# Algebra of a Branch

## A New Model for the Proof of the Collatz Conjecture


Version: 1.0 Author: Franz Ziegler, Dipl. El.Ing ETH, MAS MTEC ETH Zurich

Date: Dec. 15. 2020 ORCID 0000-0002-6289-7306


## 1. Summary


This paper is focused on the representation of the Collatz number system with loop-free unbranched fragments of Collatz number series. In contrast to the previous works [01]and [02], where the proof of the Collatz conjecture was found by splitting the reduced odd-numbered Collatz tree only, this work used the whole space of natural numbers. To get a loop-free and unbranched representation of the Collatz number system, the model, which initially only consisted of the simple Collatz rules, had been gradually enhanced by additional constraints, until the consistent "Matchstick Model" was found and thus had confirmed the Collatz conjecture again. The found model constraints are corresponding the method proposed in [01] and [02], where the Collatz number tree was cut along the proxy path. Because of it's simplicity, any proof of the Collatz conjecture could lead to the same fundamental numbers and constraints. – With the Matchstick Model the most minimalistic model was found to split the Collatz tree into loop-free, unbranched fragments of Collatz number series.


## 2. Prerequisite

Assumed are the standard Collatz number series and the resulting Collatz number structure generated, after running the Collatz series starting with any possible natural number. It was shown in [01], that in a Collatz number structure all natural numbers (the zero was not used) appear exactly once.

## 3. Methods

The idea to describe the Collatz number structure as composition of unbranched fragments of Collatz number series is used again in this work. But this time, a list of strict conditions, that must apply to a loop-free structure, is compared with the model and tested for contradictions.

## 4. The Requirements List

Because the model deals with the full space of natural numbers, there will be open connections with all matchsticks. We distinguish two different connection types: The even numbered "plug" or terminal and the odd numbered "socket" or gate. A terminal cannot plug into a gate of the same number.

In order to map a tree structure with linear, unbranched chains, the following conditions must be met:

**Condition 1:** A stop condition is needed at the foot: A tree can only start to grow, when something is there to fix the seed. In nature this can be the garden soil or a flower pot. Within the Collatz number structure it is the known loop of natural numbers: 4, 2, 1, 4..., which lie at the end of any Collatz number series. This loop should not be considered as an unfriendly irregularity of the famous number structure, but rather as an important building block for a loop-free number structure and will help to confirm the proof. Because this loop has a "gate" at the number 4 to enter, but no terminal to exit, this loop will absorb the terminal of the first branch (see below). Because the end loop is a boundary condition, it has no influence on the model design.

**Condition 2:** The branches must have one single (even number) terminal: Every tree can be cut into unbranched branches or chains. Any branch that is cut off, will naturally be given an interface. It is the terminating point of a branch, called terminal. A unbranched branch could have two terminals at both ends, but we allow only one. Other branches, which were removed from this branch also leave traces. We call them gates, because we can imagine a reconstruction of a cut tree. When the Collatz number structure is indeed a tree, it can of course also be cut into branch-free sticks.

**Condition 3:** A model branch has to be unbranched. That means it is a chain with no forking at any side. So sideways no (even number) terminals are allowed, but the number of (odd number) gates is not limited.

**Condition 4:** A branch also must not have a second (even number) terminal at the end of the chain: Since the unbranched sticks must have connectivity at one side for docking to other branches or to the "flowerpot", each branch end must have the inverse property, namely, no ability to plug into another branch. The idea behind allowing only one terminal is very simple: this principle allows no loops after assembling.

**Condition 5:** Any branch must offer one (even number) gate minimum: That condition simply enables the next connection step. Because all numbers are unique, opening a unique new gate numbers allows finding a yet unused branch to attach.

**Condition 6:** We have to deal with infinitely long series of numbers: It has to be ensured, that such series of numbers can be mastered easily with mathematical means.

## 5. Analysis of the Dynamics of the Collatz Number Generator

A dynamic (number) system differs in that it can store and change it's system state. Even very simple dynamic systems of 1st order (one state variable) can thereby behave in a complex manner. The Collatz "number generator" is such a simple 1st order system. It maintains a single counter reloaded each time the Collatz rule (3x +1) is applied and decreases the counter, if the Collatz rule (x/2) is applied. Example: Starting with the number 13:

| | | | |
|---|---|---|---|
| 13 | Counter = **0** | (13 * 2 ^ **0** = 13 ) | now applying Collatz (3x +1): |
| 3 * 13 +1 = 40 | Counter = **3** | ( 5 * 2 ^ **3** = 40 ) | now applying Collatz x/2: |
| 40 / 2 = 20 | Counter = **2** | ( 5 * 2 ^ **2** = 20 ) | now applying Collatz x/2: |
| 20 / 2 = 10 | Counter = **1** | ( 5 * 2 ^ **1** = 10 ) | now applying Collatz x/2: |
| 10 / 2 = 5 | Counter = **0** | ( 5 * 2 ^ **0** = 5 ) | again Collatz (3x +1): |
| 5 * 3 + 1 = 16 | Counter = **4** | ( 2 ^ **4** = 16 ) | ..... |

**Fig. 1:** The system state counter 2^n used by the Collatz number generator

Assuming the Collatz conjecture is true, the amazing thing about this minimalistic dynamic system is the fact, that the sequence of the reloaded counter values will never repeat (except at the known loop).

The following proposal of the Matchstick Model will touch the range of dynamics in its body and will restrict the counter values to 0, 1, and 2. The idea, that dynamics could be split into a "horizontal" and a "vertical" component, would mean the Matchstick Model will lose it's "horizontal" component. We do not further discuss this interpretation in this paper and therefore there is no need to perform deeper analysis of system dynamics here.

# 6. Introducing the Matchstick Model

The backbone of the model are the original Collatz rules. Different than in [01] or [02] where the original Collatz number system was transferred into an odd numbered system, this model will use the full range of natural numbers. As consequence, the chains are not limited at one end and we will have to deal with infinity now (Condition 6).

The Collatz number structure will be described as chains or fragments of Collatz number series. We will call such fragments matchsticks or matches. The elements x of a match m are:

$$x(m,i): \quad x(m, i+1) = C'[x(m, i)] \quad m: \{0, 1, 2, 3, ...\} \;\; i: \{0, 1, 2, 3, ...\} \qquad (1)$$

Where x is the index of the element and C' is the set of "inverse" Collatz operations. Because the Collatz number generator consisting of a set of the functions (3x+1) and (x/2) has no "inverse" since the results would be ambiguous and the graphical path branched. For that, C' contains some additional constraints to be defined later. The first 8 matchstick look like this:

m=0: (4) 8 16 **5 10 3** 6 12 24 48 96 ...
m=1: (10) 20 40 **13 26 52 17 34 11 22 7 14 28 9** 18 36 72 144 288 ...
m=2: (16) 32 64 **21** 42 84 168 336 672 ...
m=3: (22) 44 88 **29 58 19 38 76 25 50 100 33** 66 132 264 528 1056 ...
m=4: (28) 56 112 **37 74 148 49 98 196 65 130 43 86 172 57** 114 228 456 912 1824 ...
m=5: (34) 68 136 **45** 90 180 360 720 1440 ...
m=6: (40) 80 160 **53 106 35 70 23 46 15** 30 60 120 240 480 ...
m=7: (46) 92 184 **61 122 244 81** 162 324 648 1296 2592 ...
. ...

**Fig. 2:** The first 8 matchsticks of the Matchstick Model

Because a description of the additional constraints will be made later, we have to start at first with a simpler description of a matchstick starting at its head:

$$x(h,i): \quad x(m, i-1) = C[x(h, i)] \quad h: \{3, 9, 15, 21, ...\} \;\; i: \{1, 2, 3, ...\} \qquad (2)$$

Different than in [01] where the directions where called "Collatz-" or "tree-" direction, we are using now the expressions "downwards" if we mean propagation from head to foot, else "upwards". To get always unambiguous predecessors and successors within unbranched chains or unforked matchsticks, we need to check conditions and definition ranges by going downwards and upwards. This method allows introducing additional constraints exactly when needed and is intended to help designing the simplest possible model. Because all matchsticks have the same blueprint, by considering a single matchstick going downwards and upwards we will complete the design of the model and check the requirement list at the same time. If we do not encounter serious design difficulties or contradict with the conditions of the requirement list, we will end this walk with success.

## 6.1. The Algebra of the Matchstick: Downwards

The different sections of a matchstick are now considered in detail. Any matchstick is given by it's matchstick number m. A typical matchstick looks like this:

Match (m = 1): (10) 20 40 **13 26 52 17 34 11 22 7 14 28 9** 18 36 72 144 ...

Fig. 2. A typical matchstick with joining information (in brackets), foot, body (bold numbers) an infinite smoke trail. The first bold number is the proxy, the last bold number is the head.

Although all members of a matchstick can be determined by using the standard Collatz rules, we distinguish between different sections, from left to right, (see Fig. 2):

- join information, one even number (in bracket)
- foot section (2 even numbers)
- body section (bold, even or odd numbers), individual length
- infinite "smoke trail" section with only even numbers

The different sections and special elements will be considered now in detail:

### The head

It's the first odd number (from the right), and always divisible by 3, so it can be identified:

**Head identification:** $(h - 3) \bmod 6 = 0$ h: {(3, 9, 15, 21, ...} (3)

The head is the only element, from where all other elements can be determined using only the standard Collatz rules. Generally, any successors at the right side is ambiguous, but in case of the head, the even numbers at the right side of the head, the smoke trail is unbranched. Because all matchstick have at least their head and because a head is an odd number, there will be one gate at least in every matchstick. For that **Condition 5** of the requirement list is met.

### The smoke trail section

Because the head is divisible by 3, the members of the infinite number series of the smoke trail are now divisible by 6. It can easily be shown that in a standard Collatz number series, any predecessor of a number divisible by 3 can only be an even number, because the equation:

$$3 *(2n-1)+1 = 3m \qquad n, m: \{1, 2, 3, 4, ...\} \qquad (4)$$
$$2n - 2/3 = m$$

exceeds the definition range of natural numbers. Because of that, only one of the Collatz rules, namely the rule (x / 2) can be applied, but now exceptionally as bijective function with it's inverse function (2 * x). That's perfect because that means, that the smoke trail has no branching any more:

smoke trail: s: h*2 h*2^2 h*2^3 h*2^4 ... h*2^n h = head (5)

Any number within a Collatz number series, divisible by 6 (having the primes 2 and 3) has only one predecessor and only one successor. With having a smoke trail without terminals or gates, connected only to the head, the important **condition 4** of the model's requirements list can be met: prevention of connections at one side of the matchstick !

The smoke trail section is also perfect in another sense: Of all things, the infinitely long string of numbers has completely lost it's meaning and can optionally be left out completely for further treatment ... the smoke is gone and the **condition 6** of model's requirements list could also be fulfilled: mastering infinite long strings of numbers !

Compared with the methods used in [01] and [02] one can also explain why a proof of the Collatz conjecture is possible using only odd numbers at all: it's simply because the infinitely long even number series have no more branching and can be left out for consideration !

### The body section

The body of a matchstick holds one odd number at least. Generally it will contain odd and even numbers. The body is the only section with odd numbers. The calculation of the body members from the head to the foot can be done with the Collatz rules, but different than in a Collatz number series, the series have to be stopped two elements after the proxy number, an important number to cut the length of the matchsticks. The name "proxy" was adapted from [01] and [02], where it was used to mark the "first daughter" of any generation. Because now, we are using not only odd numbers in our model, the proxy marks the end of the body, but not the end of the matchstick. Nevertheless we can call it "stop condition" (for the body). It is given by the equation:

**Proxy stop condition:** $(x - 5) \bmod 8 = 0$  x: {1, 3, 5, ... }  (6)

This leads to odd numbered proxies like: proxies: {5, 13, 21, 29, 37, 45, ...}

### The foot section

The foot section contains two even numbers and starts after the body's last element, the proxy. The foot elements are labeled with f0,f1 and calculated (by using Collatz rules) with:

foot elements: $f1 = p / 2$  p: {16, 40, 64, 88, ...}  (7)
$f0 = f1 / 2$  f1: { 8, 20, 32, 44, ...}  (8)

Because the proxy stop condition is in play, when calculating matchstick elements from head to tail (downwards) and because we will have an analogous condition to calculate the elements from foot to head, we define the

### Matchstick Model downwards constraint

Within a matchstick, the body ends with the proxy stop. Two elements later, the Collatz number series is halted also. The calculation of one more element leads to the additional join information (see below).

There is a simple reason, why we needed the proxy stop condition: without we had run into the end loop 4,2,1,4... with all matches. The proxy stop condition is perfect, it allows to get the required terminal and the **condition 2** of the model's requirements list is met: the mandatory terminal !

### The join number

Is not a "physical" part of the matchstick but an information about where the match should be inserted during assembly. Although the calculation of matchstick elements is halted, we calculate on more element as important join information:

join number (from foot f0): $j := f0 / 2$  f0: {4, 10, 16, 22, ...}  (9)

The whole chain of elements for a matchstick has been calculated now. But there is more. Interestingly the proxy stop condition (6) seems to have the information about the end loop 4,2,1,4. With using (6), the lowest possible proxy is the 5. Using the Collatz functions to get j from p:

join number (from lowest proxy p): $j := (3 * 5 + 1) / 4 = \mathbf{4}$  (10)

the lowest possible join number for a matchstick is the 4. That's perfect again and makes sense, because the following elements will loop at 4, 2, 1, 4... and such a loop is not compatible with the required unilateral connection property given by the smoke trail of a matchstick.

For that, the proxy stop condition could be a fundamental rule. It not only marks the system limit of the unbranched Collatz tree section but also the correct limit of the "matchbox" definition range: the first (lowest) matchstick 0 has a join number (terminal) 4. Below, no more matchsticks can be defined and the gate of matchstick 0 can be found only in the endless loop 4,2,1,0...

The perfect match of the definition range of a matchstick could be a hint that the matchbox model will master all hurdles...

## 6.2 The Algebra of the Matchstick: Upwards

Although the model fits now the requirement 2, 3, and 4, it will be ambiguous within the body section. There is no active branching prevention in play. That can only be visible by walking the stick upwards now. For the first three elements, the calculation will not have any ambiguity because of the proxy stop condition leads to even numbers at the foot. The results calculation from the matchstick number m are:

| | | | |
|---|---|---|---|
| join number (from foot) j: | $j = 6m + 4$ | m: {0, 1, 2, 3, ...} | (11) |
| foot f0: | $f0 = 2 * j = 2 * (6m+4)$ | | (12) |
| foot f1: | $f1 = 4 * j = 4 * (6m+4)$ | | (13) |

After calculation of these even numbers, both "inverse" Collatz rules: (x –1) /3 or and (2x) could be applied, but only one is correct to have unbranched matchsticks. We have to select the former, because we know that the proxy has to be an odd number which satisfies the equation (6), so the proxy element, derived from j, or m is:

| | | |
|---|---|---|
| proxy p = f( j ): | $p = (4 * j -1) / 3$ | (14) |
| proxy p = f(m): | $p = ((4 * (6m + 4)) – 1) / 3 = 8m +5$ | (15) |

At least now when we enter into the body section, generally both Collatz rules will be used and we will have the expected ambiguity. To cut the branches we need the second constraint of the model:

**Matchstick Model upwards constraint**

The body of a matchstick is allowed to contain only fragments of Collatz number series with two consecutive even elements maximum between two odd elements.

Why this constraint works to prevent branching, can easily be shown, when we list a piece of a Collatz even number series with it's odd nodes. To prevent ambiguity, when walking upwards (see Fig. 3), one has to select always the odd numbered path to prevent more than two even numbers between the odd ones. The constraint leads to clarity.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Branches: | | 17 | | 69 | | 117 | |
| Collatz even number series: | 26 | 52 | 104 | 208 | 416 | 832 | ... |

**Fig. 3:** A piece of a Collatz even number series with it nodes.

With the additional Matchstick Model upward constraint, we can get all body elements with no ambiguity, so branching as well as generating (even number) sideways terminals are prevented. The **Condition 3** of the requirements list is met.

If we do not need to detect the head, we can continue the infinite calculation, else a head identification test (3) can be carried out. To call the matchbox model constraints "upwards" and "downwards" seems to be an appropriate naming convention but that does not mean, that only one constraint is in play in every direction. When walking on a matchstick there are always two constraints involved. When going upwards we already had used a "hidden" constraint, when we calculated the first foot element with (11) or (12), a kind of a foot filter or selection, and when we are going downwards we had used the "hidden" constraint (3), a kind of head filter or selection. The conditions are therefore strongly linked at least in pairs.

# 7. The Proof of The Collatz Conjecture

To prove the Collatz Conjecture is true, one has to show, that a Collatz number system is a composition of branches meeting the strict requirements list for being loop-free and unbranched. This could be shown in chapter 6 where the conditions 2 – 6 from the requirements list (chapter 4) were met. The Matchstick Model has also correctly shown, that condition 1 can't be part of the model and that it's definition range ends, when the number 4 was reached: the known loop can't be mapped into a matchstick, designed for unilateral connectivity. The essential end loop 4, 2, 1, 4 ... has to be added always at the beginning of an assembly, when the full Collatz number structure has to be shown.

Of course, no one will ever put the whole tree together, it would take an infinite amount of time and it would be quite confusing. But for a demonstration of the impossibility to build a loop with the matchsticks, one has to plug the matchstick 0 into the gate of the initial loop 4,2,1,4, ... at first, then the matchsticks 1, 2, 3, ... can be connected one after the other always considering the join number used to find the corresponding gate. Because every matchstick has one gate minimum but only one terminal, this terminal will always be absorbed, when connected and therefore loops can't be built with the Matchstick Model. It's worth emphasizing that the matches not only can be produced in their order (using the upwards algebra), but also can be assembled in the same order. This is practical, and an additional confirmation of the model, because every calculated match can be mounted immediately.

The method not only proved the Collatz conjecture but also explains, why a method considering only odd numbers [01] can lead to the same result. It's because the infinitely long strings of only even numbers, "the smoke trail", does not play any role for the proof.

The Matchstick Model has shown, that it is a perfect representation of the original Collatz number structure and tailored at the only allowed size given by a lonely model constraint used for each of Collatz number series propagation direction. In other words: the Collatz number structure can be mastered by the most primitive mathematical tool: a simple constraint limiting a branched structure.

# 8. Literature


[01] FRANZ ZIEGLER: Proof of the Collatz conjecture V 1.2
www.collatz.com/proof_of_collatz.pdf

[02] FRANZ ZIEGLER: A Proof of the Collatz conjecture by Synthesis V 6.7
www.collatz.com/collatz_proof_by_synthesis.pdf

[03] FRANZ ZIEGLER: Understanding the Proof of the Collatz conjecture V1.0
www.collatz.com/understanding_collatz.pdf