A solution to Collatz's conjecture

Ignacio San José

Instituto Aragonés de Estadística
Bernardino Ramazzini, 5, 50015 Zaragoza, Spain
E-mail : isanjose@aragon.es

Abstract.

A simple and innovative approach is presented to solve Collatz's conjecture, based on an equivalence relation and the derivation of equivalence classes. Therefore, it is demonstrated that the union of all equivalence classes forms the set of odd numbers.

Keywords: Collatz conjecture; equivalence classes.
MSC2020: 11B99, 65Q30

1 Introduction

The Collatz's conjecture, also known as the $3\,x+1$ problem, is a famous unsolved problem in mathematics which consists of the following: if a positive integer x is even, it is divided by two. If it is odd, it is multiplied by three and one is added. This operation can be expressed as:

$$f(x) = \begin{cases} x/2 & x \equiv 0 \ (mod\ 2) \\ 3\,x+1 & x \equiv 1 \ (mod\ 2) \end{cases}$$

In this way, by applying the function $f(x)\colon \mathbb{N}+1 \to \mathbb{N}+1$, with $\mathbb{N} \coloneqq \{0,1,,2\ldots\}$, repeatedly to any positive integer, we always reach the number 1, regardless of the initial chosen number, and in a finite number of steps. Let $n \in \mathbb{N}+1$, we use the definition of Collatz orbit given in [1], $Col^{\mathbb{N}}(n) \coloneqq \{n, f(n), f^2(n), \ldots\}$. Let $Col_{min}(n) \coloneqq min\ Col^{\mathbb{N}}(n) = inf_{k\in\mathbb{N}} Col^k(n)$, then the Collatz conjecture is expressed as $Col_{min}(n) = 1$ for every $n \in \mathbb{N}+1$.

If the previous statement were false, it would mean that there exists a number $m \in \mathbb{N}+1$ that generates a cycle where the number 1 does not belong to its cycle, i.e., $Col_{min}(m) \neq 1$. This would imply that the sequence enters a cycle that does not contain the number 1. Furthermore, the cycle could potentially increase indefinitely.

For example, if we start with $x = 7$ and apply the function repeatedly, we obtain the following sequence of numbers until reaching 1: $7 \to 22 \to 11 \to 34 \to 17 \to 52 \to 26 \to 13 \to 40 \to 20 \to 10 \to 5 \to 16 \to 8 \to 4 \to 2 \to 1$. Thus, its orbit would be $Col^{\mathbb{N}}(7) = \{7,\ 22,\ 11,\ 34,\ 17, 52,\ 26,\ 13,\ 40,\ 20,\ 10,\ 5,\ 16,\ 8,\ 4,\ 2,\ 1\}$.

In [1], the function $Aff_n\colon \mathbb{R} \to \mathbb{R}$ is defined such that for every positive integer $n \in \mathbb{N}\backslash\{0\}$, we have $Aff_n(x) = \frac{3x+1}{2^n}$, where n is the largest natural number such that 2^n divides $3x+1$, resulting in an odd number if x is odd. In this paper, we will always consider x

to be an odd number. Thus, we can define the Collatz orbit for this new function as follows: $Col^{2\mathbb{N}+1}(x) := \left\{x, Aff_{n_1}(x),\ Aff_{n_1}\left(Aff_{n_2}(x)\right), \dots\right\}$. This corresponds to the formulation known as Syracuse, so the Collatz conjecture can be expressed as $min\ Col^{2\mathbb{N}+1}(x) = 1$ for every $x \in 2\mathbb{N}+1$. This Syracuse orbit $Col^{2\mathbb{N}+1}(x)$ corresponds to the Collatz orbit but for odd numbers. Thus, in the previous example, we obtain the sequence $Col^{2\mathbb{N}+1}(7) = \{7, 11, 17, 13, 5, 1\}$.

From the direct observation of the odd numbers that reach 1 after applying the function only once Aff_n, it is observed that there is recursion between them. For example, consider the odd numbers $x = 1, 5, 21, 85, 341, \dots$ then $5 = 4 \times 1 + 1, 21 = 4 \times 5 + 1, 85 = 4 \times 21 + 1, 341 = 4 \times 85 + 1$, in general $x_{k+1} = 4x_k + 1$, obtaining the general term of the series as $4^k + \frac{1}{3}(4^k - 1) \in 2\mathbb{N}+1$.

This fact reveals a structure or pattern that is generated by the repeated application of the function Aff_n. Clearly, the numbers in this series satisfy that $Aff_{2(k+1)}\left(4^k + \frac{1}{3}(4^k - 1)\right) = 1$.

Due to this, we can consider that the series of nodes thus defined constitute a series of nodes belonging to the same branch. See figure 1.

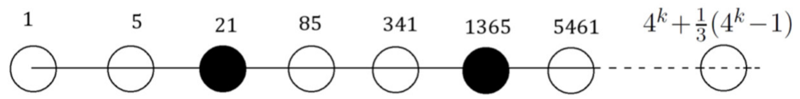

Figure 1. Representation of a branch with its nodes. Nodes that are multiples of three are marked with a black circle.

Thus, the distribution of odd numbers is ordered in the branch according to the number of 2^{2k} required to belong to it. For example, for $Aff_2(1) = 1$ with 2^2, $Aff_4(5) = 1$ with 2^4, $Aff_6(21) = 1$ with 2^6, $Aff_8(85) = 1$ with 2^8, so the number 2^{2k} varies but it is clear that $1, 5,\ 21,\ 85$ reach 1 in one step (a single application of Aff_n), hence they belong to the same branch.

In general and as a consequence of this result, we can consider all numbers $x, y\ \in 2\mathbb{N}+1$ that satisfy the following equation $Aff_{n_x}(x) = Aff_{n_y}(y)$, which means that $\frac{3x+1}{2^{n_x}} = \frac{3y+1}{2^{n_y}}$. If we solve for x, we obtain $x = 2^{n_x - n_y}\ y + \frac{1}{3}(2^{n_x - n_y} - 1)$. Since x and y are odd numbers, it can be deduced that $n_x - n_y$ must be a natural and even number, that is, $n_x - n_y = 2\ k$ with $k \in \mathbb{Z}$. Rewriting it, we obtain that $x = 2^{2k}\ y + \frac{1}{3}(2^{2k} - 1)$. The sign of $|k|$ will be determined by the relationship between x and y. If $x < y$, then this implies that $k > 0$. On the other hand, if $x > y$ then necessarily $k < 0$ since both x and y have to belong to $2\mathbb{N}+1$. For example, if $x = 1$ and $y = 5$ we have that $1 = 2^{-2}\ 5 + \frac{1}{3}(2^{-2} - 1)$, in this case $k = -1$.

In this way, we can establish the following relation: let $x, y \in 2\mathbb{N}+1$ be given. Then, $x \sim y$ if $Aff_{n_x}(x) = Aff_{n_y}(y)$ for suitable n_x and n_y. This relation is one of equivalence relation since it is reflexive, symmetric and transitive, and its proof is trivial. This equivalence class allows us to regroup the elements into equivalence classes and consider subsets or equivalence classes accordingly.

2 Nodes and Branches

Definition 2.1. The equivalence classes, which we will call branches, are defined as $b(m) := \left\{x \in 2\mathbb{N}+1 \mid x = 4^k m + \frac{1}{3}(4^k - 1),\ k \in \mathbb{Z}\right\}$. And the number $m \in 2\mathbb{N}+1$ as the initial node of the branch.

Clearly the equivalence relation holds for any $x \in b(m)$, since $Aff_n(x) = Aff_n\left(4^k m + \frac{1}{3}(4^k - 1)\right) = Aff_{n-2k}(m)$.

As mentioned earlier, the initial node of the equivalence class can be any element from it. The only thing to consider is the sign of $|k|$ so that all members of the class belong to $2\mathbb{N}+1$.

Proposition 2.1. Let $m \in b(n)$ then $n \in b(m)$ and vice versa. Therefore, $b(n) = b(m)$.

Proof. Assuming that $m \in b(n)$, then $m = 4^k n + \frac{1}{3}(4^k - 1)$, so if $k > 0$, solving for n, $n = 4^{-k} m + \frac{1}{3}(4^{-k} - 1)$ thus $n \in b(m)$. The same reasoning applies for $k < 0$. Therefore, $b(n) \subseteq b(m)$. By exchanging n with m, the other inclusion is demonstrated. □

For example, let $m \in b(1)$ then we have that 1 is the initial node of the branch and thus $m = 4^k + \frac{1}{3}(4^k - 1)$ and k could be $k = 0,1,2 \ldots$ such that $m \in 2\mathbb{N}+1$ and so $b(1) = \{1,\ 5,\ 21,\ 85,\ 341,\ \ldots\}$. If we consider the element 85 as the initial node, we obtain the same branch. Let $m \in b(85)$ then $m = 4^k\, 85 + \frac{1}{3}(4^k - 1)$ and k could be $k = -3, -2, -1, 0, 1, 2 \ldots$ and so $b(85) = \{1,\ 5,\ 21,\ 85,\ 341,\ \ldots\}$.

Node 1 is the only node that belongs to its branch since $Aff_2(1) = 1$. For any other node this does not occur, for example for $b(3) = \{3, 13, 53, 213, \ldots\}$ and $Aff_1(3) = 5$, that is, it goes down to the branch of 1. See figure 2.

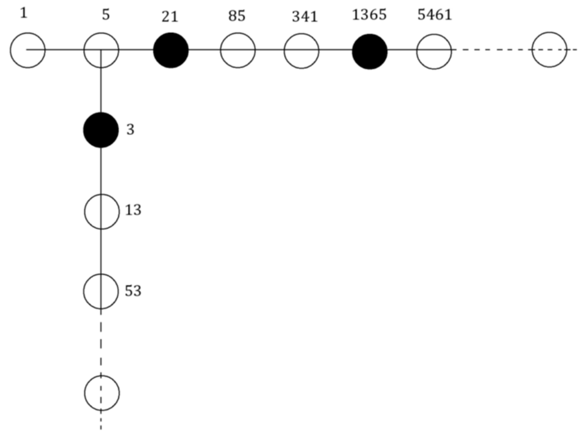

Figure 2. Representation of two branches with its nodes.

Proposition 2.2. The only initial node that belongs to its branch is 1 node, that is, let $x \in b(m)$ if $Aff_n(x) = m$ for some $n \in \mathbb{N}\backslash\{0\}$ then $m = 1$.

Proof. Just apply the function once Aff_n, $Aff_n(x) = Aff_n\left(4^k m + \frac{1}{3}(4^k - 1)\right) = \frac{3\left(4^k m + \frac{1}{3}(4^k-1)\right)+1}{2^n} = \frac{4^k(3\,m+1)}{2^n} = \frac{3\,m+1}{2^{n-2k}} = m$, then $m = \frac{1}{2^{n-2k}-3}$ and since it has to be an odd number and a positive integer there is only one solution that corresponds to $n - 2k = 2$, that is, $n = 2k + 2$ and therefore $m = 1$, thus $x \in b(1)$ and since $1 = 4^0 + \frac{1}{3}(4^0 - 1)$ then $1 \in b(1)$ with $k = 0$. □

When applying the function Aff_n to any number other than 1, we obtain another number that can be considered one step closer to the branch that contains the number 1. In other words, when applying the function Aff_n, it switches from one branch to another and the latter will be closer to the node 1. Similarly, we can think that the inverse function of Aff_n applied to a node results in a node that is one step further away from the node 1. Therefore, we can consider the inverse function of Aff_n defined as $Aff_n^{-1}: 2\mathbb{N} + 1 \to 2\mathbb{N} + 1$ such that $Aff_n^{-1}(m) := \frac{2^n\,m-1}{3}$.

Proposition 2.3. For every node m not divisible by three, belonging to any branch, another branch comes out with infinite nodes, that is, if $m \equiv 1\ (mod\ 3)$ or $m \equiv 2\ (mod\ 3)$ there exists $p \in 2\mathbb{N} + 1$ such that $Aff_n(p) = m$, equivalent $Aff_n^{-1}(m) = p$. And if $m \equiv 0\ (mod\ 3)$ no branches come from those nodes.

Proof. Let m be any node, consider the only three possible situations

1) Suppose that $m \equiv 0\,(mod\ 3)$, then let p be its superior node, that is $p = Aff_n^{-1}(m) = \frac{2^n\,m-1}{3}$ then according to the remainder of p, we have that

   a. If $p \equiv 0\ (mod\ 3)$, that is $p = 3p_0$, with $p_0 \in \mathbb{N}$, then $3p_0 = \frac{2^n\,m-1}{3}$, which simplifies to $2^n\,m = 3^2 p_0 + 1$; in order words, $2^n\,m \equiv 1\ (mod\ 3)$, which is not possible.

b. If $p \equiv 1\ (mod\ 3)$, that is $p = 3p_0 + 1$, with $p_0 \in \mathbb{N}$, then $3p_0 + 1 = \frac{2^n\, m-1}{3}$, which simplifies to $2^n\, m = 3\ (3p_0 + 1)+$; in order words $2^n\, m \equiv 1\ (mod\ 3)$, which is not possible.
c. If $p \equiv 2\ (mod\ 3)$, that is $p = 3p_0 + 2$, with $p_0 \in \mathbb{N}$, then $3p_0 + 2 = \frac{2^n\, m-1}{3}$, which simplifies to $2^n\, m = 3\ (3p_0 + 2) + 1$; in order words, $2^n\, m \equiv 1\ (mod\ 3)$, which is not possible.

This make clear that if $m \equiv 0\ (mod\ 3)$, it does not have a superior node.

2) Suppose $m \equiv 1\ (mod\ 3)$, then let p be its superior node, which means $p = Aff_n^{-1}(m) = \frac{2^n\, m-1}{3}$. Since $m \equiv 1\ (mod\ 3)$, we can write $m = 3m_0 +\ 1$, with $m_0 \in \mathbb{N}$. Therefore, $p = \frac{2^n(3m_0 + 1) -1}{3} = \frac{3\ 2^n m_0 + 2^n -1}{3} = 2^n m_0 + \frac{2^n -1}{3} \in 2\mathbb{N} + 1$ if $n \in 2\mathbb{N}$, as in that case $\frac{2^n -1}{3}$ will be an odd and integer number. Hence, there exists $p\ \in 2\mathbb{N} + 1$. As p depends on n, there are infinitely many values of p.
3) Suppose $m \equiv 2\ (mod\ 3)$, then let p be its superior node, which means $p = Aff_n^{-1}(m) = \frac{2^n\, m-1}{3}$. Since $m \equiv 2\ (mod\ 3)$, we can write $m = 3m_0 +\ 2$, with $m_0 \in \mathbb{N}$. Therefore, $p = \frac{2^n(3m_0 + 2) -1}{3} = \frac{3\ 2^n m_0 + 2^{n+1} -1}{3} = 2^n m_0 + \frac{2^{n+1} -1}{3} \in 2\mathbb{N} + 1$ if $n \in 2\mathbb{N} + 1$, as in that case $\frac{2^{n+1} -1}{3}$ will be an odd and integer number. Hence, there exists $p\ \in 2\mathbb{N} + 1$. As p depends on n, there are infinitely many values of p. □

In conclusion, given $m \equiv 1\ (mod\ 3)$, a branch is formed defined by the nodes $p = \frac{2^{2k}\, m-1}{3}$ with $k = 1, 2, \ldots$. The initial node corresponds to $k = 1$, $p = \frac{2^2\, m-1}{3}$. In other words, we have the branch $b\left(\frac{2^2\, m-1}{3}\right)$ And in the case of $m \equiv 2\ (mod\ 3)$, a branch is formed defined by the nodes $p = \frac{2^{2k+1}\, m-1}{3}$ with $k = 0, 1, 2, \ldots$. The initial node corresponds to $k = 0$, $p = \frac{2\, m-1}{3}$. In other words, we have the branch $b\left(\frac{2m-1}{3}\right)$.

For example, let's assume $m = 341$. Since $341 \equiv 2\ (mod\ 3)$, then $p = \frac{2^{2k+1}\, m-1}{3}$. If $k = 0$ then $p = 227$. If $k = 1$ then $p = 909$, graphically, see figure 3.

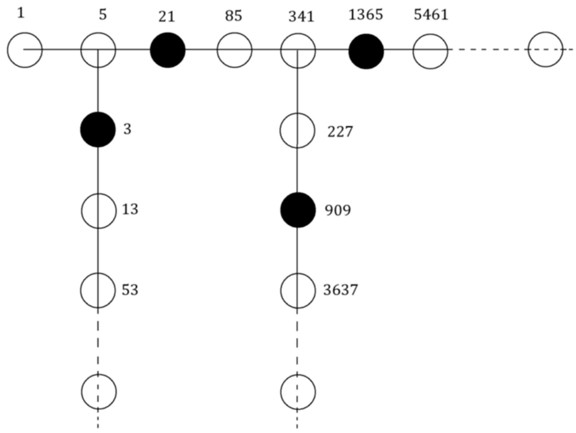

Figure 3. Representation of three branches with its nodes.

In both cases it is verified that 341 are in the orbit $Col^{2\mathbb{N}+1}(227) = \{227, 341, 1\}$ and $Col^{2\mathbb{N}+1}(909) = \{909, 341, 1\}$.

Each node m is not a multiple of 3 ($m \not\equiv 0 \ (mod\ 3)$) generates a distinct equivalence class (branch) and each branch has its own nodes.

Proposition 2.4. Let $\mathrm{n, m} \in 2\mathbb{N}+1$ consider the distinct branches $\mathrm{b(m)}$ and $\mathrm{b(n)}$ with $\mathrm{m} \notin \mathrm{b(n)}$ then $\mathrm{b(m)} \cap \mathrm{b(n)} = \emptyset$.

Proof. Assume that there exists an odd number $x \in b(m) \cap b(n)$, then $x = 4^k m + \frac{1}{3}(4^k - 1) = 4^l n + \frac{1}{3}(4^l - 1)$ with $k \neq l$; simplifying $2^{2k} m = 2^{2l} n + \frac{1}{3}(2^{2l} - 2^{2k})$ and so $m = 2^{2l-2k} n + \frac{1}{3}(2^{2l-2k} - 1)$; which implies that $m \in b(n)$ which cannot be by hypothesis, therefore there is no x that belongs to the intersection.□

3 Tree. Conclusion

As indicated earlier, we can consider a tree-shaped like structure with the initial node being 1. In other words, the main branch can be assigned to the one that contains the initial node 1, with the trunk of the tree where for each node not divisible by 3, a branch with infinite nodes emerges. Each branch represents an equivalence class, and the union of all equivalence classes forms the set of odd numbers.

Lemma 3.1. $\bigcup_{m \in 2\mathbb{N}+1} b(m) = 2\mathbb{N}+1$.

Proof. The inclusion $\bigcup_{m \in 2\mathbb{N}+1} b(m) \subseteq 2\mathbb{N}+1$ is obvious from the very definition of $b(m)$. Let us see that $\bigcup_{m \in 2\mathbb{N}+1} b(m) \supseteq 2\mathbb{N}+1$. Let $n \in 2\mathbb{N}+1$, consider $m = 4n + 1$, then $n = 4^k m + \frac{1}{3}(4^k - 1)$ therefore $n \in b(m)$, with $k = -1$. □

In this way, a set of infinite branches with infinitely many nodes is obtained. Nodes only belong to one branch by proposition 2.4. In addition, the branches are interconnected by the initial nodes of each branch by proposition 2.3 and the only initial node that is in its own branch is 1, which shows that there are no unconnected branches or unconnected nodes that generate another independent tree and this is because the union of all branches is the set of odd numbers. Furthermore, it is clear that the orbit of any odd number indicates the passage from one branch to another, each time the Aff_n function is applied, it jumps from one branch to a lower one. This indicates that every orbit $Col^{2\mathbb{N}+1}(m)$ contains the number 1, thus fulfilling the conjeture.

For example, consider the node $n = 1643861$; this node belongs to the branch of $b(401)$ since $1643861 = 4^k\ 401 + \frac{1}{3}(4^k - 1)$ with $\mathrm{k} = 6$. The branch of $b(401)$ starts at node 301, because $Aff_2(401) = 301$. This node belongs to branch $b(75)$, this branch starts from node 113 and finally, node 113 starts from node 85 which is in the main branch. Its orbit is $Col^{2\mathbb{N}+1}(1643861) = \{1643861, 301, 113, 85, 1\}$. See figure 4.

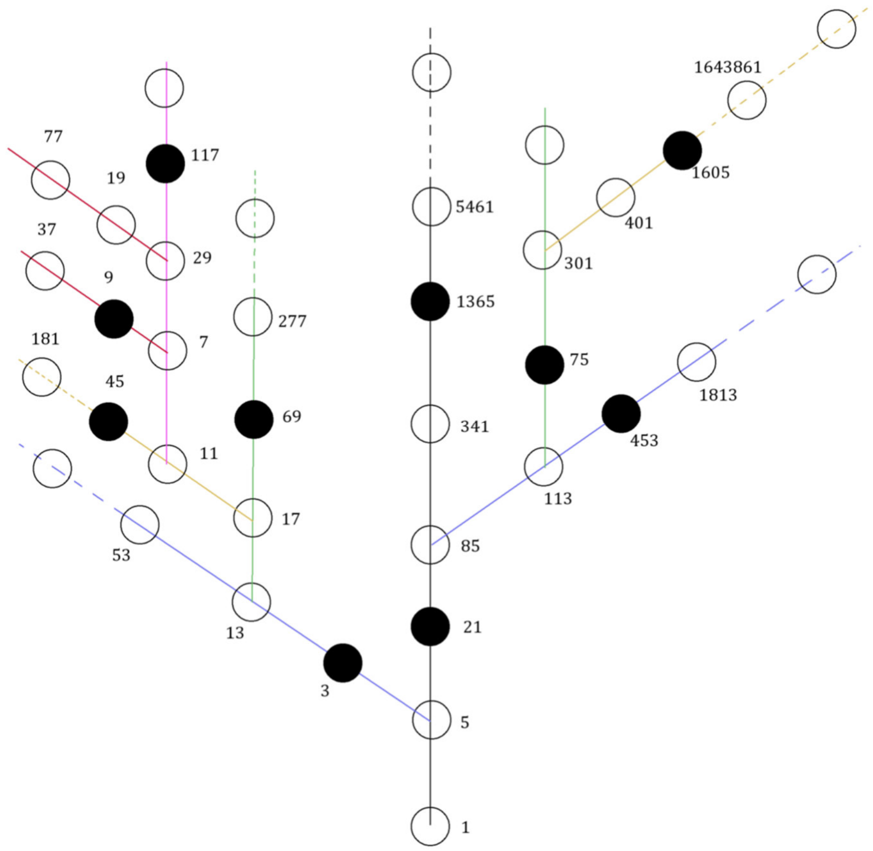

Figure 4. Representation of a part of the tree where node 1643861 and the lower branches appears

Conflict of Interest

The author declare no conflict of interest.

References

[1] T. Tao, Almost all orbits of the Collatz map attain almost bounded values, Forum of Mathematics, Pi 10 (2022).