Boolean Hypercubes, Mersenne Numbers, and the Collatz Conjecture

Ramon Carbó-Dorca

Institut de Química Computacional i Catàlisi / Universitat de Girona / Girona 17003 (Catalonia) Spain

Article Info

Abstract

This study is based on the trivial transcription of the vertices of a Boolean N-Dimensional Hypercube $\mathbf{H}_N$ into a subset $\mathbb{S}_N$ of the decimal natural numbers $\mathbb{N}$. Such straightforward mathematical manipulation permits to achieve a recursive construction of the whole set $\mathbb{N}$. In this proposed scheme, the Mersenne numbers act as upper bounds of the iterative building of $\mathbb{S}_N$. The paper begins with a general description of the Collatz or $(3x+1)$ algorithm presented in the $\mathbb{S}_N \subset \mathbb{N}$ iterative environment. Application of a defined *ad hoc* Collatz operator to the Boolean Hypercube recursive partition of $\mathbb{N}$, permits to find some hints of the behavior of natural numbers under the $(3x+1)$ algorithm, and finally to provide a scheme of the Collatz conjecture partial resolution by induction.

1. Introduction

The Collatz or $(3x+1)$ conjecture, (see the work of Lagarias [1] for a comprehensive review and reference [2] for a decade-old discussion of this problem by several authors), seems that it still has to be proven true since its description around the first third of the last century. More information on the Collatz conjecture can be easily obtained from reference [3]. Recent advances towards the problem solution can be found thanks to the insight of Tao in reference [4]. More information on the so-called Collatz or $(3x+1)$conjecture can be found in references contained within the previous publications and in the references [5, 6].

Some previous work on the application of N-Dimensional Boolean Hypercubes, $\mathbf{H}_N$, has been carried on by the present author. For instance, to describe some inner structure of vector spaces [7], and discuss several mathematical, physical, biological, and chemical problems, which can be found in references [8]-[22]. The study of computational multinomial combinatorics for colorings of Hypercubes for all irreducible representations with applications to non-rigid molecules has been recently published by Balasubramanian [23, 24].

In the present paper, Boolean Hypercubes will be chosen as the cornerstone tool to study the Collatz conjecture, because apparently, no discussion of this old problem includes the use of the *iterative construction* of natural numbers in terms of the sequence of decimal numbers, associated to the 2^N binary vertices of an N-Dimensional Boolean Hypercube $\mathbf{H}_N$.

It is intriguing that, at least as far as the author knows, no such simple computational-geometric structure has been considered to approach the Collatz conjecture.

The scheme which this paper will follow consists of the following parts: 1) Firstly, we shall provide the definitions of an appropriate Collatz algorithm and an *ad hoc* operator able to be employed in the development of a proof of the $(3x+1)$ problem. 2) Next, it will be discussed the recursive construction of Boolean Hypercubes and their translation into an iterative buildup of natural numbers. 3) Then, a section will be devoted to a computational study of the Mersenne numbers from the Collatz algorithm perspective. 4) After this, some extensions of Mersenne numbers will be studied, and the characteristic behavior of them in front of the $(3x+1)$ problem described. 5) Finally, will be studied the Collatz conjecture through the perspective of an inductive procedure.

2. The collatz algorithm

It is straightforward to describe the problem associated to the Collatz conjecture. Provided the set of natural numbers:$\mathbb{N}$, then one can expose any natural number to the simple algorithm:

Algorithm 1: Collatz or $(3x+1)$ **procedure**

> $\forall m \in \mathbb{N};$
> $I = 0;$
> $Define\ C_I[m];$
> while $m > 1;$
> > $I = I+1;\ c \leftarrow m/2;$
> > if $2*c \neq m:\ m \leftarrow 3m+1;$ else $:\ m \leftarrow c;$

When using the Collatz algorithm above, the variable I indicates the number of times (here they will be called Collatz *iterations*) that the algorithm has been applied until termination is reached, resulting in reaching the natural unit, bearing a Collatz iteration number $I = end$. When such termination occurs, a so-called Collatz *orbit* has been completed.

The name orbit possibly comes from the generalization of the Collatz conjecture with the chance that the transformation $(px+1)$, being p a prime number, is used instead of the usual $(3x+1)$ transform. While in this $(3x+1)$ study no never-ending divergence nor loop has been found, a simple test with $p = 5 \wedge x = 5$ yields a closed loop.

Because the sequence of natural numbers, provided until a $(3x+1)$ Collatz iteration completion is reached, has not a closed structure, perhaps better than naming such a sequence as an orbit, one might call it a *complete path*. Also, by a *partial path,* one can understand any non-terminal sequence of natural numbers originated by the Collatz **Algorithm 1**. One can simply refer to a *path* if the context makes it well-defined, not needing more information.

For example, the complete path, made of seven iterations, generated by Collatz **Algorithm 1** applied to the natural number 3 becomes:

$$\begin{aligned} Path&: \quad \{3 \to 10 \to 5 \to 16 \to 8 \to 4 \to 2 \to 1\} \\ Iteration&: [0 \cdots 1 \cdots 2 \cdots 3 \cdots 4 \cdots 5 \cdots 6 \cdots 7] \end{aligned} \quad, \tag{2.1}$$

which can be also imagined as a graph, whose vertices are constituted by the Collatz sequence of natural numbers generated from the application of **Algorithm 1**.

2.1. Collatz conjecture

The final stage or the path completion associated with the essence of the Collatz conjecture, affecting the whole set of natural numbers $\mathbb{N}$, might be expressed as:

$$\forall m \in \mathbb{N} : C_{end}[m] = 1. \tag{2.2}$$

Also, any natural number n fulfilling:

$$n \in \mathbb{N} \wedge C_{end}[n] = 1$$

will be called Collatz *compliant*.

3. Development of the collatz operator

Collatz **Algorithm 1** as defined above might be supposedly constructed by a non-linear operator, which conducts a Collatz path flux from a starting natural number down to the natural unit. Thus, to discuss and study the behavior of the Collatz compliance of natural numbers it is interesting to design the structure of a Collatz operator, as follows.

3.1. Definitions

One can formally use the Collatz operator as a symbolic structure, allowing to represent the path evolution under **Algorithm 1** applied to any natural number. It can be obtained via an indeterminate number of **Algorithm 1** iterations I. For instance, a partial path of the Collatz operator could be written like:

$$\forall m \in \mathbb{N} : C_I[m] = m_I.$$

Moreover, using the equation (2.2), then any complete path generated by the Collatz operator can be easily described.

To grasp the meaning of the previous notation which will be used from now on, one might suppose that the basic natural numbers $\{0, 1\}$ are also Collatz compliant, or:

$$C_{end}[0] = 1 \wedge C_{end}[1] = 1.$$

3.2. Extension to subsets of natural numbers

If a natural number subset $\mathbb{S}$ of cardinality: $Card\,[\mathbb{S}] = N$ is defined as: $\mathbb{S} \subset \mathbb{N}$, then one can define the partial path application of the Collatz algorithm over such a natural number subset, using as much as I iterations, noting:

$$C_I[\mathbb{S}] = \mathbb{S}_I$$

and meaning that:

$$\forall s \in \mathbb{S} : C_I[s] = s_I \in \mathbb{S}_I.$$

The application of the Collatz algorithm over a natural number subset as above defined, until the outcome of a Collatz complete path is reached, could be symbolically expressed with:

$$C_{\langle end|}[\mathbb{S}] = \langle 1_N|,$$

where the row vector on the right is defined as the unity vector of dimension N, that is:

$$\langle 1_N | = (1, 1, 1, ..., 1).$$

Therefore, the Collatz conjecture involving the whole natural set can be easily expressed using alternatively the equation:

$$C_{\langle end|} [\mathbb{N}] = \langle 1_\infty |. \tag{3.1}$$

Note that in this case, the iteration subindex indicator: $\langle end|$ must be considered as a vector of dimension N, bearing everyone of ending number of iterations of each element of the set S, upon application of Collatz algorithm. In the case of equation (3.1), the $\langle end|$ vector has to be supposed of ∞ dimension.

3.3. Application to natural vector spaces

This scheme, as above described, suggests that if a natural vector space of dimension N: $\mathbb{V}_N(\mathbb{N})$ is defined (see for more details references [8,9]) then one can also write an alternative to the above-defined symbolic application of the Collatz algorithm, like:

$$\forall \langle s| \in \mathbb{V}_N(\mathbb{N}) : C_{\langle end|} [\langle s|] = \langle 1_N |;$$

additionally, considering the natural number set as an infinite-dimensional row perfect vector: $\langle \mathbb{N}| \in \mathbb{V}_\infty(\mathbb{N})$, then the Collatz conjecture could be also rewritten as:

$$C_{\langle end|} [\langle \mathbb{N}|] = \langle 1_\infty |.$$

3.4. A trivial property

Here it will be considered an important but trivial property of the symbolic application of the Collatz operator. Suppose now that the Collatz operator is applied to a natural number, in such a way that after some number of iterations I the result of the generated partial path appears to be:

$$\exists n \in \mathbb{N} : C_I [n] = m_I \in \mathbb{N}.$$

If the resulting natural number m_I is *a priori* known to be Collatz compliant, then necessarily the initial number n is Collatz compliant. Because one can write:

$$C_{end} [m_I] = 1 \Rightarrow C_{end} [n] = 1.$$

Such an obvious consequence of the definition of the Collatz operator will be of further use to discuss a possible Collatz conjecture incomplete demonstration track.

3.5. Powers of 2 as collatz compliant natural numbers

When observing **Algorithm 1**, it is immediate to consider the powers of 2 as a trivial Collatz compliant set of natural numbers, and thus one can write:

$$\forall N \in \mathbb{N} : C_N \left[2^N \right] = 1. \tag{3.2}$$

Hence, because of the trivial property discussed in the previous paragraph 3.4, whenever in a Collatz path development, attached to a specific natural number, there appears a power of 2 in the path sequence, then such an initial natural number is Collatz compliant:

$$\exists P, m, N \in \mathbb{N} : C_P [m] = 2^N \Rightarrow C_{end} [m] = 1.$$

The path of number 3, as shown in the equation (2.1) is an example of this situation, as at the third Collatz iteration, $2^4 = 16$ is reached. Powers of 2 transcribed into binary vectors have a leading role in the definition of the vertices of Boolean Hypercubes. Indeed, they correspond to a binary *canonical* basis set, see for instance references [10-22], where every vector has one bit 1 in some position of the Boolean vertex vector and the rest of elements are made by 0 bits.

3.6. Forbidden partial paths

It appears that some vertices of Collatz partial paths cannot be present if Collatz conjecture in equations (2.2) or (3.1) is accepted as true. For instance, suppose that:

$$\exists m, n, I, J \in \mathbb{N} : C_I [m] = n \wedge C_{I+J} [m] = n \Rightarrow \tilde{\exists} C_{end} [m]$$

Then, if this is the case, **Algorithm 1** will enter an unending loop, and Collatz conjecture will be false.
To avoid this kind of situation, if the Collatz conjecture has to be true, the natural number sequence of a complete path generated by **Algorithm 1** must be made in any case by completely *different* natural elements.

4. *N*-dimensional boolean hypercubes and natural number sequences

The concatenation of Boolean Hypercubes (see for more information references [9]-[22]) constitute the fundament where a possible demonstration of Collatz conjecture will be built. To briefly introduce the useful structure of Boolean Hypercubes, one can start choosing a set of natural numbers of cardinality 2^N: $\mathbb{S}_N$. This set might be constructed as the ordered sequence of elements of the natural number subset, which can be associated with the decimal transcription of the set of 2^N vertices of an *N*-Dimensional Boolean Hypercube: $\mathbf{H}_N$.

Simply, the set $\mathbb{S}_N$ corresponds to the natural number subset:

$$\mathbb{S}_N = \left\{ m = 0, 2^N - 1 : \Delta = 1 \right\} = \left\{ 0, 1, ..., 2^N - 1 \right\},$$

where the last number in the set $\mathbb{S}_N$ corresponds to the so-called *Mersenne number*: $\mu_2(N) = 2^N - 1$.

At the same time, and generally speaking, any decimal Mersenne number is transcribed into a Boolean vector, a bit string, constituting a characteristic vertex of an *N*-Dimensional Boolean Hypercube $\mathbf{H}_N$, the *binary unity vector* of the adequate dimension: $\langle 1_N | = (1, 1, ..., 1)$ made with all its elements as the 1 bit. It can be called the Mersenne vertex of the Boolean Hypercube and might be considered as the farthest vertex from the Boolean zero vertex: $\langle 0_N | = (0, 0, ..., 0)$.

The most interesting property of this kind of partition of the natural numbers is the trait by which the natural number sequences $\mathbb{S}_N$ resemble a fractal structure. Indeed, the decimal translation of the 2^{N+1} vertices of an (*N*+1)-Dimensional Boolean Hypercube $\mathbf{H}_{N+1}$ can be calculated iteratively, once the sequence $\mathbb{S}_N$, attached in turn to a previous *N*-Dimensional Boolean Hypercube $\mathbf{H}_N$ is known. This extension can be obtained just by adding to each member of $\mathbb{S}_N$ the *N*-th power of two: 2^N.

That is, the above-mentioned recursion involving a Boolean Hypercube structure can be easily written with the following Algorithm:

Algorithm 2: Knowing the Initial Natural Number Subset $\mathbb{S}_N$ Iterate Towards the Natural Number Sequence $\mathbb{S}_{N+1}$

$$\begin{aligned}
\mathbb{S}_{N+1} &= \mathbb{S}_N \cup \left[2^N \oplus \mathbb{S}_N \right] \equiv \mathbb{S}_N \cup \mathbb{A}_{N+1} \\
&\equiv \left\{ 0, 1, ..., 2^N - 1 \right\} \cup \left\{ 2^N, 2^N + 1, ..., 2^{N+1} - 1 \right\} \\
&= \left\{ 0, 1, ..., 2^N - 1, 2^N, ..., 2^{N+1} - 1 \right\}
\end{aligned} \tag{4.1}$$

In **Algorithm 2** the summation symbol $\oplus$ means that the power 2^N has to be summed to every element of the set $\mathbb{S}_N$. **Algorithm 2** above is also revealing a convenient *iterative construction* of natural numbers. When stepping one dimension up from a known *N*-Dimensional Boolean Hypercube $\mathbf{H}_N$, obtaining $\mathbf{H}_{N+1}$, the new decimal translation of his bit string vertices can be collected in the set: $\mathbb{S}_{N+1}$. Then the set $\mathbb{S}_{N+1}$ corresponds to the associated natural numbers, which are the decimal transcription of the Boolean vertices of the new (*N* + 1)-Dimensional Boolean Hypercube $\mathbf{H}_{N+1}$ in the upgraded sequence.

The natural set $\mathbb{S}_{N+1}$ also can be expressed as the union of two parts. The first one is coincident with the old $\mathbb{S}_N$ sequence, associated with the entire set of vertices of the *N*-Dimensional Boolean Hypercube $\mathbf{H}_N$. The second part is the set, which could be called: $\mathbb{A}_{N+1}$, say, just corresponding to the old set $\mathbb{S}_N$, that to each of its elements has been added the value of 2^N.

The set $\mathbb{A}_{N+1}$ characterizes the decimal transcription of half the set of vertices of the (*N* + 1)-Dimensional Boolean Hypercube $\mathbf{H}_{N+1}$. It corresponds to an entirely new set, which is not included in the decimal transcription $\mathbb{S}_N$ of the Boolean vertices within the initial Boolean Hypercube $\mathbf{H}_N$. From the Boolean point of view one can also write the concatenation:

$$\langle e_{N+1} | = \{1\} \cup \langle 0_N | = (1, 0, 0, ..., 0)$$

as a new element of the Boolean canonical basis set, then:

$$\mathbf{H}_{N+1} = \mathbf{H}_N \cup \left\{ \langle e_{N+1} | \oplus \mathbf{H}_N \right\}. \tag{4.2}$$

In the equation (4.2) above, the symbol $\oplus$ means that the Boolean vector $\langle e_{N+1} |$ is summed to every vertex of the Boolean Hypercube $\mathbf{H}_N$. Of course, the same rule applies when going downwards on the Boolean Hypercube dimension, wherein this case at the end, the sequence: $\mathbb{A}_{N+1} = \left[2^N \oplus \mathbb{S}_N \right]$ shall be overridden.

5. Mersenne numbers and their extensions

5.1. Mersenne numbers and their twins

As one can easily see, the Mersenne numbers become important when, as discussed in the previous section 4, this kind of partition and iterative construction of natural numbers via Boolean Hypercubes is considered. Mersenne numbers mark the end of the recursive sequences, associated with the decimal transcription of the vertices of the Boolean Hypercubes.

Also, Mersenne numbers might appear when other kinds of natural numbers are studied. For instance, the so-called *perfect* numbers, which can be defined as:

$$p_2(N) = 2^N \cdot \left(2^{N+1} - 1 \right) = 2^N \cdot \mu_2(N+1). \tag{5.1}$$

Perfect numbers, from the Collatz conjecture and Algorithm 1 viewpoint, iterate in the same manner as the associated Mersenne number. That is, one can write:

$$C_{end_p}[p_2(N)] \equiv C_{end_\mu}[\mu_2(N+1)],$$

because the multiplicative term 2^N in the equation (5.1), being even, transforms into the unit after *N* Collatz iterations, as previously discussed in the equation (3.2). Thus, this is the same as to consider that:

$$end_p = end_\mu + N.$$

At the same time, one can observe the trivial extension of the Mersenne numbers (here will be named *twin* Mersenne numbers), corresponding to the second term (the first one is 2^N, which is Collatz compliant) in the sequence $\mathbb{A}_N = \left[2^N \oplus \mathbb{S}_N \right]$ defined in the equation (4.1):

$$v_2(N) = 2^N + 1. \tag{5.2}$$

Thus, these new numbers: the Mersenne twins (5.2), are defined in this way. Then, it can be easily deduced that:

$$\mu_2(N) v_2(N) = 2^{2N} - 1 = \mu_2(2N).$$

5.2. Mersenne twins and base 3 Mersenne-like numbers

It is easy to prove the Collatz path connection between Mersenne numbers and similar powers of 3. For instance, using the symbol $C_I \left[\mu_2(N) \right]$ to indicate the application of the Collatz algorithm at the *I*-th iteration. That is, a partial path made of *I-th* iterations, then one can write:

$$
\begin{aligned}
C_1 \left[\mu_2(N) \right] &= 3 \cdot 2^N - 3 + 1 = 3 \cdot 2^N - 2 \rightarrow \\
C_2 \left[\mu_2(N) \right] &= 3 \cdot 2^{N-1} - 1 \rightarrow \\
C_3 \left[\mu_2(N) \right] &= 3^2 \cdot 2^{N-1} - 3 + 1 = 3^2 \cdot 2^{N-1} - 2 \rightarrow \\
C_4 \left[\mu_2(N) \right] &= 3^2 \cdot 2^{N-2} - 1 \rightarrow \dots \\
C_{2N} \left[\mu_2(N) \right] &= 3^N - 1 = \mu_3(N)
\end{aligned}
\tag{5.3}
$$

Meaning that, after a partial path containing a finite number of *2N* iterations, *any* Mersenne number submitted to the Collatz **Algorithm 1** converges to an equivalent number:

$$\mu_3(N) = 3^N - 1,$$

providing a Mersenne-like number associated with a base 3 instead of 2. Note that $\mu_3(N)$ are even natural numbers.
This result implies that the number of iterations needed to complete the respective paths can be related by:

$$end_2 = end_3 + 2N.$$

Unfortunately, it cannot be shown the same property with the numbers defined in the equation (5.2), that is, one has in general:

$$C_K \left[v_2(N) \right] \neq 3^N + 1 = v_3(N).$$

However, one can study the Mersenne twins with even and odd powers of the base 2, in such a way that:

1. $N = 2n$.

Then one can write the Mersenne twin as:

$$v_2(N) = v_2(2n) = 4^n + 1 = v_4(n)$$

and the Collatz algorithm transforms the base 4 number into:

$$C \left[v_2(2n) \right] = C \left[v_4(n) \right] = v_3(n).$$

A particular case of Mersenne twins of this kind are the so-called Fermat numbers, which can be written as:

$$v_2 \left(2^N \right) = 2^{2^N} + 1$$

but one can easily write:

$$v_2 \left(2^N \right) = v_4 \left(2^{N-1} \right) \rightarrow C \left[v_4 \left(2^{N-1} \right) \right] = v_3 \left(2^{N-2} \right) = 3^{2^{N-2}} + 1$$

1. $N = 2n + 1$.

Then one can also write:

$$v_2(N) = v_2(2n+1) = 2 \cdot 4^n + 1 = \eta_4(n)$$

thus, it is easy to see that in this case the Collatz algorithm transforms the number $\eta_4(n)$ in a similar manner as in case a):

$$C \left[\eta_4(n) \right] = 2 \cdot 3^n + 1 = \eta_3(n).$$

Some numerical calculations have been performed on several $v_2(N)$ and $v_3(N)$ numbers. It has been found that the number of iterations to complete the Collatz paths might vary with varying the power *N*:

$$C_{end_{v_2}} \left[v_2(N) \right] \wedge C_{end_{v_3}} \left[v_3(N) \right] \rightarrow end_{v_3} > end_{v_2},$$

that is, the number of iterations to arrive at the completion of Collatz **Algorithm 1** appears reversed to the case of Mersenne numbers $\mu_2(N)$ and $\mu_3(N)$.

5.3. Computational tests over Mersenne primes

Numerical tests over Mersenne numbers with powers yielding prime numbers, which can be obtained from the list posted in reference [25], prove that several prime Mersenne numbers containing an exceptionally large number of digits are Collatz compliant. The same occurs with their twins.

Of course, the number of Collatz iterations is increasing with the digits of the Mersenne prime and his twin, subject to the Collatz Algorithm 1, as Table 1 indicates.

The interest of the associated Mersenne numbers of base 3, $\mu_3(N)$, which is apparent from the equation (5.3), has raised the curiosity to test them numerically from Collatz **Algorithm 1** point of view. As a result, all have been found Collatz compliant, but differing from the Mersenne numbers $\mu_2(N)$, yielding a lesser number of iterations towards Collatz completion. The difference between the complete path iterations happened to be systematically equal to $2N$, as a consequence of the demonstrated property shown in the equation (5.3).

Table 1 has not been augmented with exponents larger than the displayed in the table, because while the information of first elements was obtained in a reasonable time, the last one has taken several days to complete the Collatz path, using a desktop computer with an i7 processor. Every element of the tested sample of prime Mersenne numbers, as shown in Table 1, corresponds to a Collatz **Algorithm 1** starting on a natural number containing an exceptionally large set of decimal figures.

N	#Iterations $\mu_2(N)$	#Iterations $\nu_2(N)$
44497	598067	317327
86243	1158876	618898
216091	2906179	1562363
756839	10197081	5486483
859433	11568589	6209600
2976221	40055567	21532695
3021377	40663017	21846558
6972593	93778449	50464352

Table 1: Values of the Power N of assorted prime Mersenne numbers, $\mu_2(N)$, see reference [25], and their twin numbers, $\nu_2(N)$, and the necessary Collatz iterations to obtain complete paths

Collatz **Algorithm 1** has been also tested for the same assorted exponents employed in the Mersenne sequence, but on the numbers $\nu_2(N)$, and after this also on $\nu_3(N)$ numbers, which were also computed with the same powers as those of Table 1, to study if the $2N$ rule of the equation (5.3) was also fulfilled in these cases. Results of the termination iterations of $\nu_2(N)$ within the Collatz **Algorithm 1** can be also seen in Table 1.

5.4. Equal iteration numbers on complete path of collatz*Algorithm 1*

Due to the earlier mentioned Collatz connection between Mersenne numbers and the powers of base 3, also, as already commented, the twin of $\mu_3(N)$, $\nu_3(N)$ has been tested for the same exponents as those shown in Table 1. The result was intriguing because, without exception, it has been found the *same* number of iterations on both twins, obtained with the exponents of Table 1, to complete the Collatz **Algorithm 1**. That is, it was found:

$$C_{end_\mu}\left[\mu_3(N)\right] = C_{end_\nu}\left[\nu_3(N)\right] \rightarrow end_\mu = end_\nu. \tag{5.4}$$

Such a result was unexpected; thus, an extra search was performed to shed as much light as possible into such a peculiarity, its possible extension, and the possibility that this was a general feature of these numbers.

At the light of the obtained results, one could test the conjecture that the twin pairs $\{\mu_3(N), \nu_3(N)\}$ possessed the property described in the equation (5.4) for any value of the power N. Alternatively, one could consider that the results were obtained just by chance and that the tested twins have shown this property because only the limited number of powers of Table 1 were used.

Now one must stress that even if the iteration termination numbers become equal in these twins of base 3, this does not preclude that the iterative sequences of Collatz paths become the same. In fact, for both twins, Collatz **Algorithm 1** already starts differently, because: $\mu_3(N)$ and $\nu_3(N)$ are both different even natural numbers. It is interesting to obtain the result of the product of both twins for any power:

$$\mu_3(N)\nu_3(N) = \left(3^N - 1\right)\left(3^N + 1\right) = 3^{2N} - 1 = \mu_3(2N) \tag{5.5}$$

which yields the same result that the Mersenne twins of base 2.

5.5. Numerical tests to search for equal iteration numbers in complete paths of base 3 Mersenne twins

To test the possibility of the presence of the property of equal terminating iterations in a large range of powers, a numerical prospection has been made for the twin pairs of base 3, $\{\mu_3(N), \nu_3(N)\}$ using powers lying in the interval $\{2,...,8192 = 2^{13}\}$.

Results revealed that there exists a **large** percentage of base 3 twin numbers bearing equal Collatz iteration termination numbers. Base 3 twins possess in the tested interval above an abundant number of twins having the same terminal iteration numbers. The results yielded around **93%** of the total tested twins within the above interval.

Unfortunately, this property **is not a general one**, as **7%** of base 3 twins **do not** possess the property of possessing complete paths with equal iteration numbers. Even bearing this curious property, the Collatz complete paths appear essentially different in both twins but arrive at the terminal completion unit at the **same** iteration time. Then, as the percent of equal iterations to reach complete paths in the studied power range is so high, no wonder that the small sample of base 3 twins of Table 1 present all of them the same complete path iteration numbers.

Contrarily to the base 3 case, the true Mersenne twin numbers: $\{\mu_2(N), \nu_2(N)\}$ both in Table 1 and also studied within the power interval: $\{2-8192\}$ do not present at all coincident Collatz **Algorithm 1** complete path termination iteration numbers, although all of them are Collatz compliant. This was also a remarkable result in full contrast with the already discussed base 3 twins.

5.6. Mersenne-like twin pairs involving prime number base

One must note here that while Mersenne twins are odd, base 3 twins are even, and perhaps the equal iteration numbers shown in base 3 twins is due to this kind of odd-even startup of the Collatz **Algorithm 1**.

The property of the equation (5.5) can be extended to any base number, as one can straightforwardly write:

$$\mu_B(N)\,\nu_B(N) = \left(B^N - 1\right)\left(B^N + 1\right) = B^{2N} - 1 = \mu_B(2N) \tag{5.6}$$

A property that might be related to the behavior of these Mersenne prime twins concerning **Algorithm 1**.

Tests performed with several **even** number bases show the same behavior as Mersenne twins: **none or a scarce** number of equal termination iterations of Collatz paths.

However, tests on **prime or odd** number bases provided a similar behavior as the base 3 results. Many powers show equal Collatz complete paths termination iterations. When the number of powers in a tested sample increases, this produces an increase of twins' percent possessing equal termination iterations.

Table 2 shows the results of complete path Collatz iterations coincidences in several prime bases, calculated in a sample of powers contained within the interval: $\{2-8192\}$.

B	3	5	7	11	13	17	19	23	29	31	37	41	43	47
%	93	95	91	95	96	92	96	93	96	92	96	96	96	93

Table 2: Percentages of coincident Collatz iterations to reach the complete path using the fourteen first prime base twins $\{\mu_B(N), \nu_B(N)\}$ computed with the powers in the interval $\{2-8192\}$.

From Table 2 one can conclude that the percentages fluctuate with the prime base number, but these fluctuations keep the amount of equal terminating iterations for twins on several bases greater than 90%, with an average of 94% of tested cases. This means that there is a high chance to find out odd base twins with equal terminating iterations within a large set of powers.

This might explain the results found when calculating Table 1. Several previous tests have shown a clear trend of obtaining larger percentages of equal Collatz complete path iterations as the set of tested power range increases.

This can be illustrated by the base 7 behavior, which provides, in the power range $\{2-1024\}$, a 77% of equal termination iterations, but in the power range $\{2-2048\}$, the obtained results deliver 83%, while in the power range $\{2049-4096\}$ this percentage rises to 92%. The value for this base 7 presented in Table 2, corresponds to a raising percentage, weighted by the increasing trend associated with larger exponents.

5.7. Numerical proof of Collatz conjecture

One must also mention that the results of Table 2 can also be considered as a large sample test (around 160,000 elements) of Collatz **Algorithm 1** convergence, obtained within a large interval of natural numbers: $\left\{\left(3^2 \pm 1\right) \to \left(47^{8192} \pm 1\right)\right\}$.

In all the studied cases included in this paper, no Collatz **Algorithm 1** cyclic loops or divergences associated with any tested natural number appeared. Statistically speaking, the obtained numerical results look like one can safely accept Collatz conjecture as true, at least from a computational point of view.

6. On trying to prove Collatz conjecture

Besides the already shown computational results, one might attempt to prove the Collatz conjecture. This could be tested by using the recursive structure of the natural number set. That is, constructing Boolean Hypercube using concatenation and afterward transforming the Boolean Hypercube vertices into natural numbers, as earlier discussed in section 4 and as shown in **Algorithm 2**.

6.1. Some preliminary considerations

Under the extensive computational tests to show the Collatz compliance of the natural numbers, as it has been already commented, no natural number checked in the described numerical tests have shown cyclic or divergent Collatz **Algorithm 1** behavior. In the sampling of the commented properties in section *5.7* above **no exception** to overall convergence has been found.

Nevertheless, these previous computational-experimental findings, as many others described in reference [26], and still ongoing tests, see reference [27] for example, by themselves *do not prove* the Collatz conjecture. However, they can be surely taken as tentative evidence of the general Collatz compliance of the natural number set as a whole.

Another comment one can put forward is the obtained logical fact showing that, considering some natural numbers made as powers of some natural number base B: B^N, the number of Collatz **Algorithm 1** iterations increases as the base number B grows keeping the same exponent, and also the number of Collatz **Algorithm 1** iterations increases when keeping the base number constant and increasing the exponent.

Therefore, it seems a bit useless to try a systematic search of Collatz compliance using as much as a possibly large set of natural numbers. Like the computational endeavor promoted in reference [27], which constitutes a web page dedicated to the $(3x+1)$ problem, studied via a large computing effort. As it seems that, according to the recent study of Tao [4], it is sufficient to consider some numeric sequences, which can be associated with the Collatz compliance problem.

This is the case of Mersenne numbers $\mu_2(N)$, or the $\nu_2(N)$, $\mu_3(N)$ and $\nu_3(N)$ numbers when studied with large power sequences.

Also, in the twin sequences $\{\mu_B(N), \nu_B(N)\}$ with the first ten prime natural numbers chosen as a base B, the iteration numbers to reach the complete path of the Collatz **Algorithm 1** could be huge, as huge are the natural powers tested, which can be of the order of 10^{8000} or greater. In these cases of extremely large natural numbers submitted to Collatz **Algorithm 1** testing, no cyclic nor convergence problems have been encountered, except the challenging sizable amount of computing time needed to finish the Collatz complete paths using a standard desktop CPU computer.

6.2. Are Mersenne numbers collatz compliant?

It has been demonstrated, see the equation (5.3), that Collatz **Algorithm 1** applied to Mersenne numbers, becomes equivalent to the transformation providing a partial path via $2N$ iterations: $\mu_2(N) \to \mu_3(N)$, into Mersenne-like numbers with base 3.

Thus, the Collatz compliance problem of Mersenne numbers is equivalent to the compliance of the Mersenne-like numbers of base 3. The numbers $\mu_3(N)$ as already commented are even though. Therefore, if it can be proven that even natural numbers are Collatz compliant, then this will also show that odd true Mersenne numbers $\mu_2(N)$ are Collatz compliant.

6.3. Collatz compliance

When trying to understand the general case of the Collatz **Algorithm 1** application to any natural number, the iterative construction shown in the equation (4.1) can be recalled as an adequate step to develop a test of possible inductive proof of the Collatz conjecture.

Hence, if one has found that at some Boolean Hypercube dimension step N, all the elements of the transcribed natural number sequence $\mathbb{S}_N$ have been found Collatz compliant; that is, supposing that:

$$\forall m \in \mathbb{S}_N \to C_{end}[m] = 1$$

or using a more compact notation as previously noted:

$$C_{\langle end|}[\mathbb{S}_N] = \langle 1_{2^N}|.$$

Assuming moving one step further, with the natural transcription of the vertices of a Boolean Hypercube of dimension $N+1$, then one can study from the Collatz **Algorithm 1** the behavior of the elements of the additional set, which truly conform the transcribed augmented natural number set $\mathbb{S}_{N+1}$ as described in the equation (4.1).

Proving that the elements of the set: $\mathbb{A}_{N+1} = [2^N \oplus \mathbb{S}_N]$ are Collatz compliant, then the Collatz compliance of any natural number set, constructed in the way proposed in this work, leads to consider Collatz conjecture to be inductively proved.

6.4. Collatz compliance of natural even numbers

Moreover, the set $\mathbb{A}_{N+1}$ can be divided into the two disjoint sets of even: $\mathbb{E}_{N+1}$ and odd $\mathbb{O}_{N+1}$ elements, that is:

$$\mathbb{A}_{N+1} = \mathbb{E}_{N+1} \bigcup \mathbb{O}_{N+1} \Leftarrow \mathbb{E}_{N+1} \bigcap \mathbb{O}_{N+1} = \emptyset. \tag{6.1}$$

Now, one can easily prove that all the even elements of the set $\mathbb{E}_{N+1}$, after the first Collatz iteration, they transform into elements of the set $\mathbb{S}_N$, which have been supposedly considered as Collatz compliant.

Indeed, one can write a trivial sequence:

$$\forall m_{N+1} \in \mathbb{E}_{N+1} \wedge \forall k \in \mathbb{S}_{N-1} : m_{N+1} = 2^N + 2k \to$$
$$C_1[m_{N+1}] = 2^{N-1} + k = m_N \in \mathbb{S}_N \to \tag{6.2}$$
$$C_{end_N}[m_N] = 1 \to C_{end_{N+1}}[m_{N+1}] = 1$$

The sequence in the equation (6.2) demonstrates that all the even numbers $\mathbb{E}_{N+1}$ contained in the set $\mathbb{A}_{N+1}$ are Collatz compliant. This is so since, after iteration 1, the resultant number in the partial path belongs to the previous step set of natural numbers, and it could be considered by hypothesis Collatz compliant. Then, using the property of section 3.4, the even numbers contained in the set $\mathbb{E}_{N+1}$ are Collatz compliant.

Thus, the induction proof apparently might permit to deduce that the whole set of natural even numbers are Collatz compliant.

However, the induction proof appears *incomplete*. Because it doesn't use the whole natural set $\mathbb{A}_{N+1}$ contained in the equation (6.1), but just the even elements $\mathbb{E}_{N+1}$, and lacks to include the odd number set $\mathbb{O}_{N+1}$, which is necessary to deduce in the next induction step that even numbers are Collatz compliant.

Thus, it is **not** possible within the present analysis by induction to deduce that all even numbers are Collatz compliant.

One can use instead, this potential property as a new conjecture.

Conjecture 1: Partial Collatz compliance

"All even natural numbers are Collatz compliant".

As one can write the natural number set as the union of even and odd numbers:

$$\mathbb{N} = \mathbb{E} \cup \mathbb{O} \wedge \mathbb{E} \cap \mathbb{O} = \emptyset,$$

the induction procedure result, performed over the even numbers whole set could be now **conjectured** to behave as:

$$\forall e \in \mathbb{E} \to C_{end}[e] = 1,$$

6.5. Collatz compliance of Mersenne numbers

The **Conjecture 1** about partial Collatz compliance of the even natural numbers is sufficient to deduce that Mersenne numbers are Collatz compliant, even if they are odd natural numbers. As it has been commented in section 5.2.

Recalling equation (5.3), one can easily write:

$$C_{2N}[\mu_2(N)] = \mu_3(N) \in \mathbb{E}_{N+1} \rightarrow C_{end}[\mu_3(N)] = 1 \rightarrow C_{end}[\mu_2(N)] = 1,$$

showing that *any* Mersenne number is Collatz compliant.

One can also consider the results shown in Table 2, as an indirect computational proof of the Collatz compliance of even numbers, besides of the curious property of twin Mersenne numbers, having a large probability to converge in the same number of iterations but via different paths. One must keep in mind, though, that computational behavior, even if consistent over a large set of even Mersenne twins, constructed with base prime numbers, cannot be a complete proof of **Conjecture 1**.

This is like this because it cannot be discarded that perhaps does exist a unique natural even number, which is not Collatz compliant and has not yet computationally found. Although, besides the huge computational effort of reference [27], there are also contemporary attempts to set up a proof of the Collatz conjecture by computational means [28]. Perhaps they are a pair of useless efforts.

6.6. Collatz compliance of odd natural numbers

Thus, there is only left to try if it is possible to make use of **Conjecture 1** that if true, one can demonstrate that the elements of the odd subset $\mathbb{O}_{N+1}$ are also Collatz compliant. That is, only the natural numbers in the sequence:

$$\forall m_{N+1} \in \mathbb{O}_{N+1} \wedge \forall k \in \mathbb{S}_{N-1} : m_{N+1}(k) = 2^N + (2k+1) \tag{6.3}$$

need to be considered now, once the even natural numbers $\mathbb{E}$ are supposedly Collatz compliant.

Then, the *first* iteration of Collatz **Algorithm 1** applied to the numbers belonging to the odd set $\mathbb{O}_{N+1}$ of $\mathbb{A}_{N+1}$, generally represented by the equation (6.3), will yield:

$$\begin{aligned} C_1[m_{N+1}(k)] &= 3m_{N+1}(k) + 1 \\ &= 3 \cdot 2^N + 3 \cdot 2k + 4 = 2 \cdot (3 \cdot 2^{N-1} + 3 \cdot k + 2) \in \mathbb{E} \end{aligned} \tag{6.4}$$

which yields an even number. Then one can also write:

$$C_2[m_{N+1}(k)] = 3 \cdot 2^{N-1} + 3 \cdot k + 2 = (2^N + 2^{N-1} + 2k + 2) + k$$

therefore, one can see that further application of Collatz **Algorithm 1** will depend on the odd or even nature of the number $k \in \mathbb{S}_{N-1}$.

The equation (6.4) can be seen as demonstrating that odd numbers in $\mathbb{O}_{N+1}$ are Collatz compliant if and only if all the *whole set of even* natural numbers can be compliant. **Conjecture 1** can again be invoked to write, as it is the case to be true:

$$\forall o \in \mathbb{O} \rightarrow C_{end}[o] = 1,$$

which allows to finally write the Collatz conjecture:

$$\forall n \in \mathbb{N} \rightarrow C_{end}[n] = 1.$$

7. Conclusion

A large set of numerical experiments reported here have yielded no natural number presenting compliance problems when tested with Collatz **Algorithm 1**. The set of examples tested has been mainly focused on Mersenne numbers, Mersenne-like, and related number sets, bearing diverse base and powers.

With odd base numbers and for a large range of powers Mersenne-like twins appear which follow, in a percentage larger than 90%, different paths to completion, but present the same number of Collatz iterations to achieve the respective complete paths.

The number of numerical experiments has been sufficiently large and varied allowing to say that Collatz conjecture holds from the statistical point of view. However, this is an inconclusive proof of the conjecture.

After using the recursive building of the whole natural number set, provided by the Boolean Hypercube concatenation and the following Boolean vertex transcription into decimal numbers, it has been tested if it was possible to construct a general inductive proof of the Collatz compliance of the whole natural number set.

Results indicate that one can reach an incomplete verification of the Collatz conjecture by an inductive proof, based on a recursive construction of natural numbers. But even if this attempt was inconclusive, it leads to a new partial **Conjecture 1** involving just the even natural numbers. The attempted proof even incomplete sheds light on the fact that, if the whole set of even natural numbers can be proved Collatz compliant, then the whole natural set is compliant.

Compliance with ethical standards

Conflict of interest: The author states that there is no conflict of interest related to this work.

References

[1] J. C. Lagarias, *The 3x+1 problem and its generalizations*, Am. Math. Monthly, **92** (1985), 3-23.

[2] J. C. Lagarias (Editor), *The Ultimate Challenge: The 3x+1 Problem*, American Mathematical Society, 2010.

[3] *Collatz conjecture*, available at https://en.wikipedia.org/wiki/

[4] T. Tao, *Almost all orbits of the collatz map attain almost bounded values*, (2019), arXiv:1909.03562v2 [math.PR].

[5] J. C. Lagarias, *The 3x+1 problem: An Annotated Bibliography, I* (1963-1999), (2009), arXiv:math/0309224v12 [math.NT].

[6] J. C. Lagarias, *The 3x+1 problem: An Annotated Bibliography, II* (2000-2009), (2009), arXiv:math/0608208v5 [math.NT].

[7] R. Carbó-Dorca, *Boolean hypercubes and the structure of vector spaces*, J. Math. Sci. Modelling, **1** (2018), 1-14.

[8] R. Carbó-Dorca, *Natural vector spaces, (Inward power and Minkowski norm of a natural vector, natural Boolean hypercubes) and Fermat's last theorem*, J. Math. Chem., **55** (2017), 914-940.

[9] R. Carbó-Dorca, C. Muñoz-Caro, A. Niño, S. Reyes, *Refinement of a generalized Fermat's last theorem conjecture in natural vector spaces*, J. Math. Chem., **55** (2017), 1869-1877.

[10] R. Carbó-Dorca, *Cantor-like infinity sequences and Gödel-like incompleteness revealed by means of Mersenne infinite dimensional Boolean hypercube concatenation*, J. Math. Chem., **58** (2020), 1-5.

[11] R. Carbó-Dorca, *Fuzzy sets and Boolean tagged sets*, J. Math. Chem., **22** (1997), 143-147.

[12] R. Carbó-Dorca, B. Calabuig, *Molecular similarity and quantum chemistry*, M. A. Johnson, G. M. Maggiora (editors) Chapter 6 in Concepts and Applications of Molecular Similarity, John Wiley & Sons Inc., New York, 1990, pp. 147-171.

[13] R. Carbó, B. Calabuig, *Molecular Quantum Similarity Measures and N-Dimensional Representation of Quantum Objects II. Practical Applications (3F-Propanol conformer taxonomy among other examples)*, Intl. J. Quant. Chem., **42** (1992), 1695-1709.

[14] R. Carbó-Dorca, *About Erdös discrepancy conjecture*, J. Math. Chem., **54** (2016), 657-660.

[15] R. Carbó-Dorca, *N-dimensional Boolean hypercubes and the Goldbach conjecture*, J. Math. Chem., **54** (2016), 1213-1220.

[16] R. Carbó-Dorca, *A study on Goldbach conjecture*, J. Math. Chem., **54** (2016), 1798-1809.

[17] R. Carbó-Dorca, *Boolean hypercubes as time representation holders*, J. Math. Chem., **56** (2018), 1349-1352.

[18] R. Carbó-Dorca, *DNA, unnatural base pairs and hypercubes*, J. Math. Chem., **56** (2018), 1353-1356.

[19] R. Carbó-Dorca, *Transformation of Boolean hypercube vertices into unit interval elements: QSPR workout consequences*, J. Math. Chem., **57** (2019), 694-696.

[20] R. Carbó-Dorca, *Role of the structure of Boolean hypercubes when used as vectors in natural (Boolean) vector semispaces*, J. Math. Chem., **57** (2019), 697-700.

[21] R. Carbó-Dorca, T. Chakraborty, *Hypercubes defined on n-ary sets, the Erdös-Faber-Lovász conjecture on graph coloring, and the polypeptides and RNA description spaces*, J. Math. Chem., **57** (2019), 2182-2194.

[22] J. Chang, R. Carbó-Dorca, *Fuzzy hypercubes and their time-like evolution*, J. Math. Chem., **58** (2020), 1337–1344.

[23] K. Balasubramanian, *Combinatorial multinomial generators for colorings of 4D-hypercubes and their applications*, J. Math. Chem., **56** (2018), 2707-2723.

[24] K. Balasubramanian, *Computational multinomial combinatorics for colorings of 5D-hypercubes for all irreducible representations and applications*, J. Math. Chem., **57** (2018), 655-689.

[25] https://www.mersenne.org/primes/

[26] A.V. Kontorovich, J. C. Lagarias, *Stochastic models for the 3x+1 and 5x+1 problems*, (2009), arXiv:0910.1944v1 [math.NT].

[27] http://www.ericr.nl/wondrous/

[28] W. Ren *A new approach on proving collatz conjecture*, Hindawi J. Math., (2019), Article ID 6129836, 1-12.