ISSN: 2834-7706

Research Article

Journal of Mathematical Techniques and Computational Mathematics

Considerations on the Collatz Conjecture

Audemir Loris*

Armando Alvares Penteado University Center- FAAP, Brazil

Corresponding Author
Audemir Loris, Armando Alvares Penteado University Center- FAAP, Brazil.

Submitted: 2024, Aug 19; **Accepted:** 2024, Sep 10; **Published:** 2024, Oct 23

Citation: Loris, A. (2024). Considerations on the Collatz Conjecture. *J Math Techniques Comput Math, 3*(10), 01-26.

Abstract
Part of the scientific community has spent considerable time and resources to somehow validate Collatz's conjecture, countless efforts have achieved considerable progress in this direction, but this conjecture lacked definitive confirmation that choosing an ***odd number*** *any xi* $\in \mathbb{N}^*$*, we will obtain* $_{x(i+1)} = 3x_i + 1$*, this being an even number* $x_{(i+1)}$*, divide it if the same by the number two (successively) until another* ***odd number*** $\in \mathbb{N}^*$ *is obtained, the process* $x_n = 3x_{(n-1)} + 1$ *and divisions by two until the result is a number equal to 1. This work presents deductions, algorithms and equations that corroborate this proposition, supporting this perception and conclusion that Collatz's conjecture points to the final cycle 4 → 2 → 1.*

Keywords: Collatz Conjecture, Chaotic Dynamics, Limit Cycle, Periodic Orbit, Principle of Mathematical Induction, Python and 'R' Language

1. Introduction
The direct approach that seeks proofs of convergence to conjunction has proven undecidable, at least any algorithm based on formal logic has been only partially successful, that is, there is currently no algorithm that definitively proves such a conclusion, probably with the advent of and algorithms and quantum computers it is possible to model and prove such a conjecture [1]. The use of transfinite numbers ($\aleph_0$, $\aleph_1$, ...) as well as the sets they represent allows a trend analysis when xi → ∞ being $x_i \in \mathbb{N} \rightarrow| \mathbb{N} |= \aleph_0$. The present approach, based on processes and simulations obtained with the aid of public domain software, aims to conduct part of the research towards obtaining a proof that Collatz conjecture has a final cycle restricted to the sequence 4→2→1.

In the body of this article, results are presented that, based on the principle of induction, when $x_i \rightarrow \infty$ point to the cycle 4→2→1. Using simple tools and a programming language accessible to the general public, in some cases abusing "brute force" in the solution of the same algorithms. This work focuses on presenting the conjecture and its behavior in a succinct manner, taking into account the boundary conditions. Statistical and programmatic approaches that surround the Natural numbers will be explored in a very simple way. Finally, an 'alternative' form of the conjecture will be presented in addition to the programs used in this search [2].

It would not be reasonable to omit that the scientific community in a certain way advises to stay away from such a conjecture given the fact that the mathematical resources for solving such a problem are still unknown (or have not been listed) [3,4].

It is also worth mentioning that the Collatz conjecture has aroused enormous interest, especially among the young community that usually ventures into this wonderful world of Mathematics, in which sometimes due attention is not given to common and trivial statements, such as this simple example: $x_i \in \mathbb{N}^{>1} = \{2, 3, 4, 5...\}$, it can be said that for any $x_i \in \mathbb{N}^{>1}$, x_i will always have as a divisor one or more prime numbers (Fundamental Theorem of Arithmetic → direct consequence of the factorization of integers > 1) [4].

1.1 Background
This article uses algorithms developed in Python and the 'R' language. When presented, they will be duly notified, as well as their relevance [5]. Initially, the environment used for developing the codes in 'R' is presented, and later in Python (this language and environment being preferably used in this work). Remember that the programming interface in 'R' also supports programming in Python. When possible, both solutions, i.e. in 'R' and Python, will be presented, thus allowing the reader to choose the environment that is most suitable for them.

1.1.1 Installing PyCharm

It is assumed that the reader has previously installed the Python language on his/her machine, if he/she has not done so, the following link provides the subsidy for this [6]:
https://python.org.br/ (in Portuguese)
https://wiki.python.org/moin/BeginnersGuide/Download (in English)
There you will be directed to the available solutions and platforms. Then install the PyCharm environment (IDE) from the link:
https://www.jetbrains.com/pt-br/pycharm/download/

Select your platform and download the Community version.

1.1.2 Installing Rstudio

A good guide to download and install the 'R' environment can be found at:
https://livro.curso-r.com/1-instalacao.html
There, follow the necessary steps for your platform.

1.1.3 Testing the Installations

Starting with the 'R' environment, open RStudio and create a new file (R script) with the following code:

```
# Função Collatz presente na biblioteca numbers
library(numbers)
collatz(7)
```

Figure 1: Code: collatz_1.R

After saving the file with the name collatz_1.R and 'running' it, you should get the following output [1]: 7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1

The first line of code instructs the environment to load (install) the library (package) numbers which among its various functions contains: the function Collatz displaying the sequence for: x_i = 7 [7].

Testing the PyCharm + Python environment. Open PyCharm (see appendix C), create a new project called Python_Collatz, in addition to the main.py file, include another Collatz_Files.py, insert the following code:

```
# Implementação vetorial
def collatz_seq(x):
    seq = [x]
    if x < 1:
        return []
    while x > 1:
        if x % 2 == 0:
            x = x // 2
        else:
            x = 3 * x + 1
        seq.append(x)  # Inclui resultado na sequência
    return seq
```

Figure 2: Code: collatz_seq()

The main program main.py will be seen in due time (appendix C), activate the Col- latz_Files.py tab and run the file using the Run File in Python Console option (access with the right mouse button on the function), access the console and activate the collatz_seq(7) function, you should get the following output: [7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1]

The program collatz_1.R (Figure1) and the function collatz_seq(x) (Figure 2) are versions of the Collatz sequence defined by equation (1) [8]. Once the environment is installed and certified, the next items will consider the development of the functions related to the Collatz sequence presented in (1) and later modified in (2).

$$x_{i+1} = f(x_i) = \begin{cases} \frac{x_i}{2^1} & : \text{se } x_i \text{ é par} \qquad (1a) \\ 3 \times x_i + 1 & : \text{se } x_i \text{ é impar} \qquad (1b) \end{cases} \qquad (1)$$

Note that the exponent of the number 2 (two) being 1 (one) implies a single division per 'step' or 'cycle', that is, for successive iterations when x is even, only one division per cycle, a fact that will be adapted to 2^{ρ} later, where a single cycle may include more than one division by 2 (two).

2. Expanding the Collatz Conjecture

Consider the Collatz sequence shown below for x1 = 7:
collatz_d(7) = [7, 22, 11, 34, 17, 52, 13, 40, 5, 16, 1] The Python code for the function collatz_d() can be seen in Figure 3, the results correspond respectively to the values $[x_1][x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}]$, in other words for this sequence five odd numbers are identified xi (before the last x11 = 1 and, five even numbers, that is [9]:

$x_i \in \mathbb{N}_I$ = [7, 11, 17, 13, 5] = [x1, x3, x5, x7, x9] | i = {1, 3, 5, 7, 9}

The values of $x_i \in \mathbb{N}_I$ being odd will be multiplied by 3 and added to the unit $(3x_i + 1)$, this operation will result in a necessarily even number that will be divided (when even) by a power of 2, they are:
$x_i \in \mathbb{N}_p$ = [22, 34, 52, 40, 16] = [x2, x4, x6, x8, x10] | i = {2, 4, 6, 8, 10}
Often only $x_i \in \mathbb{N}^*$ or $x_i \in \mathbb{N}_I \cup \mathbb{N}_p$, remembering that indexes *i odd* numbers represent odd numbers, and *i even numbers* represent even numbers (for the result obtained by the function *collatz_d*(7)). Adjusting the equations in (1) with the necessary modifications we obtain:

$$collatz\_d(x) = \begin{cases} \frac{x}{2^{\rho}} & : \text{se } x \text{ é par, e } \rho \in \mathbb{N}^* \\ 3 * x + 1 & : \text{se } x \text{ é impar} \end{cases} \tag{2}$$

where:

$$x_2 = 3 * x_1 + 1 \text{ being } x_1 = 7, \quad \Longrightarrow x_2 = 22$$
$$x_3 = \frac{3*x_1+1}{2^1} \; ou \; \frac{x_2}{2^1}, \text{ resulting in } \quad \Longrightarrow x_3 = \frac{22}{2^1} = 11$$
$$\ldots$$
$$x_{11} = \frac{x_{10}}{2^4}, \text{ resulting in (this is the last term) } \Longrightarrow x_{11} = \frac{16}{2^4} = 1$$

Collatz sequence seen in the equation (2) and the following code:

```
# Implementação impar/par
def collatz_d(x):
    seq = []
    while True:
        seq.append(x)
        if x == 1:
            break
        if x % 2:
            x = x * 3 + 1
        else:
            while (x % 2) == 0:
                x = x // 2
    return seq
```

Figure 3: Code: *collatz_d*()

Once the values of xi are known, as per the previous example (*collatz_d*(7) where $x_1 = 7$) the ratio between the successive xi is calculated as follow

$$\rho_1 = \frac{x_3}{x_2} = \frac{11}{22} = \frac{1}{2^1}$$
$$\rho_2 = \frac{x_5}{x_4} = \frac{17}{34} = \frac{1}{2^1}$$
$$\rho_3 = \frac{x_7}{x_6} = \frac{13}{52} = \frac{1}{2^2}$$
$$\rho_4 = \frac{x_9}{x_8} = \frac{5}{40} = \frac{1}{2^3}$$
$$\rho_5 = \frac{x_{11}}{x_{10}} = \frac{1}{16} = \frac{1}{2^4}$$

Generalizing we have: $\rho_i = \frac{x_{(2i+1)}}{x_{(2i)}}$, where: ρ_i[10]$= [\frac{1}{2^1}, \frac{1}{2^1}, \frac{1}{2^2}, \frac{1}{2^3}, \frac{1}{2^4}] = [\frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}]$. Rewriting the equations in (2):

$$x_3 = (3 * x_1 + 1) * \rho_1$$
$$x_5 = (3 * x_3 + 1) * \rho_2$$
$$x_7 = (3 * x_5 + 1) * \rho_3$$
$$x_9 = (3 * x_7 + 1) * \rho_4$$
$$x_{11} = (3 * x_9 + 1) * \rho_5$$

or with the appropriate substitutions:

$$x_{11} = (3 * ((3 * ((3 * ((3 * ((3 * x_1 + 1) * \rho_1) + 1) * \rho_2) + 1) * \rho_3) + 1) * \rho_4) + 1) * \rho_5$$

The final term (in this case) $x_{11} = 1$, with the appropriate operations we will obtain the following equation

$$1 = \underbrace{x_1 * 3^5 * \prod_{i=1}^{5}(\rho_i)}_{(A)} + \underbrace{3^4 * \prod_{i=1}^{5}(\rho_i)}_{(B)} + \underbrace{3^3 * \prod_{i=2}^{5}(\rho_i) + \cdots}_{(C)} + \underbrace{3^0 * \prod_{i=5}^{5}(\rho_i)}_{(D)} \tag{3}$$

The equation (3) can be rewritten as follows:

$$\underbrace{x_1 * 3^I * \prod_{i=1}^{I}(\rho_i)}_{(A)} + \underbrace{3^{I-1} * \prod_{i=1}^{I}(\rho_i)}_{(B)} + \underbrace{\sum_{j=(I-2)}^{1}\left(3^j * \left[\prod_{i=(I-j)}^{I}(\rho_i)\right]\right)}_{(C)} + \underbrace{3^0 * (\rho_I)}_{(D)} = x_n \tag{4}$$

In the equations (3) and (4) it is assumed that the Collatz sequence always ends at the number $x_n = 1$ (one)(final cycle 4→2→1), rigorously we obtain an orbit defined by the numerical sequence 1→4→2→1→4→2→1→4→2→1→4→2→1→4→2→1→....

3. Considerations on the Collatz Conjecture

The equation presented in (4) will be broken down into its four parts (A), (B), (C) and (D) for the purpose of study.

3.1 Head of the Collatz Conjecture (A)

The *head* $(A) \Longrightarrow h\ (x_i)$ stands out from the equation (4) where the initial term $x_i = x_1$ can be seen:

$$(A) \Longrightarrow h(x_1) = x_1 * 3^I * \prod_{i=1}^{I}(\rho_i) \tag{5}$$

It can be seen that to calculate $h(x_1)$ it is necessary to know the values of I and $\prod_{i=1}^{I}(\rho_i)$, the function *r_collatz*(7) (shown in Figure 4 below) will provide the necessary subsidies. When executing the function *r_collatz*(7) the answer is →[16,5,11] which corresponds to [C,I,P], that is, 16 Cycles (operations) before obtaining 1(one), with 5 being Odd and 11 being Even [11]

The function *r_collatz*(7) also informs that [I] = 5 (equation (5) $I = 5$), however *r_collatz*(7) does not provide the values of ρ_i, but indicates that the total number of pairs [P] = 11, so we can calculate $\prod_{i=1}^{I}(\rho_i) = \prod_{i=1}^{P}(\frac{1}{2}) = 2^{-P}$ or:

$$\prod_{i=1}^{I}(\rho_i) = 2^{-P} \qquad \longmapsto \qquad \prod_{i=1}^{5}(\rho_i) = 2^{-11}$$

In fact, from the function *collatz_d*(7) seen in figure 3, we obtained the sequence [7, 22,11, 34, 17, 52, 13, 40, 5, 16, 1] and from the generic formula $\rho_i = \frac{x_{2i+1}}{x_{2i}}$ (presented after figure 3):

$$\prod_{i=1}^{5}(\rho_i) = [\frac{1}{2^1} * \frac{1}{2^1} * \frac{1}{2^2} * \frac{1}{2^3} * \frac{1}{2^4}] = 2^{-11}$$

Which allows us to calculate the *head* (A) of Collatz(7):

$$h(x_1) = x_1 * 3^I * \prod_{i=1}^{I}(\rho_i) \quad \rightarrow \quad h(7) = 7 * 3^5 * \frac{1}{2^{11}} \simeq 0.83056640625$$

3.1.1 Head of the Collatz Conjecture: How are the Powers of 2

Consider a number $M \in \mathbb{N}^* = \{1, 2, 3, 4, 5, ...\}$ and the integer 2^M when expressed in binary will have only one of the bits with the value one, being divisible by two "M" times until the number is obtained one as the final answer (in binary the bit whose value is one will be shifted to the right "M" times). By reevaluating the equation presented in (5) and rewriting it taking into account $x_i = 2^M$, $I = 0$ and $P = M$, we arrive at the following expression $\Longrightarrow h(x_i) = 2^M * 3^0 * \rho \mid \rho = \frac{1}{2^M}$, in short $h(x_i) = 1$ in these cases, confirming the validity of the equation described in (5) for $\forall\, x_i = 2^M$.

3.2 Neck of the Collatz Conjecture (B)

In a similar way, the term (B) seen in the equation (4) can be calculated:

$$(B) \Longrightarrow 3^{I-1} * \prod_{i=1}^{I}(\rho_i) \quad \rightarrow \quad 3^4 * \frac{1}{2^{11}} \simeq 0.03955078125$$

It can be easily verified that $3 * x_1 * (B) = (A)$.

3.3 Tail of the Collatz Conjecture (*D*)

Continually, the term (D) is calculated according to equation (4), the term (C) will be calculated later:

$$(D) \Longrightarrow t(x_i) = 3^0 * \rho_n \quad \rightarrow \quad t(7) = 1 * \frac{1}{2^4} \simeq 0.0625$$

It can be seen that the term (D) is constant (for $x_i >= 3$ and $x_i \neq 2^M$) as it corresponds to the final sequence $16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1$ within which the cycle $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$ is included, that is, from 16 to reach 1 it is divided by 24.

3.4 Collatz Conjecture Body (C)

In equation (4), we can see that we need to know the values of x_1, I and ρ_i, remembering that $\rho_i = \frac{1}{\gamma_i}$, the following program fragment, shown in Figure 4, presents the function *r_collatz*(x_1) which will provide such values: *r_collatz* (x_1) $\longrightarrow [C, I, P]$.

```
def r_collatz(num):
    # funcao retorna array [Total de Ciclos,Impares,Pares]
    P = 0
    I = 0
    C = 0
    resp = np.array([C, I, P])
    # verifica se nBIN eh impar se sim continua, caso par RSH
    nBIN = converte(num)
    while nBIN[-1:] == '0':  # bit a direita menos sig.
        nBIN = nBIN[:-1]  # elimina bit zero a direita, RSH
        P += 1
        C += 1
    # criamos nBIN_t  somamos 1 a nBIN e RSH,
    # isto enquanto len(nBIN) > 1
    while len(nBIN) > 1:  # existem bits a serem processados
        if nBIN[-1:] == '1':  # eh impar
            nBIN_t = nBIN
            nBIN = add_binary_nums(nBIN, '1')
            nBIN = nBIN[:-1]
            nBIN = add_binary_nums(nBIN_t, nBIN)
            I += 1
            C += 2
            P += 1  # estou x3 + 1 e dividindo por 2
        else:
            nBIN = nBIN[:-1]
            P += 1
            C += 1
    resp = [C, I, P]
    return resp
```

Figure 4: Code: *r_collatz*()

In the item (C) of the equation (4) considering the Collatz sequence relative to the number 7, that is, using the data of *r_collatz*(7)→[16, 5, 11] ([C,I,P]), where $I = 5$ we obtain:

$$\underbrace{\sum_{j=(I-2)}^{1}\left(3^{j} * \left[\prod_{i=(I-j)}^{I}(\rho_i)\right]\right)}_{(C)}$$

$$\underbrace{3^3 * \prod_{i=2}^{5}(\rho_i)}_{(\alpha)} + \underbrace{3^2 * \prod_{i=3}^{5}(\rho_i)}_{(\beta)} + \underbrace{3^1 * \prod_{i=4}^{5}(\rho_i)}_{(\delta)}$$

Previously it was shown that $\rho_i = [\frac{1}{2^1}, \frac{1}{2^1}, \frac{1}{2^2}, \frac{1}{2^3}, \frac{1}{2^4}]$. In this way we calculate (α), (β) and (δ):

$$\underbrace{3^3 * [\frac{1}{2^1} * \frac{1}{2^2} * \frac{1}{2^3} * \frac{1}{2^4}]}_{(\alpha)} + \underbrace{3^2 * [\frac{1}{2^2} * \frac{1}{2^3} * \frac{1}{2^4}]}_{(\beta)} + \underbrace{3^1 * [\frac{1}{2^3} * \frac{1}{2^4}]}_{(\delta)}$$

$$\underbrace{0.0263671875}_{(\alpha)} + \underbrace{0.017578125}_{(\beta)} + \underbrace{0.0234375}_{(\delta)} = \underbrace{0.0673828125}_{(C)}$$

3.5 Complete Collatz Conjecture

Recovering the previous results we will have:
(A)≃ 0.83056640625, (B) ≃ 0.03955078125, (C) ≃ 0.0673828125, (D) ≃ 0.0625
Adding the terms of the equation (4) we will finally have: (A) + (B) + (C) + (D) = 1 As can be seen from the function *abcd*('7') presented in the program in Figure 5:

abcd('7') → ('7', 0.8305664062499994, 0.039550781249999986, 0.0673828125, 0.0625, 0.9999999999999994).

```
def abcd(num): #Retorna os valores (A)A,(B)B,(C)C e (D)D
    [Ci, I, P, gama] = r_collatz1(num)
    gamas = np.array(gama)
    B = 2**(((I-1)*math.log2(3))-P)
    D = 1/(2**gamas[-1])
    C = 0
    A = 2 ** (math.log2(int(num)) + math.log2(3) * I - P)
    for J in range(I-2,0,-1):
        sgama = 1
        for i in range(I-J-1, I):
            sgama = sgama * (1/(2**gamas[i]))
        C = C + (3**J)*sgama
        # qq Xi = ((2**e)-1) divisivel (int) por 3 somente tera um impar,I, o proprio
          ...
        # sendo assim (A) > 0, (B) > 0, (C) = 0 e (D) = 0
        if I == 1:
            D = 0
    return (num, A, B, C, D, A+B+C+D)
```

Figure 5: Código: *abcd*()

Several numbers from table 1 were tested with the function *abcd*(x_i) and returned the sum $A + B + C + D \approx 1$

x_1	C	I	P	$\sim I/P\%$	$\sim x_1 * \frac{3^I}{2^P}$	$\sim A+B+C+D$[6]
9	19	6	13	46.15	0.8009	1
97	118	43	75	57.33	0.8428	1
871	178	65	113	57.52	0.8639	1
6 171	261	96	165	58.18	0.8395	1
77 031	350	129	221	58.37	0.8084	1
837 799	524	195	329	59.27	0.8373	1
8 400 511	685	256	429	59.67	0.8423	1
63 728 127	949	357	592	60.30	0.8450	1
670 617 279	986	370	616	60.06	0.8450	1
9 780 657 630	1132	425	707	60.11	0.8683	1
75 128 138 247	1228	461	767	60.10	0.8683	1
989 345 275 647	1348	506	842	60.09	0.8942	1
Big_num1[7]	10466	3455	7011	49.28	0.8472	1
NN4[8]	36780	12293	24487	50.20	0.8078	1

Table 1: $x_1 \longrightarrow [C, I, P]$

Note that the column x_1 * corresponds to the Collatz function *head* presented in the equation (5), being equal to the term 'A' presented in the function *abcd*(x_i), shown in penultimate column of this table, in several tests it is shown that $h(x_1) = x_1 * \frac{3^I}{2^P}$ tends to be >0.7, a fact that will be addressed later [13].

4. Exploring Some Sequences

Computers have routinely tested the Collatz Conjecture for increasingly larger numbers (see issue NN4 note 8 above), using powerful machines and improved algorithms that indicate that the Collatz Conjecture 'apparently' ends in its cycle $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$. In the previous item it was shown that for $\forall x_i = 2^M, M \in \mathbb{N}$ the final cycle will always be $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$.

This chapter explores a restricted set of data that, after being processed, allowed some graphs to be drawn and some considerations to be made about them. Because they are restricted (data and graphs), it is clear that they can and should be improved as the tests to be carried out advance.

4.1 Limit Cycle or Orbit

Previously it was considered that the final cycle for the Collatz Conjecture is $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$, a brief explanation of the concept of limit cycle [7] or Orbit is in order. Consider the following transformation:

$$f^n(x_i) = x_{(i+n)} \quad (6)$$

Since $f^n(x_i)$ is the process of transforming the variable x_i, implying one or more times the application of the Collatz Conjecture initially on the variable $x_i \in \mathbb{N}^*$, it is not just the possibility of applying a singular function, as the stages may involve multiple steps where the variable x_i will increase and subsequently decrease, as seen in the definition of the code for the function *collatz_d*(x); the index n in $f^n(x_i)$ denotes that there are several steps, with n being the steps that correspond to increases and decreases.

A limit cycle or Orbit is considered in mathematical operations systems that present the occurrence of the fact that $f^n(x_i) = x_i$ being $x_i \in \mathbb{N}^*$, in this way the existence of a cycle (may be repetitive) or Orbit (may be periodic) within the sequence is verified, it is observed that any limit cycle that may exist in the Collatz Sequence where $x_i \neq 1, x_i \neq 2, x_i \neq 4 \mid x_i \in \mathbb{N}$ leads to the collapse of the Conjecture, as it will inevitably not reach the $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$ orbit, since it was 'captured by another orbit' [9].

In the case of the Collatz Conjecture, using the definition found in equation (2) and used in the construction of the *collatz_d*(x) function, it is possible to verify the existence of (*probably just*) *a 'limit' cycle* where $x(i) = f^n(x_i)$.

From the equation (2) we can group the 'growth' and 'decay' operation of the system into a single equation (for the sequence $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$), namely:

$$\frac{3 * x_i + 1}{2^\gamma} = x_{(i+n)} \qquad \text{(remembering that } x_i \text{ is odd)} \tag{7}$$

With the necessary adjustments where $x_i = x_{(i+n)}$ is obtained from the equation

$$\frac{1}{(2^\gamma - 3)} = x_i \qquad \text{(remembering that } x_i, x_{(i+n)}, \gamma \in \mathbb{N}^*) \tag{8}$$

It is easy to see that the only values that solve the equation (8) are $\gamma = 2$ and $x_i = x_{(i+n)} = 1$, thus the expression $(2^2 - 3) = 1$ evidencing the existence of a ‘limit‘ cycle in the Collatz Conjecture, which will be seen later at the end of section 4.1.2.

4.1.1 Periodic Orbit

It was previously shown that if there are other limit cycles (orbits) within the Collatz Conjecture other than the final cycle 1 → 4 → 2 → 1 it is false. Strictly speaking, the system in (2) cannot be treated as a single procedure function (linear) since there are two possible approaches (use of the conjecture), one for even numbers and the other for odd numbers. In this way, an alternative way of considering the two operations in just one algebraic equation is sought. The candidate equation is presented below, which will certainly facilitate the study of the orbits:

$$x_{i+1} = [(3 * x_i + 1) * (1 - cos^2(x_i * \frac{\pi}{2}))] + [\frac{x_i}{2} * cos^2(x_i * \frac{\pi}{2})] \tag{9}$$

The equation coded above 10 in Python as per Figure 6 below, it operates identically to the function (or code) *collatz_d*(x_i).

```
# Implementação função Collatz
def collatz_ang(x):
    seq = [x]
    if x < 1:
        return []
    while x > 1:
        #x = ((3 * x +1)*(1-int((np.cos(np.pi/2*x))**2))) + x*((int((np.cos(np.pi/2*x))**2))/2)
        x = ((3 * x + 1) * (1 - round((np.cos(np.pi / 2 * x)) ** 2))) + x * ((round((np.cos(np.pi / 2 * x)) ** 2)) / 2)
        seq.append(x)
    return seq
```

Figure 6: Code: *collatz_ang*(x_i)

Note that for even values of x_i the term $[(3 * x_i + 1) * (1 - \cos^2(x_i * \pi/2)]$ becomes null and the operation is just division by two, i.e. $x_{i+1} = \frac{x_i}{2}$. However, if x_i is odd, the term $[\frac{x_i}{2} * cos^2(x_i * \frac{\pi}{2})]$ is nullified, leaving only the result $x_{i+1} = (3 * x_i + 1)$. From the equation (9) it is possible to study the periodic orbit and its respective equilibrium points, considering the following sequence:

$$x_2 = f(x_1)$$
$$x_3 = f(x_2) = f^2(x_1)$$
$$\ldots$$
$$x_n = f(x_{n-1}) \quad ou \quad x_n = f^{(n-1)}(x_1)$$
$$x_{n+1} = f(x_n) \quad ou \quad x_{n+1} = f^{(n)}(x_1)$$

A periodic orbit is established when $x_i = f^{(n)}(x_i)$, that is, $x_{n+m} = x_n$. According to Monterio[7] to know the characteristics of an orbit it is necessary to know its eigenvalue λ, which corresponds to the product of each eigenvalue relative to the fixed points x^*_i in the orbit, as follows:

$$\lambda^{11} = \frac{df^{(n)}(x_i)}{dx}\Big|_{x_i^*} = \frac{df(x)}{dx}\Big|_{x_1^*} \quad \frac{df(x)}{dx}\Big|_{x_2^*} \quad \cdots \quad \frac{df(x)}{dx}\Big|_{x_i^*}$$

By deriving the equation (9) we obtain:

$$\frac{df(x)}{dx} = \frac{1}{2}[(-5(cos^2(x\frac{\pi}{2}))) + (\pi(5x+2)(sen(x\frac{\pi}{2}))(cos(x\frac{\pi}{2}))) + 6] \tag{10}$$

It can be seen that the expression $(sen(x\frac{\pi}{2})cos(x\frac{\pi}{2}))$ will always be zero for $x_i \in \mathbb{N}*$ consequently cancelling the term that multiplies, in this way for the proposed purposes it is possible to calculate λ_j through the following expression: $\lambda_j = \frac{df(x_i)}{dx}|_{x_i^*}$ or according to equation (10) modified to $x_i \in \mathbb{N}*$:

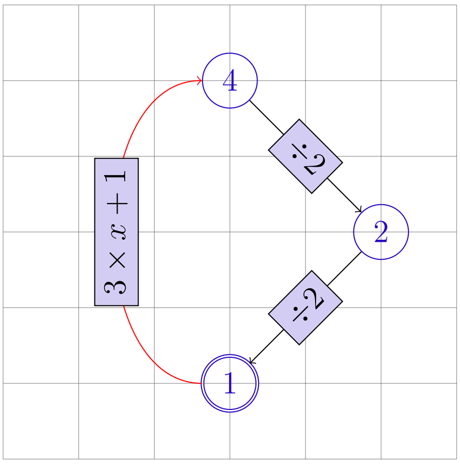
$$\lambda_i{}^{12} = \frac{1}{2}[(-5(cos^2(x_i^*\frac{\pi}{2}))) + 6] \tag{11}$$

Considering the known periodic orbit 4 → 2 → 1→ 4 we proceed to calculate the eigenvalue λ for this orbit:

$$\lambda = \lambda_{(x^*=4)} \times \lambda_{(x^*=2)} \times \lambda_{(x^*=1)} \qquad \text{or} \qquad \lambda = \frac{1}{2} \times \frac{1}{2} \times 3 = \frac{3}{4}$$

Also according to Monteiro[7] as λ < 1 the orbit of period 3 (4 → 2 → 1→ 4) is stable, which allows us to state that such a cycle repeats indefinitely once any of the points belonging to the orbit is reached.

The sequence produced by the Collatz Conjecture for $\forall x_i \in \mathbb{N}*$ being xi < 268[10] will end upon reaching the fundamental orbit 4 → 2 → 1 since this is stable, we will see in the next subsections if there are other possible orbits in the Collatz Conjecture.
Here are some observations about the orbit 1 −→ 4 → 2 → 1



grafo 1 - ciclo fundamental

The Collatz conjecture as seen in (1) is formed by the first three prime numbers 1, 2, 3 ∈ ℕ*, with the main pole being the number ONE, this independent term having the function of making any odd number pre- viously multiplied by THREE even, the even result will be continuously divided by TWO (while even). Note that the number ONE is the smallest that can be ‘calculated‘ by the conjecture after several divisions by TWO. The number FOUR (4 = 3 × 1 + 1), constitutes the other pole of the orbit that when divided by TWO generates the third element of the same orbit, this in turn ‘ends‘ in ONE.

4.1.2 Limiting Case $x_i \to \infty$

Considering that $x_i \to \infty$, that is, is a very large number13, making $x_1 = 2(68) + 1$ for analysis purposes x1 = 295147905179352825857 the program *r_collatz*(x_1) provides as answer [562, 191, 371] namely, Cycles = 562, Odd (rising) = 191 and Even (falling) = 371. Note that the initial cycle will be 3 × x1 + 1, the independent term ONE is much smaller than the product $3 \times x_1$, (1 <<< ($3 \times x_1$)), and can be neglected14 for study purposes, in this way the conjecture will be simplified to according to the following sequence:

$x_2 \simeq \frac{3}{2^{\gamma_1}} \times x_1$

$x_3 \simeq \frac{3}{2^{\gamma_2}} \times x_2$ or $x_3 = \frac{3}{2^{\gamma_2}} \times \frac{3}{2^{\gamma_1}} \times x_1$ in short: $x_3 = \frac{3^2}{2^{(\gamma_2+\gamma_1)}} \times x_1$

...

$x_n \simeq \frac{3^{n-1}}{2^{(\sum_{i=1}^{n-1}\gamma_i)}} \times x_1$ as seen previously in the equation (5):

$x_n \simeq x_1 * 3^I * \prod_{i=1}^{I}(\rho_i)$ or even $x_n \simeq \frac{x_1 * 3^I}{2^P}$

Starting from $x_1 = 295147905179352825857$ we obtain: $x_n \simeq \frac{x_1 * 3^{191}}{2^{371}}$

By evolving the previous expression we obtain:

$x_n \simeq x_1 \times 2^{(191*\log_2(3)-371)}$

...

$x_n \simeq (2^{(68)} + 1) \times 2^{(-68.27216236225917)} \simeq 0.8280774634724271$

The same result can be verified through the function:
abcd(268 + 1) = (x1, A, B, C, A + B + C + D) which presents the following output:
(295147905179352825857,0.8280774634724271,9.352118592531982e-22,0.10942253652757464,0.0625,1.0000000000000018)

As previously mentioned, 13 numbers ≦ 268 were experimentally tested for the Collatz Conjecture and all of them invariably ended in the cycle 4→2→1, empirically demonstrating the non-existence of another repeating cycle other than the trivial, therefore if any other cycle exists, it must have as its origin numbers greater than 268 and of course have its fixed points all greater than 268, (xk > 268, $x_k \in \mathbb{N}^*$), however it was found that such a statement ($x_k = x_n \approxeq 0.8280774634724271$) shows that when $x_i = 2(68) + 1$ is obtained, $x_n \leqq 1$ is obtained, thus evidencing the decrease of the 'series'.

In addition and based on the equation (4) and its alternative form 15 below equation (12), assuming the existence of xk a fixed point of a repetitive cycle we obtain the following equation (13):

$$\underbrace{x_k * 3^I * \prod_{i=1}^{I}(\rho_i)}_{(A)} + \underbrace{\sum_{j=(I-1)}^{0}\left(3^j * \left[\prod_{i=(I-j)}^{I}(\rho_i)\right]\right)}_{(B)+(C)+(D)} = x_k \tag{12}$$

$$(A) = \Psi x_k, \quad (B) + (C) + (D) = \Omega \quad \Psi x_k + \Omega = x_k$$

Multiplying both terms by x_k results in: $x_k^2(\Psi - 1) + \Omega x_k = 0$.

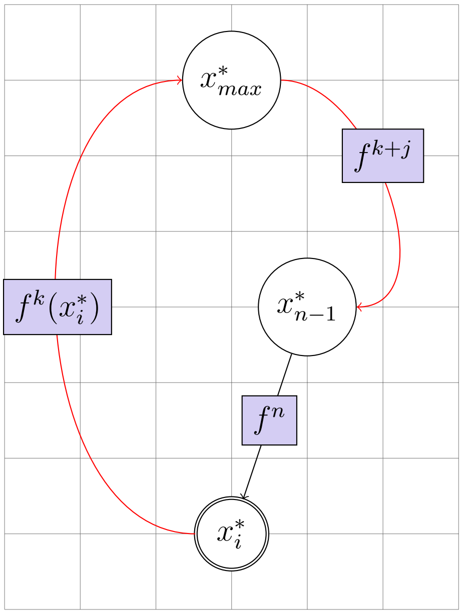
Therefore, in addition to the trivial root $x_k = 0$ the other root will be (this valid one):

$$x_k = \Omega\left(-\tfrac{1}{\Psi-1}\right), \quad \text{remembering that } -\tfrac{1}{\Psi-1} = \tfrac{1}{1-\Psi} \text{ we obtain:}$$

$$\Psi x_k + \Omega = x_k \qquad ou \qquad x_k = \frac{\Omega}{1-\Psi} \tag{13}$$

Which can also be obtained from $\Psi x_k + \Omega = x_k$, it is also observed that in accordance with what was seen previously 16 $0 < \Psi < 1$, and $\Omega \geqq 0$, consequently the terms (A), (B), (C) and (D) are all positive.

From the equation (12) we obtain that $\psi = 3^I * \Pi_{i=1}^{I}(\rho_i)$ or even $\psi = \frac{3^I}{2^P}$, seen previously, we know that $0 < \Psi < 1$, in this way we can write that: $0 < \frac{3^I}{2^P} < 1$, adopting the base two we rewrite it in the following way $0 < 2^{(I*\log2(3)-P)} < 1$ (where P represents simple divisions by 2), for it to be true (that is: valid for $\forall x_i \in \mathbb{N}*$) we have that $(I * \log_2(3) - P) < 0$ or even, $I * \log_2(3) < P$ 17. This limit can also be seen in table 1, also verifying that according to $x^* \to \infty$ we have that $\Psi = \frac{3^I}{2^P} \to 0$ (without ever being zero).



grafo 2 - any closed cycle

Consider the figure to the side (graph 2) which represents any closed (repetitive) cy- cle, as suggested previously if such a cycle exists it must have as fixed points x^*_i numbers greater than 2^{68} ($x^*_i > 2^{68}, x^*_i \in \mathbb{N}^*$) can be seen that in this cycle there 'exists' amaximum value $x^*_i = x^*_{max}$ being the result of the transformations $f^k(x^*_i)$, which take $x*$ in an increasing way up to the maximum value of the cycle. Similarly, the transformations $f^{k+j}(x_{max}) = x_{n-1}$ and finally $f^n(x_{n-1}) = x^*_i$, lead to a decrease in the maximum value x^*_{max}. It is clear that the 'largest divisions by two' are obtained by such transformations f^{k+j} and f^n from the values $\frac{1}{2^{\rho_i}}$, where each value of ρ is part of the vector ρ_i.

Still according to equation (12) the term $\mathbf{\Omega}$ tends towards 'small' finite values when compared to x_i ($> 2^{68}$), it can be seen that the series of values for $\mathbf{\Omega} = (B) + (C) + (D)$ does not include the initial term x_i, depending only on the values of **I** and **P**.

For demonstration purposes, consider the Sequence formed by 10 ascents and 10 descents presented below for the number 57, remembering as observed for Lagarias [6] the cycles of ascents $(3x_i + 1)$ are the same number as those of descents (division by 2^ρ):
r_collatz1(57) = [32, 10, 22], [2, 1, 2, 2, 4, 1, 1, 2, 3, 4] 18
collatz_d(57) = [57, 172, 43, 130, 65, 196, 49, 148, 37, 112, 7, 22, 11, 34, 17, 52, 13, 40, 5, 16, 1]
Expanding the equation (12) using the previous values we have:

$$\underbrace{\frac{x_k * 3^{10}}{2^{22}}}_{(A)} + \underbrace{\frac{3^9}{2^{22}}}_{(B)} + \underbrace{\frac{3^8}{2^{20}} + \frac{3^7}{2^{19}} + \frac{3^6}{2^{17}} + \frac{3^5}{2^{15}} + \frac{3^4}{2^{11}} + \frac{3^3}{2^{10}} + \frac{3^2}{2^9} + \frac{3^1}{2^7}}_{(C)} + \underbrace{\frac{3^0}{2^4}}_{(D)}$$

Using the function *abcd(57)* ($x_i = 57$) we obtain:

$$\underbrace{0.8024675846099856}_{(A)} + \underbrace{0.004692792892456051}_{(B)} + \underbrace{0.1303396224975586}_{(C)} + \underbrace{0.0625}_{(D)}$$

where $\Omega = 0.19753241539001465$, as is evident $\Omega << x_i$.

Taking into account the complexity of the vector $\rho i = \left[\frac{1}{2^2}, \frac{1}{2^1}, \frac{1}{2^2}, \frac{1}{2^2}, \frac{1}{2^4}, \frac{1}{2^1}, \frac{1}{2^1}, \frac{1}{2^2}, \frac{1}{2^3}, \frac{1}{2^4}\right]$, which are the divisors responsible for the 'descents', it is impossible to construct a generic formula for Collatz sequences without completely knowing the vector ρ_i, therefore the study is carried out based on the average value between ascents and descents which, according to Lagarias [6], points to a value lower than the unit $(\longrightarrow \frac{3}{4})$.

Previously it was shown that the minimum limit between **I** and **P** is $\log_2(3)$ or $P \geqslant I \times \log_2(3)$, rewriting the equation (12) taking into account 'as an average value' $\rho_i = \log_2(3)$ we obtain the following equation (14):

$$\underbrace{\frac{x_k * 3^I}{2^{(I*\log_2(3))}}}_{(A)} + \underbrace{\sum_{j=(I-1)}^{0}\left(3^j * \prod_{i=(I-j)}^{I} 2^{-(I*\log_2(3))}\right)}_{(B)+(C)+(D)} = x_k \tag{14}$$

evolving it we have:

$$\underbrace{\frac{x_k * 3^{10}}{2^{(10*\log_2(3))}}}_{(A)} + \underbrace{\frac{3^9}{2^{(10*\log_2(3))}}}_{(B)} + \underbrace{\frac{3^8}{2^{(9*\log_2(3))}} + ... + \frac{3^1}{2^{(2*\log_2(3))}}}_{(C)} + \underbrace{\frac{3^0}{2^{(1*\log_2(3))}}}_{(D)} = x_k$$

or:

$$\underbrace{x_k}_{(A)} + \underbrace{\frac{1}{3}}_{(B)} + \underbrace{\frac{1}{3} + ... + \frac{1}{3}}_{(C)} + \underbrace{\frac{1}{3}}_{(D)} \text{ or even: } \underbrace{x_k}_{(A)} + \underbrace{\sum_{j=1}^{I}\left(\frac{1}{3}\right)}_{(B)+(C)+(D)} = x_k$$

in short: $x_k = x_k + \left(\frac{I}{3}\right)$, which expresses an incoherent relationship, unless I were equal to zero, but as previously mentioned I = 10 [19].

The function *equation_12(x_i, type)*= $\left(x_i, \psi, A, \Omega, \frac{\Omega}{x_i}, x_n\right)$ (coded in Python) returns the following where (*tipo* = 3 =⇒ $\rho = j \times \log_2(3)$):
(57, 1.0000000000000007, 57.000000000000036, 3.333333333333335, 0.058479532163742715, 60.333333333333337)
compatible with such a statement as $\Omega << x_i$ or x^*.

Suppose that ∃ is an orbit in the Collatz conjecture such that $x_k = x^*$ is a fixed point of the orbit, we have: $x^* = \frac{\Omega}{1-\Psi}$, assuming $x^* \to \infty \quad \therefore \quad \Psi$[20] $= \frac{3^I}{2^P} \to 0$ (without reaching zero), it results from the equation (13) that:

$$\lim_{(x^*,\Psi)\to(\infty,0)} \left(\frac{\Omega}{1-\Psi}\right) = \Omega$$

$$\lim_{x^*\to\infty} (x^*) = \Omega$$

However as seen previously $x^* \to \infty$ consequently $0 < \Omega << x^*$ it is again evident that ∄ any value of $x^* \neq 1, x^* \neq 2, x^* \neq 4 \mid x^*21 \in$ N that satisfies the equation (13), thus
we conclude that:

In the Collatz Conjecture there is only one limit cycle formed by the stable points: 1→4→2→1[22].

5. Stochastic Models, Deterministic Process!

The Collatz sequence has been described in several texts under different names, one of which is: *Hailstone Numbers* [5], just as hailstones in clouds before being precipitated go through several 'ascents' and 'descents', the numbers jump from one place to another before reaching the final cycle 4→2→1[23].

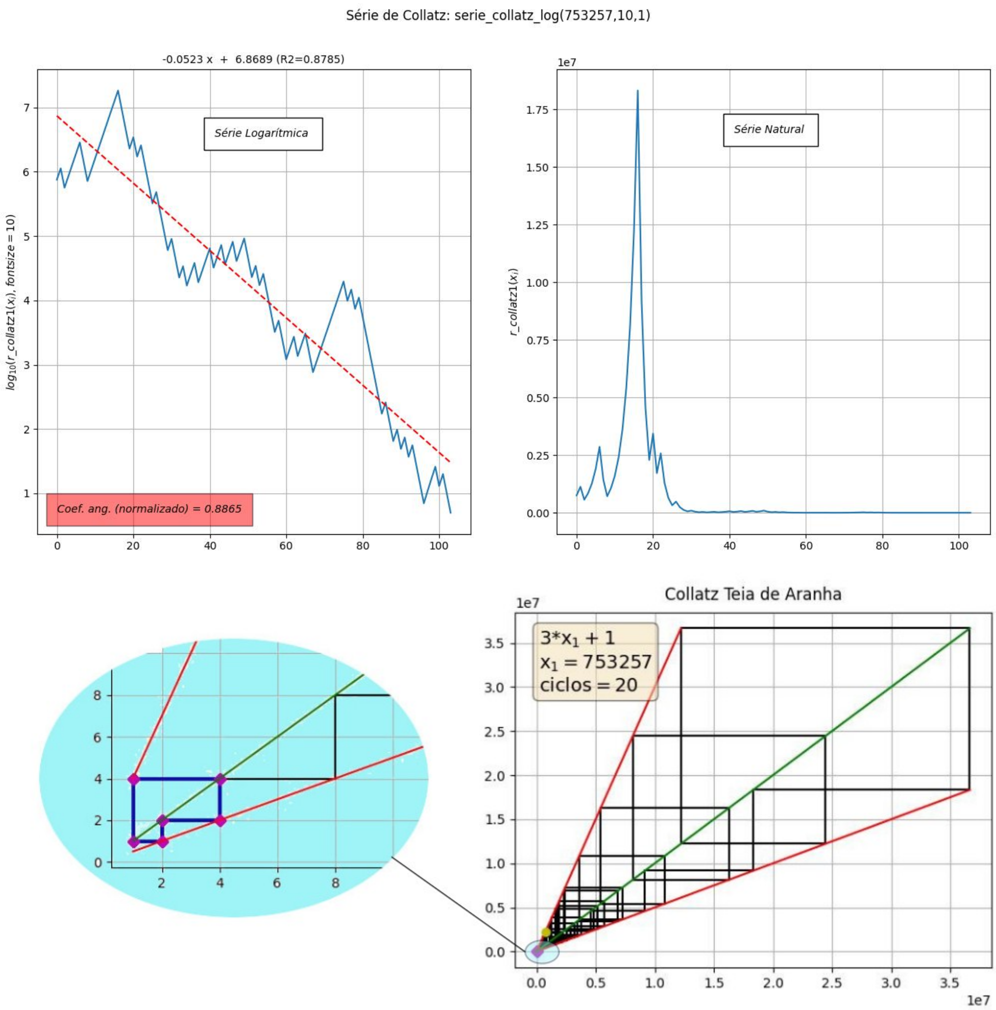


Figure 7: Collatz sequence for x_1 = 753257

Several attempts to understand the Collatz sequence through computer simulations point to the previously seen final cycle 4→2→1, Figure 7 presents an example when x_1 = 753257, you can see the Natural sequence, upper right graph, and next to it the log_{10} of the same sequence, below we see the sequence in a spider web graph and in detail the cycle 4→2→1 [24,25].

It can be seen that when the cycles approach the end (in this case 110 iterations) the response tends towards the final cycle 4→2→1, the negative coefficient (-0.0523) of the approximated line stands out in the logarithmic graph, which causes the successive values of x_i to decrease, the same normalized coefficient $10^{-0.0523}$ = 0.8865... shows that being less than ONE in the successive iterations the value of xi should decrease. In fact, the 'progression' factor of this 'apparent' series is on average less than unity, according to predictions made by Lagarias [6] thus converging in successive iterations to the final cycle 1 → 4 → 2 → 1.

In item 4.1 it has been demonstrated that the sequence contains only one cycle 4→2→1, which is sufficient proof as such for all numbers; however, numbers $xi \in \mathbb{N} \mid x_i \to \infty$ allow for a complementary approach using probabilistic models, presenting significant results that point to this conclusion:

".. a basic probabilistic model of iterations of the function 3x + 1 proposes that most trajectories for iterations 3x + 1 have equal numbers of even and odd iterations" [6](translated by the author).

In the items 2 and 3.1 we saw that the number of ascents is identical to the number of descents, that is, the number of odd numbers is the same as the number of even numbers (equations (2) and (5)), Lagarias also presents this fact [26].

The following items present in a simplified manner some subsidies linked to statis- tical processes that infer a similar conclusion.

6. Relationship between Even and Odd cycles

The following image contains four graphs generated from the function *nucleo1*() and *nucleo2*() present in the main code (main.py) detailed in appendix A.

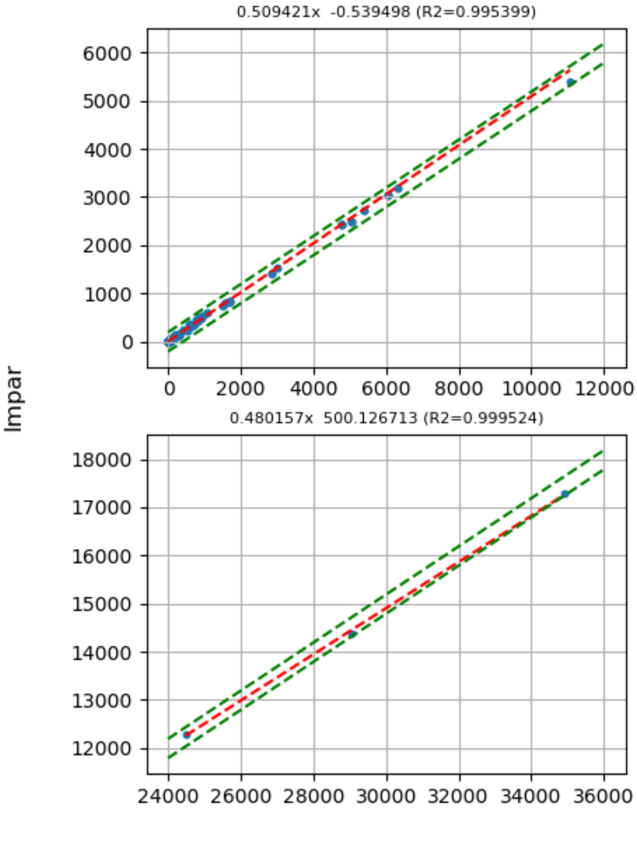


Figure 8: Graphs P x I

The fourth graph shows all the points (*P xI*) that are part of the studied database[27], the first three detail excerpts from the fourth graph. It can be seen that the points (in blue) represent the total number of *P* even cycles and the respective number of *I* odd cycles for a given x_1 to reach *ONE*, while the internal lines (in red) are obtained as the best (linear) approximation to the cited points; the two external lines (in green) limit the values within a specific region that includes all the points in blue of the studied base, that is, the confidence index in this region is 100%. The approximation lines (in red) have their components presented at the top of each graph. In the last graph, we can see a strong convergence between the values *P* x *I*, showing a relationship between them.

7. Limit Region

The two straight line segments in green represent the limit region where all the even points (blue) *P* x *I* meet (referring to the study base). The straight lines have the following components:

$$R_s(\text{reta superior}) \qquad \longrightarrow 0.500000 * P + 189.304 \tag{15}$$

$$R_i(\text{reta inferior}) \qquad \longrightarrow 0.500000 * P - 210.500 \tag{16}$$

It can be seen that they are in fact parallel (same angular coefficients) and the width (spacing) between them is ≃ 400 points, which means that for a certain number of cycles *P* we will obtain within the range presented the value of ~ *I* ± 200 points.

From the presented adjustment equation (last of the four graphs) we have a correlation coefficient of $R^2 = 0.999607$, very close to ONE indicating adherence of the Linear model to the presented distribution:

$$Ciclos_{Impares} = 0.500000 \cdot Ciclos_{Pares} + 0.194725 \tag{17}$$

It can be seen that the correlation between the number of P even cycles and *I* odd cycles ($I \simeq 0.5 \times P$) is in accordance with the Normal distribution of the numbers $\in \mathbb{N}^*$ as will be seen later. It should be noted that although the precision is not absolute, as the equations work with Real numbers subject to rounding, the cycles *P, I* are still positive natural numbers, and the Collatz sequence is deterministic, despite the lack of knowledge about the evolution of the same sequence in relation to all possible numbers $\in \mathbb{N}^*$, these approximations and the use of statistical methods are used. The equation (9) does not assertively indicate that ∀ number xi will end in um after the Total Cycles (=P+I) however the terms A,B,C and D shown in the code in Figure 5 or the equivalent terms (A), (B), (C) and (D) present in the equation (4) provide resources for an approximation based on the initial number xi and possible points IxP , ($C = I + P$).

It is also noted that the limit lines can be better defined by adopting smaller segments for the values x_i.

In the following graphs (Figure 9) it is possible to see for the studied database the Histograms [8] relating to term A of the equation (4), the relationship between odd and even cycles (even cycles include all cycles in which there is a simple division by two) and it is also observed that the number of odd cycles is smaller than the number of even cycles (total), and in the last graph that any cycle (in the studied database) has more than 60% of even cycles.

It is important to note that the value A_{max} is obtained for $x_i = 87381$, $I = 1$, $P = 18$, this combination also generates the smallest relation *I/P* present in the studied base, on the other hand it represents the largest relation *P/C* as can be seen in the last graph of Figure 9.

It is also observed that the value of A which depends on x_1 accounts for more than 79% of the value of x_n.

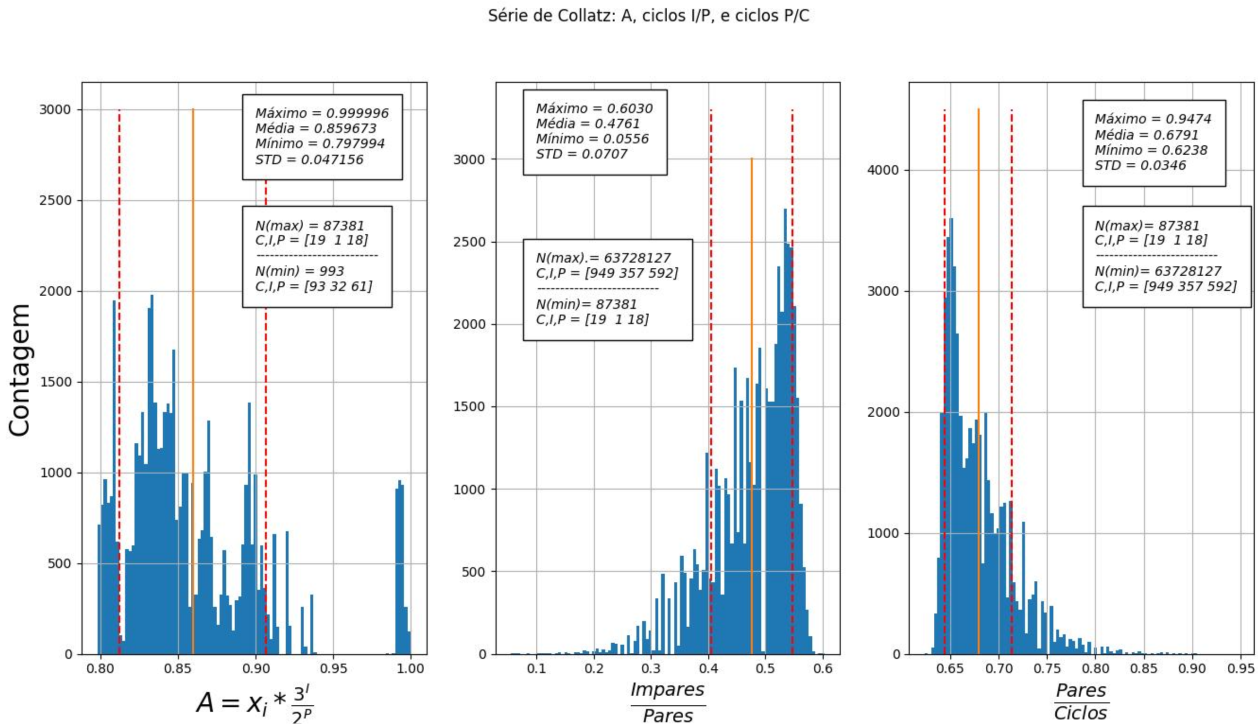


Figure 9: Distribution graphs of A, $\frac{I}{P}$, $\frac{P}{C}$

8. Maximum value for x_i

It is assumed that in the Collatz Sequence there is a maximum value where $x_i = x_m$ from which any subsequent (or previous) value will be less than this maximum x_m, it can be concluded that the maximum value is even, and it will be divided by a number 2^γ. Previously, it was seen that *collatz_d*(7) produces the following output:
collatz_d(7) = [7, 22, 11, 34, 17, 52, 13, 40, 5, 16, 1]

The number 52 corresponds to the maximum (in this sequence), which will later be divided by $2^\gamma = 22 = 4$ becoming the odd number 13. Once a maximum is reached, the subsequent numbers will necessarily be smaller, evidencing a decrease in the series. If the same series were only of increases, it would not be limited in a final cycle, that is, it would tend to infinity beyond what is observed that according to Lagarias [6] heuristic predictions made using tested probabilistic models indicate that this factor is on average lower than ONE, thus converging in successive iterations to the final cycle 1 → 4 → 2 → 1.

Considering the set of natural numbers $\mathbb{N} = \{0, 1, 2, 3, ...\}$ and the set of non- negative even numbers $\mathbb{P} = \{0, 2, 4, 6, ...\}$, it can be seen that both have the same Car- dinality 28 $| \mathbb{N} |=| \mathbb{P} |= \aleph_0$ [9], the same occurs with the set of numbers that are powers of $2\mathbb{E}_2 = \{0, 2, 4, 8, 16, ...\}$, the set of odd numbers I also has the same cardinality of $\mathbb{N}$. In short, when dealing with infinite (countable) sets that have the same cardinality ($| \mathbb{N} |=| \mathbb{P} |=| \mathbb{E}_2 |=| \mathbb{I} |= \aleph_0$), it follows that the probability distribution for ∀ and the number xi present in these sets are equivalent, since the same sets are equipotent, making any prediction difficult to make or that it presents objective trends (taking into account infinite countable sets). In the previous item, it was seen that every Collatz series has a final cycle 1 → 4 → 2 → 1 and consequently a maximum establishing for each series a particular set of finite numbers, having the same amount of odd and even numbers [6] (before reaching the number 1).

Consider also the set $\mathbb{D}_7$ = [7, 22, 11, 34, 17, 52, 13, 40, 5, 16]29 whose cardinality is:
$|\mathbb{D}_7|$= 10, it is clear that the amount of even numbers is identical to the amount of odd numbers present in $\mathbb{D}_7$, both subsets (even $\mathbb{D}_{7P}$ and odd $\mathbb{D}_{7I}$) have the same finite cardinality, in item 2 previously seen and adapted here we have:

$$x_i \in \mathbb{D}_{7I} = [7, 11, 17, 13, 5] = [x_1, x_3, x_5, x_7, x_9] \mid i = \{1, 3, 5, 7, 9\}$$

$$x_i \in \mathbb{D}_{7P} = [22, 34, 52, 40, 16] = [x_2, x_4, x_6, x_8, x_{10}] \mid i = \{2, 4, 6, 8, 10\}$$

A fact that allows us to assume that any $x_i \in \mathbb{D}_7$ has the same probability of occurrence, that is, there is the same number of increases and decreases in the series, however, it is observed that in the decreases or decrements made by divisions by even numbers there are compound divisions, dividing more than once by two, which is in accordance with the observations seen in the same item 2, that is:

$$\rho_i = \frac{x_{(2i+1)}}{x_{(2i)}}, \text{ being: } \rho_i = [\tfrac{1}{2^1}, \tfrac{1}{2^1}, \tfrac{1}{2^2}, \tfrac{1}{2^3}, \tfrac{1}{2^4}] = [\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}].$$

Therefore, due to the various cumulative divisions, an even number present in the Collatz sequence will be divided more than once by the number two. The function *r_collatz*(*x*) (Figure 4) already presented the composition of the sequence in terms of [C, I, P]30, that is, C_{iclos}, I_{mpares} and P_{ares}, where we can see that: $C = I + P$, and $P > I$.

The result of the function *r_collatz*(7) is [16,5,11], the terms C_{iclos}, I_{mpares} and P_{ares} correspond exactly to the cardinality of the subsets of the series *collatz_seq*(7), that is $|\mathbb{C}_7|=|\mathbb{I}_7|+|\mathbb{P}_7|= 16$ where $|\mathbb{I}_7|= 5$ and $|\mathbb{P}_7|= 11$ for the respective sets:
$\mathbb{C}_7$ = [7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2],
$\mathbb{I}_7$ = [7, 11, 17, 13, 5] and $\mathbb{P}_7$ = [22, 34, 52, 26, 40, 20, 10, 16, 8, 4, 2].

Table 2 shows the distribution of numbers that are powers of 2 (less than 5000) as a function of the exponent γ, remembering that any even number within the series (result of $3 \times x_i + 1$) will necessarily be divided by an even number of the type 2^γ , highlighting the non-zero probability that it (result of $3 \times x_i + 1$) will be divided more than once by 2.
The function *nPares*(*x*) (Figure10) provides the values from table 2 for x = 5000, in the console of the Python: *divisores* = *np.array*(*nPares*(5000))[: 2] which will give the following answer:
divisores[0] = *array*([2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096])
and
divisores[1] = *array*([1250, 625, 313, 156, 78, 39, 20, 10, 5, 2, 1, 1]])

γ	divisor	occurrence	occurrence (%)
1	2	1250	50
2	4	625	25
3	8	313	12.52
4	16	156	6.24
5	32	78	3.12
6	64	39	1.56
7	128	20	0.80
8	256	10	0.40
9	512	5	0.20
10	1024	2	0.08
11	2048	1	0.04
12	4096	1	0.04

Table 2: Even divisors where: 2^γ <= 5000

Whose sum (sum(divisors[1])) is 2500, that is, there are 2500 available divisors, these being powers of 2, the function nPaires(numero)[2:] indicates the original number and, if necessary, the immediately higher pair (to be used if the number informed be odd).

```python
def nPares(num):
    import numpy as np
    conta = int(num)
    #valores=[["n","C","I","P","nucleo"]]
    valores=[]
    while (conta > 1):
        c_num = str(conta)
        resp = divByPot2(c_num)
        valores.append(resp[1])
        conta -= 2
        if conta <= 1:
            break
    matrizV = np.array(valores)
    expoente, ocorrencia = np.unique(matrizV, return_counts=True)
    return expoente, ocorrencia
```

Figure 10: Code: *nPares*()

Previously in table 1 we presented some results of the relationship between I and P, that is: between the number of Odd cycles in relation to the number of Even cycles [31].
where 'apparently' we have:

$$\lim_{x_1 \to \infty} r\_collatz(x_1) \longrightarrow \frac{I}{P} \approx 0.50$$

Table 3 shows the distribution for number NN8, where:

$$[C, I, P], E\gamma = r\_collatz1(NN8) \longrightarrow [74671, 25016, 49655], E\gamma.$$

The matrix $E\gamma$ (for the number NN8) is composed of 25016 elements, which correspond to 2^{γ}, γ in the first column of the table. 3 [32,33].

γ	divisor	occurrence	$\sim \div 2$	occurrence (%)	Heuristic[6]
1	2	12536	12536	50.11	1.2253
2	4	6350	12700	25.38	1.1390
3	8	3137	9411	12.54	1.0072
4	16	1485	5940	5.94	0.9119
5	32	754	3770	3.01	0.8491
6	64	379	2274	1.52	0.8107
7	128	185	1295	0.74	0.7885
8 512	62	558	0.25	0.7675 0.7596	
13	8192	4	52	0.02	0.7587
14	16384	3	42	0.01	0.7579
	totals	25016	49655	100	—

Table 3: Even divisors: r_collatz(NN8)

A number x_i (even) will be divided by 2^{γ} , it can be seen that the number of divisions by 2γ is exactly equal to the number of odd operations, which is in agreement with the function *collatz_d*(x_1), that is, 50% of the operations occur on odd numbers and 50% of the operations occur on even numbers, however the distribution of division operations follows an exponential distribution, column occurrence (%) and Figure 11. The last column of Table 3 refers to the heuristic argument presented by Lagarias[6], in which the multiplicative factor (MF) between two consecutive odd integers should be $\sim 3/4 < 1$, it can be seen in Table 3 that MF $\cong$ 0.7579: *"this heuristic argument suggests that, on average, iterations in a trajectory tend to decrease in size, so that there should be no divergent trajectories"*(translated by the author).

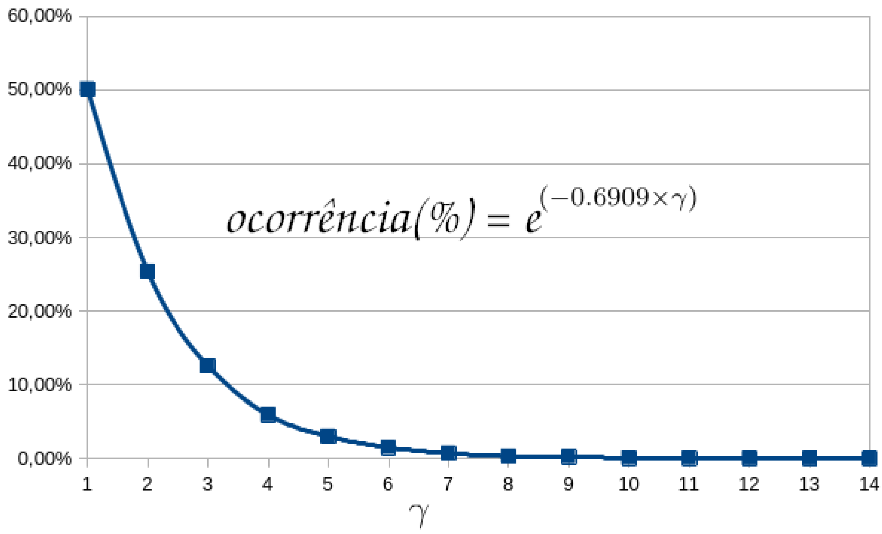


Figure 11: *Occurrence of Pairs (r_collatz1(NN8))*

The graph in (Figure 11) illustrates the exponential distribution of even numbers according to occurrence (%) in Table 3 (NN8) and approximated equation. It can be seen from Table 3 that the probability of an even number subsequent to the operation $3 \times x_i + 1$ being divided by TWO is 100%[34], and in this process being divided again by a power of TWO, that is, being divided again by 2^γ, where γ[35] $\in \mathbb{N}^*$ between 22 and 2∞ is given by:

$$\mathbb{P}(1.51 \leqq \gamma \leqq 14) = \int_{1.51}^{14} e^{(-0.6909\times\gamma)} \simeq 50\%$$

Table 3 also shows that Σ(*occurrences*) when $\gamma > 1$ is $\Sigma \simeq 50\%$, which also implies that for 25016 increases there will be 49655 divisions by 2, which on average corresponds to: $2^{\left(\frac{49655}{25016}\right)} \simeq 3.9584$, and this average value in turn approaches the heuristic argument presented and proposed by Lagarias, that is: $\frac{3}{4} \simeq \frac{3}{3.9584} \simeq 0.75$ [36].

It is worth noting that the occurrence of divisions by two, that is, divisions by 2^γ, is in accordance with the exponential distribution of numbers $\in \mathbb{N}^*$, as also seen in Table 2.
The following graph in Figure... shows the average of the exponential factor $e^{(-0.6909\times\gamma)}$ for $2^{68} + 1 \leqslant x_i \leqslant 2^{68} + 10001$ (only the odd numbers).

9. Conclusion

The Big Question
Do all natural numbers when subjected to the Collatz sequence always end in the cycle 4→2→1?
Statistical evidence and some auxiliary programs point in this direction, but they are not emphatic in admitting such a conclusion. This is expected since the operation is on a set of infinite numbers $\in \mathbb{N}^*$, and the tools used are limited in relation to the internal representation of the numbers supported by the machines used. However, with the help of binary operations on numbers expressed in text form (string), huge numbers such as NN8 were worked on. The operation $3 \times x_i + 1$ and even division by TWO are relatively simple to perform with binary operations and are a great help in overcoming the representation barrier. numerical intrinnsect to current computers. Such results operating on 'large' numbers of the order of 103000 were in accordance with the expected and proclaimed results.

It is worth highlighting based on what was previously presented in item 4.1.2 and appendix A:

In the Collatz Conjecture there is only one limit cycle formed by the stable points: 1→4→2→1.

The non-existence of other internal limit cycles in the Collatz Conjecture, and also in accordance with the equation (12) transcribed here: $\Psi + \Omega = 1$, remembering that $0 < \Psi < 1$ and $0 < \Omega < 1$ result in:

Any and all numbers $x_i \in \mathbb{N}^*$ (even xi → ∞) when subjected to the Collatz Conjecture will invariably end in the limit cycle formed by the stable points: 1→4→2→1.

The Collatz Conjecture is definitely not only a challenge, it is also a fertile field for using the tools available in mathematics. At each stage, new tools and/or observations (sometimes previously neglected) are present. In one of these surprises, we can say:
Any number $x_i = 2^\sigma - 1, \mid \sigma \in \{0, 2, 4, 6, 8, ...\}$ will always be divisible by three.

The function *divide_3*(*expI, expF*) helps in the verification, but the demonstration of such a statement is beyond the scope of this work, and will be left for a possible future work!

References

1. Baker, G. L., & Gollub, J. P. (1996). *Chaotic dynamics: an introduction*. Cambridge university press.
2. Barina, D. (2021). Convergence verification of the Collatz problem. *The Journal of Supercomputing, 77*(3), 2681-2688.
3. Grauer, J. A. (2021). Analogy Between the Collatz Conjecture and Sliding Mode Control.
4. Hartnett, K. (2019). Mathematician Proves Huge Result on Dangerous Problem. *Quanta magazine.*
5. Hayes, B. (1984). On the ups and downs of hailstone numbers. *Scientific American, 250*(1), 10-13.
6. Lagarias, J. C. (1985). The 3 x+ 1 problem and its generalizations. *The American Mathematical Monthly, 92*(1), 3-23.
7. Monteiro, L. H. A. (2002). *Sistemas dinâmicos*. Editora Livraria da Física.
8. Helene, O. A., & Vanin, V. R. (1991). *Tratamento estatístico de dados em física experimental*. Editora Blucher.

Appendix A

Comparison Between Ω Calculated and Approximated

The text in this appendix was highlighted and treated as complementary, serving to support the observations seen in the item 4.1.2, in which when $x_k \to \infty$ it is observed that $\Psi = \frac{3^I}{2^P} \to 0$ and due to this fact the value **1** present in the formula $3 \times x_i + \mathbf{1}$ was disregarded, since this is much smaller than the product $3 \times x_1, (1 <<< (3 \times x_1))$. Therefore, the value of Ω was calculated using both methods (I) and (II), namely:

$$\underbrace{\sum_{j=(I-1)}^{0} \left(3^j * \left[\prod_{i=(I-j)}^{I} (\rho_i) \right] \right)}_{(B)+(C)+(D)=\Omega} = \Omega \qquad (I)$$

$$\lim_{(x^*,\Psi)\to(\infty,0)} \left(\frac{\Omega}{1-\Psi}\right) = \Omega \quad \therefore \quad \left(\frac{\Omega}{1-\Psi}\right) = \Omega \qquad (II)$$

The following graphs using the function *omega_1e2*$(x_i, x_n, salto)$ show the modulus of the difference between the values obteined in (I) and (II), i.e. $|\Omega(I) - \Omega(II)|$:

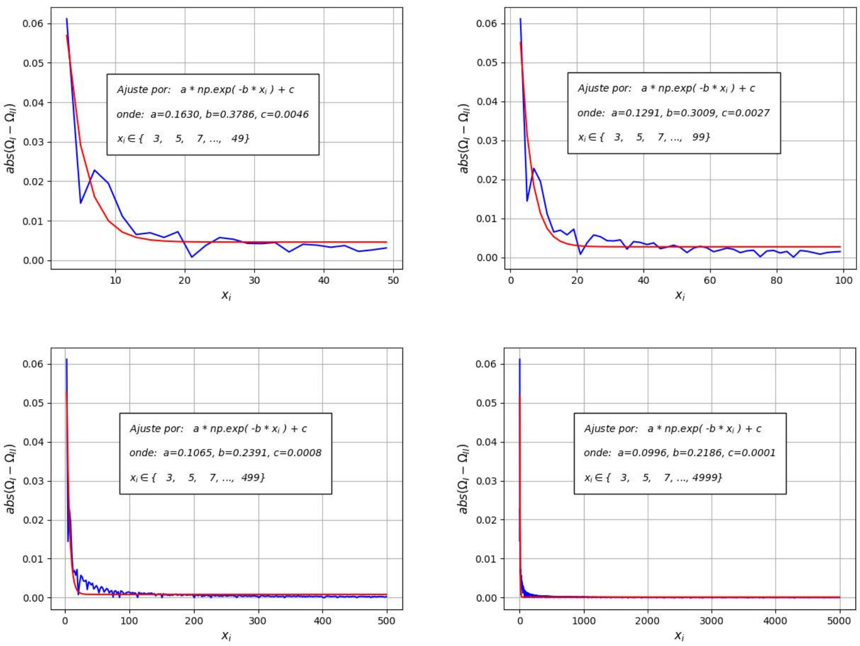


Figure 12: Graphs: $abs(\Omega(I) - \Omega(II)) \times x_i$

In Fig. 12, the sequences in blue are the calculated values of the modulus of the differences, while in red the approximated exponential trend line according to values of **a, b** and **c** in each graph, note that the residual value **c** when $x_i \to \inf$, we have $\lim_{x_i \to \infty} \mathbf{c} \to 0$, demonstrating that equations (I) and (II) lead to similar results, and that such an approximation (when $x_i \to \inf$) is acceptable as well as the limits:

$$\lim_{(x^*,\Psi)\to(\infty,0)} \left(\frac{\Omega}{1-\Psi}\right) = \Omega, \qquad \lim_{x^* \to \infty} (x^*) \longrightarrow \Omega$$

The following graphs (Fig. 13) present both methods for calculating Ψ, (I) and (II), note that when $x_i \to \inf$ the value of Ψ also tends to zero ($\Psi \to 0$).

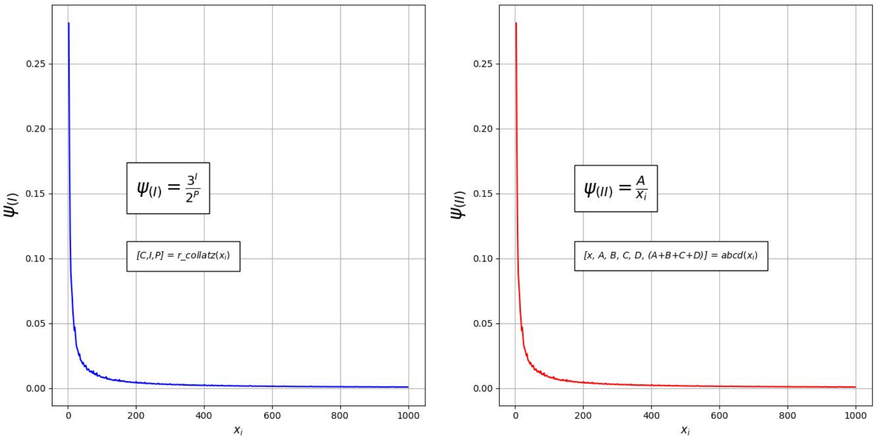


Figure 13: Graphics: $\Psi \times x_i$

Thus, with what has been exposed in this appendix and seen previously, it is evident that $\nexists$ is any value of $x^* \neq 1, x^* \neq 2, x^* \neq 4 \mid x^* \in \mathbb{N}$ that satisfies the equation (12), which leads to the conclusion that $\exists$ is just and exclusively a limit cycle in the Collatz Conjecture.

In addition, hypothetically consider the equation (12) assuming that $\Psi \to 1$ (which is in disagreement with the graphs in Fig. 13), as seen previously $\Psi x_k + \Omega = x_k$ we will necessarily have that $\Omega \to 0$, we obtain that: $\Psi x_k + 0 = x_k$ or $\Psi x_k = x_k$

Given that $\Psi = \frac{3^I}{2^P}$, it follows that $x_k \simeq x_k \times 2^{(I * \log_2(3) - P)}$, for the expression $2^{(I * \log_2(3) - P)} = 1$ to be true we have: $I * \log_2(3) = P$, remembering $\Omega = B + C + D$ where $B = \frac{\Psi}{3}$, it follows that (when $\Psi \to 1, \Omega \to 0$) $\Omega = \frac{\Psi}{3} + C + D \therefore$ we have that $\Omega > \frac{1}{3}$ which contradicts the hypothesis previously formulated. The following graphs (Fig. 14) illustrate that: $\Omega > 0$ and $\Psi \neq 1$, reinforcing that the hypothesis $\Psi \to 1, \Omega \to 0$[37] is not valid.

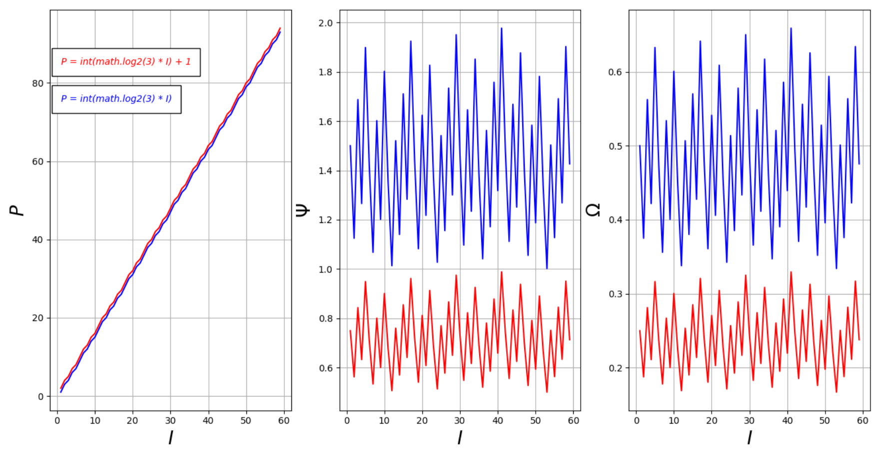


Fig. 14 – Graphs: $P \times I, \quad \Psi \times I, \quad \Omega \times I$

Figure 14: Graphs: $P \times I$, $\Psi \times I$, $\Omega \times I$

Appendix B

Relationship between I and P

1. Equation (4) Possibilities as a Function of I, P

The term (C) of the equation (4) presents a certain complexity in the formation of the sets that contain the values of γ_i, having a distribution of the elements γ_i in arrangements $(AR(\gamma_i) = (I-1)!)$ different[38], the function $r\_collatz1('7')$ displays as the answer:

$$[C, I, P, E\gamma] = r\_collatz1('7') \longrightarrow ([16, 5, 11], \underbrace{[1, 1, 2, 3, 4]}_{(E\gamma)})$$

The original set $E\gamma = [1, 1, 2, 3, 4]$, which are the exponents to be applied to $2^{-E\gamma}$, resulting in the set $\rho_i = [\frac{1}{2^1}, \frac{1}{2^1}, \frac{1}{2^2}, \frac{1}{2^3}, \frac{1}{2^4}] = [\frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}]$. The set $E\gamma$ can be rewritten in several ways, i.e. in $(I-1)!$ arrangements, without changing the result of $\prod_{i=1}^{n}(\rho_i)$.

However, such arrangements will impact the term (C) of equation (4) as can be seen below for two specific sets [39]:

$$\rho_i^E = [\tfrac{1}{2^1}, \tfrac{1}{2^1}, \tfrac{1}{2^2}, \tfrac{1}{2^3}, \tfrac{1}{2^4}], \qquad \rho_i^M = [\tfrac{1}{2^3}, \tfrac{1}{2^2}, \tfrac{1}{2^1}, \tfrac{1}{2^1}, \tfrac{1}{2^4}]$$

Rremembering the item (C) of the equation (4), where $I = n = 5$:

$$\underbrace{\sum_{j=(n-2)}^{1}\left(3^j * \left[\prod_{i=(n-j)}^{n}(\rho_i)\right]\right)}_{(C)}$$

Adopting ρ_i^E, ρ_i^M seen previously, it is possible to calculate the dyadic fractions $(\alpha), (\beta)$ and (δ) for the two series of ρ:

$$\underbrace{3^3 * [\frac{1}{2^1} * \frac{1}{2^2} * \frac{1}{2^3} * \frac{1}{2^4}]}_{(\alpha^E)} + \underbrace{3^2 * [\frac{1}{2^2} * \frac{1}{2^3} * \frac{1}{2^4}]}_{(\beta^E)} + \underbrace{3^1 * [\frac{1}{2^3} * \frac{1}{2^4}]}_{(\delta^E)}$$

$$\underbrace{0.0263671875}_{(\alpha^E)} + \underbrace{0.017578125}_{(\beta^E)} + \underbrace{0.0234375}_{(\delta^E)} = \underbrace{0.0673828125}_{(C^E)}$$

$$\underbrace{3^3 * [\frac{1}{2^2} * \frac{1}{2^1} * \frac{1}{2^1} * \frac{1}{2^4}]}_{(\alpha^M)} + \underbrace{3^2 * [\frac{1}{2^1} * \frac{1}{2^1} * \frac{1}{2^4}]}_{(\beta^M)} + \underbrace{3^1 * [\frac{1}{2^1} * \frac{1}{2^4}]}_{(\delta^M)}$$

$$\underbrace{0.10546875}_{(\alpha^M)} + \underbrace{0.140625}_{(\beta^M)} + \underbrace{0.09375}_{(\delta^M)} = \underbrace{0.33984375}_{(C^M)}$$

It becomes clear from previous considerations that a large part of the complexity in solving the equation (4) consists in solving the term (C) of the same, since the possible arrangements with the coefficients γ_i are of the order of $(I-1)!$.

2. The Conjecture Considering Real Numbers

Certainly considering even and odd numbers in the case of $x_i >= 0 | x_i \in \mathbb{R}^*$ makes little sense since these numbers, as they can be fractional or irrational, do not present parity. The conjecture must be adapted, and in this case only:

$$collatz\_reais(x_i)^{40} \implies x_n = \frac{3 \times x_i + 1}{4}$$

The factor $\frac{3}{4}$ is in agreement with Lagarias' heuristic argument[6], in which the multiplicative factor (MF) between two subsequent numbers must be $\sim \frac{3}{4} < 1$, in the graph in Fig. 14 the cycle is observed limit defined by the circle with radius **UM** to which all sequences converge[41]. Remembering that according to **Monterio**[7] "to know the characteristics of an orbit it is necessary to know its eigenvalue λ, which corresponds to the product of each eigenvalue relative to the fixed points x_i^* in the orbit", as follows:

$$\lambda = \frac{df(x)}{dx}|_{x_i^*} = \frac{3}{4}$$

As $\lambda < 1$ the orbit is stable, which allows us to state that such a cycle repeats indefinitely once the value **UM** is reached.

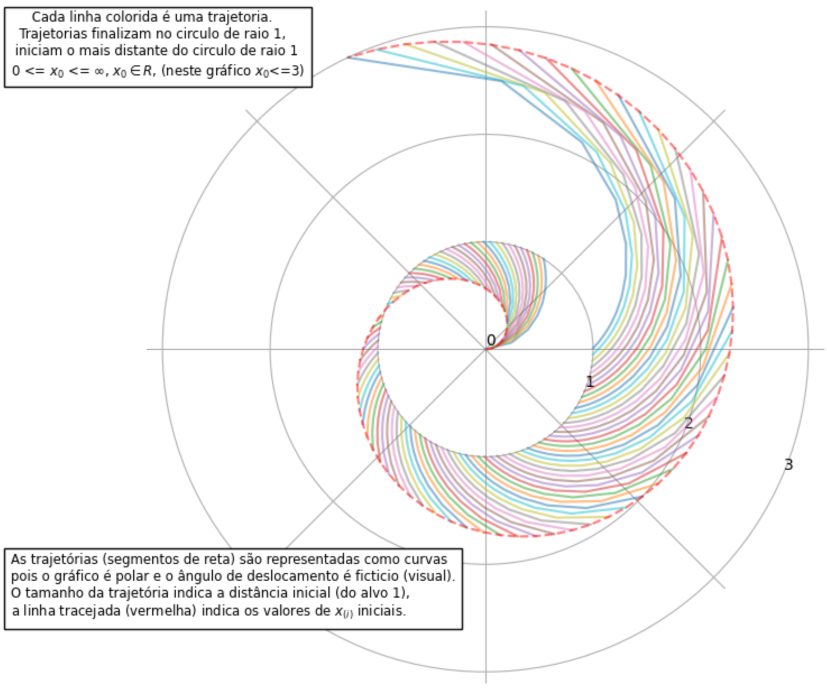


Figure 15: Polar Graph: $\frac{(3 \times x_i + 1)}{4}$

3 A Practical Example (Brute Force)

Consider the number 753257 whose graphs were seen in Fig. 7 item 4.2, from the equation (5) doing: $\prod_{i=1}^{n}(\rho_i) = \prod_{i=1}^{P}(\frac{1}{2}) = \frac{1}{2^P}$

we will have:

$(A) \implies h(753257) = 753257 * \frac{3^I}{2^P}$ $\qquad$ $(= 0.80984638847774229)$

Remembering that this value *"generally"* is ≥ 0.78 (data in table 1), we must look for the values of I and P that meet the above, also observing the equation (11), the function *acha_ABD(n,m,c)* available in the file Collatz_Files constructed in accordance with the limit range (two lines in green) seeks to find the values that satisfy the condition above, it is important to note that up to 400 possible answers of $I \times P$ can be presented, since a region was delimited where the confidence index for the study base is 100%. [42]

Executing the function *acha_ABD(n,m,c)* (n=753257, m=0.5, c=0.194725, **m** and **c** according to equation 16) we obtain the answers (only 14 are presented) [43], in the following list the correct one is highlighted:

Item	A	Pares	Impares	A+B+D	I/P(%)
1	0.886717	72	33	0.949217	0.458333
2	0.841688	80	38	0.904189	0.475000
3	0.798946	88	43	0.861447	0.488636
4	0.898814	91	45	0.961315	0.494505
5	0.853172	99	50	0.915672	0.505051
6	0.809846	107	55	0.872347	0.514019
7	0.911077	110	57	0.973578	0.518182
8	0.864812	118	62	0.927312	0.525424
9	0.820895	126	67	0.883396	0.531746
10	0.923507	129	69	0.986008	0.534884
11	0.876610	137	74	0.939111	0.540146
12	0.832095	145	79	0.894595	0.544828
13	0.936107	148	81	0.998607	0.547297
14	0.789840	153	84	0.852341	0.549020
...	...	...	...	...	...

Table 4: *acha_ABD*(753257, 0.5, 0.194725)

Among the possibilities that the function *acha_ABD(n,m,c)* presented, the correct one corresponds to item 6, that is $P = 107, I = 55$ (r_collatz(str(753257)) $\Longrightarrow$ $[162, 55, 107]$). This method (via brute force) presents many candidate values, and only serves to show the most appropriate one among these results without having to calculate the complete Collatz Conjecture as presented in the equation (2).

Appendix C
Setting up the Python / PyCharm environment
Creating the project

Once the Python environment is set up, the version currently used is: Python 3.11.6 in a Linux environment 6.5.12-200.fc38.x86_64 and the IDE interface PyCharm 2022.1.3 (Community Edition), let's create the initial project (those who have already done so can skip this step). Start the PyCharm program and create a new project the so-called Python_Collatz as shown in the following figure:

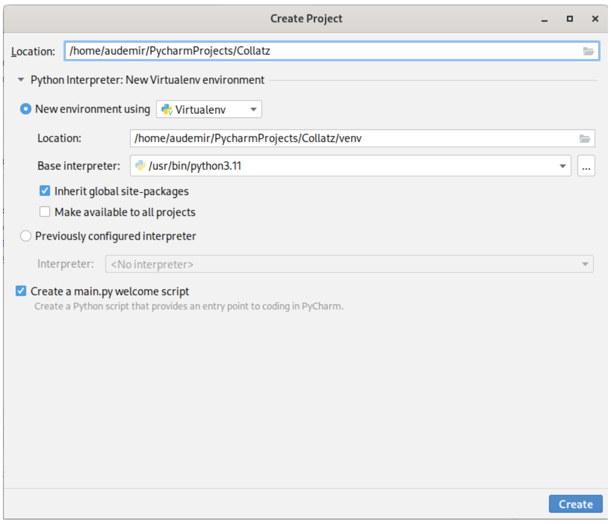


Figure 16: Creating the Project

When creating the project, the initial screen with the main.py file (initial) is displayed (Fig.16). Note that the print_hi(name) function is automatically created and triggered internally in the main code.

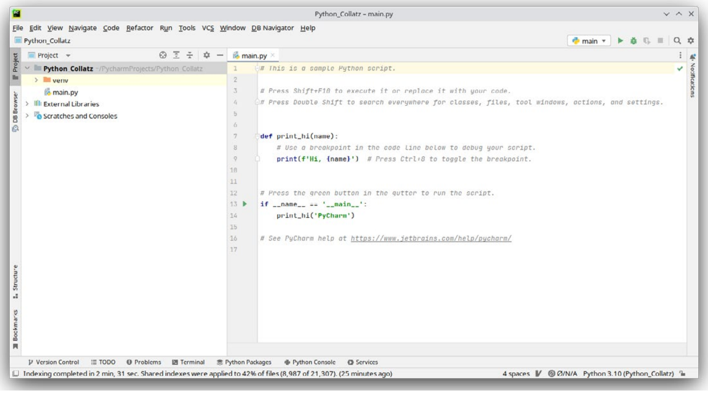


Figure 17: Main.py file (initial)

By making the changes to the code (example), our main.py file (initial) will look like this:

```
1  import matplotlib.pyplot as plt
2  import pandas as pd
3  import numpy as np
4  import math
5  import sys
6  import os
7  import csv
8
9  from Collatz_Files import *
10
11 # Press the green button in the gutter to run the script.
12 if __name__ == '__main__':
13     n = math.log(int(f'{0b101010000011001:#0}'))
14     print(n)
```

Figure 18: Code: main.py

Lines 13 and 14 shown in the code in Figure 17 were included only so that it is possible to test the environment beforehand.

1.1 Including Files

Before running the main.py program we must include other files:

___init___.py[44],

BaseDados.py,

funcoes.py,

Collatz_Files.py

The following list refers to the __init__.py file and adjusts the environment to import local functions developed in files separate from main.py within the working directory.

```
# todo comentário em Python inicia com o caracter #
# este arquivo tem o objetivo de indicar para o programa
# main.py a localização
# das funções auxiliares

from BaseDados import *
from funcoes import funcoes
from Collatz_Files import Collatz_Files
```

Figure 19: *Code: ___init___.py*

The files cited: main.py, funcoes.py, Collatz_Files.py, BaseDados.py and conjecturas.py present almost two thousand lines of code, the probability of incorrect typing is very high, due to this the files will be made available in electronic form when requested by email to the author, after publication they will be available in the directory indicated here.

Appendix D
Conflict of Interest

This article is based on research and various articles of a public nature, the developments and information made public are acknowledged in this article, applying due reference to the same authors cited. This is an analysis of a public domain topic, which does not generate conflicts of interest in general!