture says that all the positive integers generate sequences that are part of this tree.

This tree has some branches that go on forever without branching again, made by multiples of 3, but they are not very relevant for this work, I just wanted to mention them so you know I didn't forget about them.

I decided to focus on the vertices where the tree branches, what I'm gonna call for now the "branching vertices". For a branching vertex v I noticed it seemed to always be connected to three other vertices. The one bellow (in the direction of the trunk) is $v/2$, one of the vertices above is $2v$ and finally the last vertex is $(v-1)/3$. Let's prove that this is the case.

The rules say that if you have a number r which is odd you should multiply it by three and add 1. Since r is odd we can write as $r = 2n + 1$ and then the next vertex in the sequence is $3(2r + 1) + 1 = 6n + 4$. This number will be even, and the rules say that to get the next number in the sequence we have to divide it by 2, that is $3n + 2$. But since the vertex $6n + 4$ is even there must be another number $12n + 8$ which is the previous number to $6n + 4$ in some other sequence. In this way we have found that $v = 6n + 4$ will be a branching vertex and the three vertices that surround it: $v/2 = 3n + 2$ is bellow, $2v = 12n + 8$ is the even number above and $(v - 1)/3 = 2n + 1$ is the odd number above.

In order for any other vertices to be connected to $6n + 4$ there should be an integer $r \neq n$ such that $6r + 4 = 6n + 4$, which is clearly impossible, so we know for a fact that every branching vertex is of the form $6n + 4$ and it is connected to only three other vertices.

2 Node taxonomy

2.1 Proving there are only three families of nodes

Since “branching vertex” is too long to say, I will use the word “node” to refer to them exclusively. I apologize to graph theorists if there was better terminology I could have used, I wasn’t aware of it.

Every node is connected to two other sequences above, which themselves continue to branch into other sequences, but below there’s only one sequence, although it is possible that at some point another another sequence joins that sequence too. Notice that this would be true even if the conjecture was false. In that case it would just mean that there is an infinite amount of vertices bellow, and the value of these vertices increases as the sequence continues, or it could be that the vertices eventually form a loop.

I noticed that there seemed to be only three possible ways that a node could be connected to another node. First, it could be that you divide by two and reach an odd number

$$\frac{6n+4}{2} = 2k + 1 \rightarrow n = \frac{2k-1}{3} \tag{1}$$

Once you reach an odd number you just apply the rules and you end up in a node $6r + 4$. Another possibility was that you divide by two, reached a vertex that is not a node, divide by two again and then you reach another node:

$$\frac{6n+4}{4} = 6k + 4 \rightarrow n = 4k + 2 \tag{2}$$

And yet another possibility was that you divided by two twice like before, but instead of reaching a node you reached an odd number, at which point we know for a fact the next one has to be a node.

$$\frac{6n+4}{4} = 2k + 1 \rightarrow n = \frac{4k}{3} \tag{3}$$

It seems obvious that all of these possibilities exist, but are they really the only ones? Yes. One way to prove is is to see that for every n there is always a number k that satisfies one of these three equations.

As we increase k the numerator in the first equation advances by 2 starting on -1, which means it goes through all odd numbers, and when it hits a multiple of 3 it cancels out with the three in the denominator, leaving us with an odd number.

The same happens with the third equation excepts the numerator advances by 4 starting on 0, so every third number it hits a multiple of 3 and 4, the 3 cancels out, and we end up with a multiple of 4.

Finally the second equation takes a number, multiplies it by 4 and adds 2, which results in all the even numbers that are not multiples of 4, which are exactly the numbers the third equation doesn't produce.

This means that for every n one of these equations is true, which means that every node $6n+4$ is either one or two vertices away from the next node in the sequence.

Since we have three families of nodes I decided to call them Stark, Lannister and Targaryen, because I couldn't think of three rival families in real world history, all the examples I could think of were only two rival families.

In the next sections we will find the formulas that define the properties of each family of nodes.

2.2 Targaryen nodes

The idea of dividing by 4 and always reaching another node seems incestuous somehow, so I will call this family of nodes "Targaryen".

What we want is to find a set of formulas that characterize all nodes in this family, for this we need to look at the relationship this family of nodes have with the next node:

$$\frac{6n+4}{4} = 6s+4$$

The solutions to this equation define all the properties of this family of nodes: $(n = 4a+2, s = a)$.

Let's call a the "seed" of the node.

2.3 Stark and Lannister

The family of nodes that become odd numbers when divided by 2 will be called "Stark", and the family of nodes that become odd when divided by 4 will be called "Lannister".

For the Stark nodes we need the solutions to this equation:

$$3\left(\frac{6n+4}{2}\right) + 1 = 6s+4$$

Which are $(n = 2a+1, s = 3a+2)$. Notice how $s > a$, this is because this is the kind of node that can actually make the sequence increase in value.

Finally for Targaryen nodes we solve =:

$$3\left(\frac{6n+4}{4}\right) + 1 = 6s+4$$

And we get $(n = 4a, s = 3a)$.

3 Streaks and Constellations

3.1 From the Collatz Tree to Constellations

Having found the families we can see the Collatz Tree as sequences of nodes of different kinds. In fact, we can represent the numerical sequences with just the kinds of nodes, for example, if we use the initial for each family of nodes to represent them, the sequence: 34-17-52-26-13-40-20-10-5-16-8-4-2-1 can be rewritten as $SLTSTL$. Except for the fact that there are many other sequences with identical structure elsewhere in the tree, so we also have to specify the point where the sequence starts, but once we do, it is uniquely defined (unless there are loops, but we'll come back to that later). Also, notice that the sequences don't have to go all the way to the bottom, for instance the sequence STL starting in 34 ends in 40.

When we have a sequence of nodes with an unspecified starting point I will call it a "constellation", and presumably the same constellation can appear many times all over the tree.

3.2 From Constellations to tiles

We can represent each kind of node visually with a sort of "tile". I call these tiles "the super awesome tiles", since I promised my younger self that if I ever discovered something I would call it "super awesome", whatever it was, and I'm a man of my word. We could have represented the super awesome tiles as simple shapes, but instead we commissioned an artist to make them more visually appealing.

Figure 1: The Stark Tile

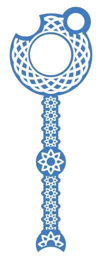

Figure 2: The Targaryen Tile

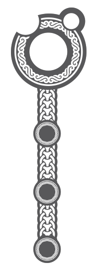

Figure 3: The Lannister Tile

The way they work is that the number of node goes in the middle of the big circle, and the little circles below it represent the numbers before reaching another node, except the last one, which is the connection to the next node. As you can see Stark and Lannister nodes connect to other nodes through the notch, which represents an odd number, while the Targaryen nodes connect to other nodes through the nub, which represents even numbers.

Notice that we don't have choice when it comes to Targaryen nodes, every node must be connected to one, but we have a choice when it comes to Stark or Lannister nodes. This means that we could use them to tile periodically or aperiodically. Which means there must be huge periodic and aperiodic sections of the Collatz Tree. Finding the numbers that start such sections, and the borders between them, seems like a very interesting problem, just saying.

But the point is that we can just start placing these tiles in any random order we want, without worrying about the numbers, and by the end we should have a constellation and it should be possible to find it in the Collatz Tree, and we will develop an algorithm to do that, but before we just want to make sure that there are no restrictions to the constellations we can construct, and we will find that there isn't. You can find any constellation, but they cannot start at any point.

3.3 Why all constellations are possible

If we have a node we can always calculate the following node, the question is, can we say anything about which family that new node should belong to?.

Let's begin with Targaryen nodes. We know that they can be written as

$$6(4a+2)+4=24a+16$$

And that the next node will be:

$$6(4a+2-3a-2)+4=6a+4$$

Which is the general form for all the nodes, which means that a Targaryen node can always be followed by any kind of node. Let's do Stark nodes now, they can be written as:

$$6(2a+1)+4=12a+10$$

And the next node will be:

$$6(3a+2)+4=18a+16$$

Notice that if $a=0$ we immediately get that 10 is a Stark node that can be followed by a Targaryen node which is 16, which we already knew, but it's nice to see the formulas work properly.

Anyway this $18a+16$ is not the general form of the nodes, so let's see if it can fit the form of the other nodes. First, we have to simplify it a bit:

$$18a+16=6n+4 \Longrightarrow n=3a+2$$

This is for a Stark node followed by a Lannister node:

$$3a+2=4\lambda$$

Notice that we are using the formula for n from the Lannister family of nodes, which is $4n$, but I'm writing it as 4λ to avoid confusion with $n=3a+2$. Anyway, this expression has integer solutions because there are multiples of four that are two units away from a multiple of three, like 8 and 6, which means that a Stark node could always be followed by a Lannister node.

We can do the same for a Stark node followed by a Targaryen node:

$$3a+2=4\tau+2$$

Once again this always has integer solutions because there are many numbers that are multiples of 4 and 3, so a Stark node could always be followed by a Targaryen node.

Finally we will see if a Stark node can be followed by another Stark node

$$3a+2=2s+1$$

Which, you guessed it, always has integer solutions. There's no restrictions at all as to what kinds of nodes can follow a Stark node, let's see if there are restrictions for Lannister nodes.

First, a Lannister node is:

$$6(4a) + 4 = 24a + 4$$

And the next node will be:

$$6(3a) + 4 = 18a + 4$$

Which means that we can get that n should be:

$$18a + 4 = 6n + 4 \rightarrow n = 3a$$

And we can use this to very easily see if a Lannister node can be followed by:

A Stark node:

$$3a = 2s + 1$$

A Targaryen node:

$$3a = 4\tau + 2$$

And another Lannister node:

$$3a = 4\lambda$$

And in each case we find that yes, they all have integer solutions, there are multiples of three two units away from a multiple of 4, and there are numbers that are multiples of three and four.

In conclusion, there are no restrictions for the construction of the constellations.

3.4 Streaks of nodes

Our previous conclusion seems wrong, or at least dangerous. I mean, Stark nodes are followed by nodes of higher value, if there's no restriction to how many you can have in a row, does that prove the conjecture is false? Not at all, and to see why, we need to introduce the concept of streaks, which will be fundamental for the rest of this work.

A streak of nodes will be a constellation of nodes all of the same kind, and the length of the streak will be the number of nodes of that type in the sequence.

What we want now is to use the formulas we've found to see if we can always calculate the start point and the end point of any streak, and we will begin with streaks of Stark nodes, since they are the most interesting ones (in my opinion).

Let's imagine a streak of Stark node starting in a node with a seed a_0, so that $n = 2a_0 + 1$, while for the next node $n = 3a_0 + 2$. If that is itself another Stark node we have that:

$$3a_0 + 2 = 2a_1 + 1 \rightarrow a_1 = \frac{3}{2}a_0 + \frac{1}{2}$$

If the next node was also a Stark node we would have:

$$a_2 = \frac{9}{4}a_0 + \frac{5}{4}$$

And in general:

$$a_m = \left(\frac{3}{2}\right)^m a_0 + \left(\frac{3}{2}\right)^m - 1 \tag{4}$$

The solutions to this formula are of the form:

$$a_0 = 2^m b + 2^m - 1 \tag{5}$$

$$a_m = 3^m b + 3^m - 1 \tag{6}$$

Where b can be any positive integer (I struggled so much finding that general form and its solutions, you have no idea). This means that if you want me to find a streak with m Stark nodes all I have to do is just to give some value to b. For example, if you want 3 Stark nodes in a row then I can just say $b = 0$, get $a_0 = 3$, $a_2 = 8$, and then I can confidently tell you that this sequence starts in the node 46, passes through the Stark nodes 70 and 106, and finally arrives at 160, which is a Targaryen node, ending the streak (test it for yourself if you are not convinced, that's what I did). And I could find infinitely many more such streaks by using other values of b.

This proves that there can't be an infinite number of Stark nodes in a row, although it proves that number can be arbitrarily large. Let me explain why.

Imagine there was a fixed number a_0 for which, if we start a streak of Stark nodes, we could always find a valid a_m for any m, then the conjecture would certainly be false, because this sequence of Stark nodes would result in numbers of increasing value, forever. But that's not the case. We know that a_0, the starting point of a streak, grows exponentially with the length of the streak, even if you choose $b = 0$, which means that if you wanted an infinite streak, you'd need to start at infinity, which is impossible. This notion will be crucial later, so I really hope it is correct.

This formula for the changing seed in a streak of nodes seems really cool, let's find the same for the other families of nodes.

3.5 Finding the formulas for the seeds of streaks of the other kinds of nodes

Let's first find the recursive formula for the seed of the Targaryen nodes (that sentence makes no sense out of context).

$$4a_0 + 2 - 3a_0 - 2 = 4a_1 + 2 \rightarrow a_1 = \frac{1}{4}a_0 - \frac{1}{2} \tag{7}$$

And in general:

$$a_m = \left(\frac{1}{4}\right)^m a_0 + \frac{2}{3}(4^{-m} - 1) \tag{8}$$

The solution for this formula is:

$$a_0 = 4^m a_m + \frac{2}{3}(4^m - 1) \tag{9}$$

Where a_n can be any positive integer, and don't be fooled by that 2/3, the constant is always an integer, because $4^n - 1 \; mod \; 3 = 0$ (it fooled me for a bit). This means that the constellations of Targaryen nodes can begin, or end, anywhere, which meakes sense, because if you are in a node you can always multiply by 4 to reach another Targaryen node.

Notice how the starting point grows exponentially the longer the streak is, just like with Stark nodes, which once again means that an infinite streak would need to start at infinity.

Now for the Lannister nodes:

$$4a_0 - a_0 = 4a_1 \rightarrow a_1 = \frac{3}{4}a_0$$

And in general

$$a_m = \left(\frac{3}{4}\right)^m a_0 \tag{10}$$

The solutions for this nodes are:

$$a_0 = 4^m b \tag{11}$$

$$a_m = 3^m b \tag{12}$$

Where b can be any positive integer. Once again we can see that the starting point grows exponentially with the length of the streak.

The reason the solutions for the Lannister and Stark nodes depend on b and the ones for the Targaryen nodes don't is because once you are in any node you can always multiply by two twice and reach another Targaryen node, so we can use any number to generate our streak. But with the Stark and Lannister nodes you have no guarantee that the next node will be of the same kind, so to generate our streak we need to "connect" the start and end points with the variable b.

Anyway, this is nice. Using this recursive formulas we can clearly see that the Stark nodes will always increase the value, because $3/2 > 1$ while the others will always decrease it because $1/4 < 1$ and $3/4 < 1$.

3.6 An observation about cycles

We can see that each kind of tile can create an infinite cycle. For example, an infinite sequence of Lannister nodes represents the sequence 4,2,1, while an infinite sequence of Stark nodes represents the sequence -2,-1, and finally an infinite sequence of Targaryen nodes represents the sequence 0,0,0.

As I will mention when we generalize the Collatz Tree, it seems that every super awesome tile can create a cycle, and this could have implications for the stable states of chaotic systems.

4 Multi-streak drifting

4.1 Intro to multi-streak constellations

Now we know the properties of streaks of a single kind of node, but of course, constellations can drift back and forth between different kinds of streaks of different lengths, even streaks of length 1, I call this multi-streak drifting.

To figure out the where such constellations could start and end we could use the formulas we just found recursively. For example, let's say that we had a constellation like this:

$$SSSLLLL \to S^3L^4$$

If you want to actually find this constellation somewhere in the Collatz Tree you can just calculate a_2 like we did before, and then we already know the next node can be written as $6(3a_2 + 2) + 4$ regardless of what kind of node it is, but if we want it to be the start of a streak of 4 Lannister nodes we know that

$$3a_2^s + 2 = 4a_0^\lambda \to a_0^\lambda = \frac{3}{4}a_2^s + \frac{1}{2}$$

Where I'm using the s and λ indices to differentiate the constants of different sequences. Anyway, we can then use this a_0^λ to calculate a_3^λ, and by the end we should have an equation we can solve to get formulas that should help us find all the nodes where this constellation begins and ends.

I am gonna call this whole process "threading", because it's like we are using a thread to connect different sequences of beads to make a collar, or a bracelet... Trees, constellations, tiles, and now bracelets, I'm all over the place with these metaphors, I know.

There are three interesting observations about threading I want to make. The first one is that the order matters.

$$S^3L^4 \neq L^4S^3$$

The second one is that whatever formula we find for the starting point of the constellation should be compatible with the formula for the starting point of the first kind of streak. If we think of it as sets, the set of solutions for the starting point for any constellation is a subset of the solutions for the starting point of the first streak, and the same is true for the solutions of the endpoints and the last streak.

Finally the third observation is about how we find the solutions to these equations. If we had something like this:

$$\alpha a_n = \beta a_0 + \gamma$$

We know that we can find the solutions using the extended Euclidean algorithm.

$$a\alpha X - \beta Y = GCD(\alpha, -\beta)$$

$$a_0 = \alpha b + Y\gamma$$

$$a_n = \beta b + X\gamma$$

Where b can be any integer.

And in fact one can see from the formulas for the streaks that α will always be some power of 2 and β will always be some power of three.

But there is one problem with this. We can use negative or positive numbers with the Collatz Rules and we get fundamentally different results, for instance, with negative numbers there are three cycles and not just one. For this reason if we are looking for a constellation we should be able to at least control whether we are looking for it with negative or positive numbers, and we can do that with a simple modification of the solutions:

$$a_0 = \alpha b + Y\gamma GCD(\alpha, -\beta)(mod\ \alpha)$$

$$a_n = \beta b + X\gamma GCD(\alpha, -\beta)(mod\ \beta)$$

By taking taking the modulo of the first coonstant on the second one we ensure that if $b < 0$ then $a < 0$... Until we discover that b has to take non integer values, but we'll get to that.

For now let's continue thinking about the Collatz Conjecture.

If the end point for any of these constellation can be arbitrarily larger than the start point, then the conjecture is false, and if the start and end points of any sequence with positive numbers can be the same, the conjecture is also false. For this reason we can reformulate the Collatz Conjecture in terms of the general solutions for an arbitrary constellation...

4.2 Infinite multi-streaks constellations require convergence for infinite sums of integers

If you thread any constellation you are quickly swamped with fractions, so let's simplify things by using Euler's number as a common base:

$$a_m = \left(\frac{3}{2}\right)^m a_0 + \left(\frac{3}{2}\right)^m - 1 \to a_m = e^{m(ln3-ln2)}(a_0+1) - 1 \tag{13}$$

$$a_m = \left(\frac{1}{4}\right)^m a_0 + \frac{2}{3}(4^{-m} - 1) \to a_m = e^{-2mln2}\left(a_0 + \frac{2}{3}\right) - \frac{2}{3} \tag{14}$$

$$a_m = \left(\frac{3}{4}\right)^m a_0 \to a_m = e^{m(ln3-2ln2)}a_0 \tag{15}$$

Now we can see that as we thread the equations we get these long sums in the exponent, something like this:

$$(ln3)\sum_{i=0}^{\infty} m_i$$

Where m_i stands for the length of each streak. Some of these sums are positive and some are negative, and this can bring the start and end points lower or higher. This is why a number like 82 is allowed to start an extremely long constellation that grows up to 9232 before coming back down, but here's the thing: Even if every finite sum will have a finite value, an infinite sum won't, because we are not summing fractions, each m_i is an integer, and an infinite sum of integers always diverges, no matter how slowly.

I could have made a mistake somewhere, but by my lights, this seems to prove that there cannot by any finite integer that starts an infinite constellation.

Alternatively, maybe this means that we cannot construct an infinite sequence, but it could still exists, it's just that it would be impossible to calculate its equation, because the streaks follow no pattern we can use to express the infinite sum. In that case if we found the a_0 that starts an infinite sequence, we would start calculating that sequence and we would never know if we are ever gonna get to the 4,2,1 cycle or not, and in fact, it would be impossible to know, like a Turing machine that doesn't know if it will halt. So even if we found this a_0, we would never know...

But then again, if it is impossible calculate the equation for that constellation, that means it is impossible to compute a_0, right?, and if a_0 is impossible to compute... Is it even a number? Is it even possible to find it by random chance? Because a random number generator is still performing a kind of computation, isn't it? I guess we could conclude that a_0 could exist, but if it does, there's no way to get to get to it, and even if you did, it would be so large no computer could make use of it, because by definition it is not computable... In summary, every number you can think of will generate a sequence that reaches the cycle 4,2,1, because if a number doesn't do that it's literally not a number you can think of

4.3 The formula to find all cyclical constellations

Sadly we weren't able to find a proof that a cyclical constellation has to use negative numbers, BUT we were able to find a formula whose integer solutions correspond to cycles, so if someone helped me find all the integer solutions we could prove or disprove the conjecture. This is the formula:

$$b = \frac{X\gamma GCD(2^w, -3^r)(mod\ -3^r) - Y\gamma GCD(2^w, -3^r)(mod\ 2^w)}{2^w - 3^r} \tag{16}$$

Where w and r are integers, and so is gamma And you probably don't even need γ nor the modulo, I'm just keeping them for convenience so that it matches my examples in just a moment.

Finding this formula was surprisingly easy: We just realized that if there was a cyclical constellation we could calculate its equation, and since it's a cycle $a_0 = a_n$, so all we need to do is take the general formulas for the solutions, set them equal to each other, and solve for b. The numbers for which this has integer solutions have to be cycles. However I wasn't able to find all the solutions to this equation, because doing the euclidean algorithm with arbitrary powers of 2 and 3 is really difficult, I couldn't figure out a way to express the solution in a single expression, but I'm sure someone else will figure it out and we can share the credit.... Assuming this formula is correct of course, which may it's not.

In the case of all 5 known cycles if we use this formula we do get a value of b that gives us a value of a_0 which when used to generate a sequence does result in the correct cycle.

For the cycle: $SSSLSSTLS$ the equation and solutions are:

$$2048a_n = 2187a_0 + 695$$

$$a_0 = 2048b + 2043$$

$$a_n = 2187b + 2182$$

And plugging these values into the formula for b we get:

$$b = \frac{2182 - 2043}{2048 - 2187} = -1$$

Using that value of b we get that this cycle is made by:

$$-50, -74, -110, -164, -122, -182, -272, -68, -50$$

For the cycle: SLS the equation and solutions are:

$$8a_n = 9a_0 + 2$$

$$a_0 = 8b + 6$$

$$a_n = 9b + 7$$

And plugging these values into the formula for b we get:

$$b = \frac{7 - 6}{8 - 9} = -1$$

Using that value of b we get that this cycle is made by:

$$-14, -20, -14$$

For the cycle: SS the equation and solutions are:

$$2a_n = 3a_0 + 1$$

$$a_0 = 2b + 1$$

$$a_n = 3b + 2$$

And plugging these values into the formula for b we get:

$$b = \frac{2-1}{2-3} = -1$$

Using that value of b we get that this cycle is made by:

$$-2, -2,$$

4.4 Discussion of the nature of cyclical streaks

Since these are cycles we could have started them in any node, and we don't have to stop after one revolution, we could keep going, expressing that as long and longer constellations, all of which have different equations, and these equations can help us find infinitely many examples of such constellations, but they are only cyclical with this particular value of b. We have to keep this in mind if we want to find all the possible solutions to this formula, so that we can safely ignore all the solutions that correspond to the cycles we already know.

5 Generalization of results

5.1 What if b wasn't an integer though?

First of all, there can only be cycles when b is an integer, because 2^w and 3^r have obviously no prime factors in common, so the only way a_0 is gonna be an integer is if b is an integer too... But if the constellation isn't a cycle, then it is possible to have a fractional b and end up with a valid constellation.

For example, with the constellation $SSSLLLTS$ (which was chosen at random) we find that we can find valid sequences with any integer value of b, but also with $b = 1/2$ and $1/3$ and $-2/3$ and many, many other fractions, positive and negative.

For every constellation you can think of there are instances of that constellation we cannot find, unless we use fractional values of b.

This caused me to have a crisis of faith when I first found it, but I came to accept it, and then I wondered: Why should we only accept the values of b that result in sequences of integer numbers?

If we are gonna accept some non integer values of b, we should accept all of them, even if it results in non integer values of a_0.

The obvious problem with this that if a_0 is not an integer then the node we make using it is not gonna be an integer either, which means that it's not even nor odd, how can we continue the sequence then? Easy, we use the formulas for the streaks of nodes to find the value of each node in the sequence, because they work with any real number, and in this way we are dodging the Collatz rules which would force us to work with integers only.

In this way we have generalized the Collatz Tree from integer numbers to all real numbers, in a manner that reminds me of how the Gamma Funcion generalized factorials to all complex numbers.

I know there existed extensions of the Collatz rules to fractional numbers, but these used the parity of the denominator of the fraction, and we don't need to do that here.

Now we can see that the Collatz Tree was just a subsection of a much larger tree definied for all real numbers, and which I have to call the "Super Awesome Tree", just in case someone discovered the super awesome tiles already and gave them some other name.

This tree has uncountably many branches, but amazingly we can still find our way through it because we can tile any path through it using the super awesome tiles. However this has a couple of surprising implications. First, any number can be a node of any kind, and we can always make a cyclical constellation, it just so happen that these sequences will contain numbers that are not integers.

But at this point it is obvious that the "Super Awesome Tree" is not special, there must be infinitely many similar trees we can make if we generalize a few aspects about it.

5.2 Creating more "Star-making functions"

To find other structures like the Super Awesome Tree we just need to find more functions that create these "constellations", which I'm gonna call "star-making functions".

Every star-making function is composed of a set of N functions, let's call this set the "design" of the tree. Each of these functions has a different domain, their intersection is the empty set, but they map onto the same codomain, and in fact, the intersection of their images is not the empty set. For example, the design of the Collatz star-making function is made of $f_1(x) = x/2$ and $f_2(y) = 3y + 1$. Where x represents only even numbers, and y represents only odd numbers. However the domain and codomain of these functions intersect in interesting ways. The domain of f_1 is all the even integers, and it maps onto all the integers, which can be even or odd, while the domain of f_2 is all the odd numbers, and it maps onto the integers of the form $6n+4$, which is a subset of the even numbers.

As you can see there are three ways to move between these different sets, this creates the branches in the tree, this is the reason there are three different tiles, and in general, if you design a tree you can know how many super awesome tiles you need to move through it by looking at how the images of the functions intersect. This gets me to my first conjecture:

Conjecture one: Every kind of super awesome tile can create an ouroboros.

In the case of the Collatz Tree this is easy to see: Lannister tiles create the cycle 4,2,1, Stark tiles create the cycle -2,-1 and finally Targaryen tiles create the cycle 0,0 (which has $b = -2/3$ by the way, yet another reason why we need fractional values of b). The question is if this could happen with every possible design of every possible star-making function.

But here's something weird: Even if we allow b to be any real number and we have sequences made of numbers that are not integers, we still have discrete steps in our sequence, but what if we didn't? It seems to me that just like we can transform series of discrete sums into continuous integrals it should be possible to transform these discrete constellations into continuous ones, although I have no idea how. BUT if this is possible my second conjecture is about it:

Conjecture two: Threading through continuous constellations requires the use of pi and the Feigenbaum constants, perhaps even Euler's number. I have no reason to conjecture this, I just feel this is the case.

I guess it could also be possible to end up in a tree that requires infinitely many different kinds of super awesome tiles, maybe that's how they are continuous, but as long as these tiles follow some logical set of rules you should be able to tile with them.

5.3 Possible applications to chaotic systems

Maybe these tiles could be applied to the study of chaos. Consider a double pendulum and imagine its initial conditions as a point in phase-space. As we let the pendulum go the point moves through phase space, making loops, and arcs, going up and down, but now let's consider there was friction, in that case we know that eventually the double pendulum will stop, we know it will eventually reach a state where it is just hanging straight down. This is the root of the tree.

As the point moves it will eventually reach the point that represents hanging straight down, and it will stay there, and the same will be true of all other possible points. They all flow to this one point, in the same way that all the positive integers flow to the constellation of 4,2,1.

For this reason it should be possible to study how the paths of these points eventually meet, forming a tree in phase space, but these paths would join in different ways, and we could figure out what these possible connections can be using the laws of physics, and those would be the super awesome tiles for this tree. It doesn't even matter if there are infinitely many possible tiles, as long as we know the rules to generate them we should be able to thread any possible behavior we want, and find the set of all possible initial conditions that create the desired behavior. Then if we could reproduce those initial conditions in real life we could have the double pendulum behave however we wanted, and stop it once the desired behavior has concluded.

Now, you may argue that not all chaotic systems eventually reach the same stable state, some have many different attractors, and you are right. I think these attractors can be understand as the root of the different trees that describe that chaotic systems. In fact, if we figure out the tiles that describe a chaotic system we could use the tiles to look for cycles and find all the possible attractors. In fact, this is my third conjecture:

Conjecture three: If we find the tree for the three body problem we will find its super awesome tiles can form five non trivial cycles, one for each Lagrangian point.

The double pendulum is a simple example, without any applications I can think of, but it is a proof of concept. If we can do this with a double pendulum, it should be possible to do it with any other chaotic system, like the weather, and the economy. Even if we cannot control the initial conditions of these systems I know people are clever and they will find a way to use it. Perhaps being able to calculate the possible constellations of these systems would have some predictive power.

In summary, while it is impossible to tame the storm that is chaos, with these tiles, we should be able to flow through it unscathed, letting the current take us wherever we want to go, after all, even if we miss our chance, there should always be another one nearby.

I have no way to know how much of this work is correct or incorrect, but I hope that even if most of it is incorrect there is still a portion of it that has some value. Thanks for reading.