A proof of the Collatz conjecture

by

Fabrice Trifaro

Abstract. Using a comprehensive approach, this paper aims to demonstrate, clearly and rigorously, the validity of the Collatz conjecture. To this end, the original $3n+1$ iteration is reformulated by isolating the odd terms into sequences referred to as R-Cz sequences. These sequences are analyzed through their structural properties and their distribution among the odd natural numbers. As a first essential result, it is shown that they do not admit non-trivial cycles: the only possible cycle is the trivial one, of value and length 1.

Two independent proofs that all R-Cz sequences converge are then presented. The first, combinatorial in nature, relies on the finiteness of intervals that could possibly separate terms of the sequences. The second, set-theoretic, is based on a contradiction between the countability of the odd integers and the uncountable cardinality of the hypothetical divergent R-Cz sequences. Both methods lead to the same conclusion: all Collatz sequences eventually enter the cycle $(1, 4, 2)$.

Contents

1 Introduction

First of all, let us define a Collatz sequence. Let $\{u_n\}_{n\in\mathbb{N}}$ be the sequence such that $u_0 = p$, where $p\in\mathbb{N}^*$, and such that:

$$\text{for } n\in\mathbb{N}^*, u_n = \begin{cases} 3u_{n-1}+1 & \text{if } u_{n-1} \text{ is odd} \\ \frac{u_{n-1}}{2} & \text{if } u_{n-1} \text{ is even} \end{cases}$$

According to the conjecture, there exists $l\in\mathbb{N}^*$ such that $u_l = 1, u_{l+1} = 4, u_{l+2} = 2, u_{l+3} = 1, u_{l+4} = 4, u_{l+5} = 2$, and so on. In other words, from rank l the sequence enters a cycle that repeats the numbers $1, 4, 2$ *ad infinitum* (see [2], [3], [4] and [6] for the background of this conjecture). We can express this sequence in another way, indeed, if p is odd then we have:

$$u_1 = 3p+1 \text{ is even,}$$

$$u_2 = \frac{3p+1}{2^1}, \text{ if } u_2 \text{ is even then } u_3 = \frac{3p+1}{2^2},$$

$$\dots,$$

$$\text{until } u_{1+\alpha_0} = \frac{3p+1}{2^{\alpha_0}} \text{ is odd.}$$

Let :

$$v_0 = u_{1+\alpha_0} = \frac{3p+1}{2^{\alpha_0}}$$

Where α_0 is the exponent corresponding to the number of times u_1 must be divided by 2 to obtain an odd number. Repeating the same process, we have:

$$u_{1+\alpha_0+1} = 3u_{1+\alpha_0} + 1 \text{ is even,}$$

$$u_{1+\alpha_0+1+1} = \frac{3u_{1+\alpha_0}+1}{2^1}, \text{if } u_{1+\alpha_0+1+1} \text{ is even then } u_{1+\alpha_0+1+2} = \frac{3u_{1+\alpha_0}+1}{2^2},$$

$$\dots,$$

$$\text{until } u_{1+\alpha_0+1+\alpha_1} = \frac{3u_{1+\alpha_0}+1}{2^{\alpha_1}} \text{ is odd.}$$

Let:

$$v_1 = u_{1+\alpha_0+1+\alpha_1} = \frac{3u_{1+\alpha_0}+1}{2^{\alpha_1}}$$

Where α_1 is the exponent corresponding to the number of times $u_{1+\alpha_0+1}$ must be divided by 2 to obtain an odd number. By reformulating v_1, we have:

$$v_1 = \frac{\left(3\left(\frac{3p+1}{2^{\alpha_0}}\right)+1\right)}{2^{\alpha_1}} = \frac{3\left(3p+1\right)+2^{\alpha_0}}{2^{\alpha_0+\alpha_1}}$$

And by an easily verifiable recurrence (see Appendix 6.1), we obtain that for all $l \in \mathbb{N}^*$:

$$v_l = \frac{3^l\,(3p+1) + \sum_{i=0}^{l-1}(3^{l-1-i}(2^{\sum_{j=0}^{i}\alpha_j}))}{2^{\sum_{i=0}^{l}\alpha_i}} = \frac{3^l\,(3p+1)}{2^{\sum_{i=0}^{l}\alpha_i}} + \sum_{i=0}^{l-1}\frac{3^{l-1-i}}{2^{\sum_{j=i+1}^{l}\alpha_j}}$$

The resulting sequence $\{v_l\}_{l\in\mathbb{N}}$ has all its values in $2\mathbb{N}+1$, and it is equal to the sequence $\{u_n\}_{n\in\mathbb{N}}$ without the first term and the even-valued terms of the latter. Thus, the cycle of length 3 and values $(1,4,2)$ of the sequence $\{u_n\}$ corresponds to the cycle of length 1 and value (1) of the sequence $\{v_l\}$.

If p is even, there exists $\alpha \in \mathbb{N}^*$ such that $p = 2^\alpha q$, where $q \in 2\mathbb{N}+1$, and it is sufficient to replace p with q in the expression of the term v_0, which does not change the demonstration.

Definition 1.1. *Let $f = 2^\alpha p$, where $\alpha \in \mathbb{N}$ and $p \in 2\mathbb{N}+1$. The sequence $\{v_l^p\}_{l\in\mathcal{N}}$, the so-called reformulated Collatz sequence, is defined as follows:*

$$v_l^p = \begin{cases} p & \text{if } l=-1 \\ \frac{3p+1}{2^{\alpha_0}} & \text{if } l=0 \\ \frac{3^l(3p+1)}{2^{\sum_{i=0}^{l}\alpha_i}} + \sum_{i=0}^{l-1}\frac{3^{l-1-i}}{2^{\sum_{j=i+1}^{l}\alpha_j}} & \text{if } l>0 \end{cases}$$

Where p in the sequence $\{v_l^p\}$ means that the sequence depends on p, α_0 is the exponent such that v_0^p is odd, the α_i, for $i \in \{1,\dots,l\}$, are the exponents such that $v_i^p = \frac{3v_{i-1}+1}{2^{\alpha_i}}$ is odd, and $\mathcal{N} = \mathbb{N} \cup \{-1\}$.

Throughout the paper, this type of sequence will be called *R-Cz sequence*, and we will consider as many different sequences as there are different values of p.

Definition 1.2. *Let $R_{Cz} = \left\{\{v_l^{p_j}\}\right\}_{(j,l)\in\mathbb{N}\times\mathcal{N}}$ be the set of all R-Cz sequences, where $p_j \in 2\mathbb{N}+1$. Let R_{Cz}^{c} and R_{Cz}^{d} be the sets of convergent and non-convergent R-Cz sequences, respectively, then $R_{Cz} = R_{Cz}^{\mathrm{c}} \cup R_{Cz}^{\mathrm{d}}$.*

Theorem 1.1. *Let $f : 2\mathbb{N}+1 \to R_{Cz}$ be a function defined for all $p \in 2\mathbb{N}+1$ as follows:*

$$f(p) = \{v_l^p\}$$

Then f is bijective.

Proof. Let $(p,p') \in (2\mathbb{N}+1)^2$ such that $p \neq p'$. Then, since $\{v_l^p\}$ and $\{v_l^{p'}\}$ are distinct R-Cz sequences (at least by their first term), we have:

$$f(p) = \{v_l^p\} \neq \{v_l^{p'}\} = f(p').$$

Hence, f is injective. Let $\{v_l^p\} \in R_{C_z}$, by definition we have $f(p) = \{v_l^p\}$. Therefore, f is bijective. □

2 Cycles of R-Cz sequences

We are going to study whether an R-Cz sequence can enter a cycle, under what conditions and what cycles are possible. We will begin with cycles of lengths 1, 2 and 3, and then move to study the general case. The mathematical expression of the terms in the sequence, involving the exponents ${\alpha_i}^{(\mathrm{i})}$, is the same as the one presented in the introduction, and to simplify matters, we will start indexing the exponents from 0.

The common condition to the cycles of lengths 2, 3 and t lies in the fact that the values within a cycle must be distinct. Otherwise, due to the definition of the sequence $\{v_l^p\}$, the repetition of a value would result in a cycle of a shorter length.

2.1 Cycle of length 1

It will be shown that the only cycle of length 1 that any R-Cz sequence $\{v_l^p\}$ can enter, is the cycle of value (1).

Theorem 2.1. *For all $p \in 2\mathbb{N}+1$, the only cycle of length 1 that the R-Cz sequence $\{v_l^p\}$ can enter from a certain rank is the cycle of value* (1).

Proof. The sequence $\{v_l^p\}$ has a cycle of length 1, if there exists $L \in \mathcal{N}$ such that for any $l \geq L, v_{l+1}^p = v_l^p$. Let q the value of the term of rank l in the sequence, then $v_{l+1}^p = v_l^p = q$ if:

$$q = \frac{3q+1}{2^{\alpha_0}} \implies q = \frac{1}{2^{\alpha_0} - 3}$$

This is only possible if $2^{\alpha_0} = 4 \implies \alpha_0 = 2$, and it follows that $q = 1$. Reciprocally, we check that if $v_l^p = 1$, then for all $k \in \mathbb{N}^*, v_{l+k}^p = 1$.

The only cycle of length 1 into which the sequence $\{v_l^p\}$ can enter from a certain rank is therefore the cycle of value (1), which corresponds to the cycle of values $(1, 4, 2)$ in the sequence $\{u_n\}$. □

2.2 Cycle of length 2 or 3

It will be shown that no R-Cz sequence has either a cycle of length 2 or length 3.

Theorem 2.2. *For all $p \in 2\mathbb{N}+1$, the R-Cz sequence $\{v_l^p\}$ has neither a cycle of length 2 nor a cycle of length 3.*

(i) We recall that for all $i \in \mathbb{N}, \alpha_i \geq 1$.

Proof there is no cycle of length 2. The sequence $\{v_l^p\}$ has a cycle of length 2, if there exists $L\in\mathcal{N}$ such that for all $l\geq L, v_{l+2}^p = v_l^p$. Let q be the value of the term of rank l in the sequence, then $v_{l+2}^p = v_l^p = q$ if:

$$q = \frac{3\,(3q+1) + 2^{\alpha_0}}{2^{\alpha_0+\alpha_1}} \implies q = \frac{3 + 2^{\alpha_0}}{2^{\alpha_0+\alpha_1} - 9}$$

This assumes that $2^{\alpha_0+\alpha_1} > 9$, which implies $\alpha_0 + \alpha_1 \geq 4$. For $\alpha_0 + \alpha_1 = 4$, we find that $q = 1$ when $(\alpha_0, \alpha_1) = (2, 2)$, and that q is not an integer for the other values of (α_0, α_1). For $\alpha_0 + \alpha_1 > 4$, q is not integer because $2^{\alpha_0+\alpha_1} - 9 > 3 + 2^{\alpha_0}$, for all $\alpha_0 \geq 1$.

Since we have established that the sequence becomes stationary from rank l when $q = 1$, the sequence $\{v_l^p\}$ cannot enter a cycle of length 2. □

Proof there is no cycle of length 3. The sequence $\{v_l^p\}$ has a cycle of length 3, if there exists $L\in\mathcal{N}$ such that for all $l\geq L, v_{l+3}^p = v_l^p$. Let q be the value of ther term of rank l in the sequence, then $v_{l+3}^p = v_l^p = q$ if:

$$q = \frac{3^2\,(3q+1) + 3.2^{\alpha_0} + 2^{\alpha_0+\alpha_1}}{2^{\alpha_0+\alpha_1+\alpha_2}} \implies q = \frac{9 + 3.2^{\alpha_0} + 2^{\alpha_0+\alpha_1}}{2^{\alpha_0+\alpha_1+\alpha_2} - 27}$$

This assumes that $2^{\alpha_0+\alpha_1+\alpha_2} > 27$, which implies $\alpha_0 + \alpha_1 + \alpha_2 \geq 5$. For $\alpha_0 + \alpha_1 + \alpha_2 = 5$, q is not integer. For $\alpha_0 + \alpha_1 + \alpha_2 = 6$, we obtain that $q = 1$ when $(\alpha_0, \alpha_1, \alpha_2) = (2, 2, 2)$, and q is not integer for the other values of $(\alpha_0, \alpha_1, \alpha_2)$.

Finally, for $\alpha_0 + \alpha_1 + \alpha_2 > 6$, q is not integer because :

$$2\left(2^{\alpha_0+\alpha_1+\alpha_2} - 27\right) > \left(9 + 3.2^{\alpha_0} + 2^{\alpha_0+\alpha_1}\right)$$

Therefore, as in the previous case, we conclude that the sequence $\{v_l^p\}$ cannot have a cycle of length 3. □

2.3 General case

Having demonstrated that no R-Cz sequence can exhibit a cycle of length 2 or 3, we now demonstrate that no R-Cz sequence can have a cycle of length greater than or equal to 4.

Theorem 2.3. *Let $t \geq 4$, for all $p \in 2\mathbb{N}+1$, the R-Cz sequence $\{v_l^p\}$ has no cycle of length t.*

Proof. The sequence $\{v_l^p\}$ has a cycle of length $t\geq 4$, if there exists $L\in\mathcal{N}$ such that for all $l\geq L, v_{l+t}^p = v_l^p$. Let q the value of the term of rank l in the sequence, such that $q > 1$ to exclude the cycle of length 1 and value (1), then $v_{l+t}^p = v_l^p = q$ if:

$$q = \frac{3^{t-1}(3q+1) + \sum_{i=0}^{t-2}\left(3^{t-2-i}\left(2^{\sum_{j=0}^{i}\alpha_j}\right)\right)}{2^{\sum_{i=0}^{t-1}\alpha_i}} \implies q = \frac{3^{t-1} + \sum_{i=0}^{t-2}\left(3^{t-2-i}\left(2^{\sum_{j=0}^{i}\alpha_j}\right)\right)}{2^{\sum_{i=0}^{t-1}\alpha_i} - 3^t}$$

The cyclicity condition for a cycle of length t can also be expressed, such that for any $k\in\mathbb{N}^*$,$v_{l+kt} = v_l$, which is equivalent to:

$$q = \frac{3^{kt-1} + \sum_{i=0}^{kt-2}\left(3^{kt-2-i}\left(2^{\sum_{j=0}^{i}\alpha_j}\right)\right)}{2^{\sum_{i=0}^{kt-1}\alpha_i} - 3^{kt}}$$

Hence, for all $k \geq 2$, $v_{l+t}^p = v_{l+kt}^p$, that is:

$$q = \frac{3^{t-1} + C_1}{2^{e_1} - 3^t} = \frac{3^{kt-1} + C_k}{2^{e_k} - 3^{kt}}$$

Where:

- $e_1 = \sum_{i=0}^{t-1}\alpha_i$ and $C_1 = \sum_{i=0}^{t-2}\left(3^{t-2-i}\cdot 2^{\sum_{j=0}^{i}\alpha_j}\right)$
- $e_k = \sum_{i=0}^{kt-1}\alpha_i$ and $C_k = \sum_{i=0}^{kt-2}\left(3^{kt-2-i}\cdot 2^{\sum_{j=0}^{i}\alpha_j}\right)$

This yields:

$$2^{e_1}\left(3^{kt-1} + C_k - 2^{e_k-e_1}\cdot 3^{t-1} - 2^{e_k-e_1}\cdot C_1\right) = 3^t\left(C_k - 3^{kt-t}\cdot C_1\right)$$

Let :

- $A = 3^{kt-1} + C_k - 2^{e_k-e_1}\cdot 3^{t-1} - 2^{e_k-e_1}\cdot C_1$
- $B = C_k - 3^{kt-t}\cdot C_1$

Finally, we find that the cyclicity condition reduces, for all $k \geq 2$, to the following equation:

$$\mathbf{2^{e_1}}\boldsymbol{A} = \mathbf{3^t}\boldsymbol{B} \tag{2.1}$$

First case $A = 0$ ***and*** $B = 0$
The equation (2.1) can have solutions if $A = B = 0$. However:

$$B = \sum_{i=0}^{kt-2} 3^{kt-2-i}.2^{\sum_{j=0}^{i}\alpha_j} - \sum_{i=0}^{t-2} 3^{kt-2-i}.2^{\sum_{j=0}^{i}\alpha_j} = \sum_{i=t-1}^{kt-2} 3^{kt-2-i}.2^{\sum_{j=0}^{i}\alpha_j} > 0$$

The term B is a series that diverges to infinity as $k \longrightarrow +\infty$, and given that it is always strictly positive, this case is excluded.

Second case $A < 0$ ***and*** $B > 0$
If $A < 0$ then $2^{e_1}A < 3^tB$, this case is also excluded.

Third case $A > 0$ ***and*** $B > 0$
As 2 and 3 are prime numbers, and the factorization of an integer into prime factors is unique, according to the equation (2.1) we have:

$$A = 3^t m \text{ and } B = 2^{e_1} m$$

And because A is odd and B even, $m \in 2\mathbb{N}+1$. Insofar as we are considering the possibility that the sequence $\{v_l^p\}$ can have a cycle length of t, for all $i \in \{0, \ldots, t-1, \ldots, kt-2\}$, according to Definition 1.1, we will have $\alpha_i \in \{\alpha_0, \ldots, \alpha_{t-1}\}$.

Then, since A and B depend on k, the equations $A = 3^t m$ and $B = 2^{e_1} m$ have solutions, if there exist $m \in 2\mathbb{N}+1$, $t \geq 4$ and $(\alpha_0, \ldots, \alpha_{t-1}) \in (\mathbb{N}^*)^t$, such that for all $k \geq 2$, $A = 3^t m$ and $B = 2^{e_1} m$.

Concerning the equation $B = 2^{e_1}$, for $m = 1$, for all $t \geq 4$ and for all $(\alpha_0, \ldots, \alpha_{t-1}) \in (\mathbb{N}^*)^t$, it suffices to take $k = 2$ to obtain $B > 2^{e_1} m$.

Now, suppose there exists $m > 1$, $t \geq 4$ and $(\alpha_0, \ldots, \alpha_{t-1}) \in (\mathbb{N}^*)^t$, such that for some $k \geq 2, B = 2^{e_1} m$, then, given that B diverges to infinity as $k \to +\infty$, and that it is possible to take k as large as desired, it would be sufficient to consider $k+1$ for that $B > 2^{e_1} m$ (since $2^{e_1} m$ does not depend on k).

Therefore, there exists no $m \in 2\mathbb{N}+1, t \geq 4$ and $(\alpha_0, \ldots, \alpha_{t-1}) \in (\mathbb{N}^*)^t$, such that for all $k \geq 2$, we have $B = 2^{e_1} m$, and thereby the equation (2.1) has no solution.

This implies that for all $p \in 2\mathbb{N}+1$, no R-Cz sequence can exhibit a cycle of length $t \geq 4$, and it also confirms that any R-Cz sequence does not have a cycle of length 2 or 3, since nothing prevents t from taking the value 2 or 3 in equation (2.1). □

3 R-Cz sequences converging to 1

In this section, we will study the convergent R-Cz sequences (i.e., those that eventually reach 1) to determine their relationship to the set $2\mathbb{N}+1$. This will lead us to partition this set into the subsets $6\mathbb{N}+1$, $6\mathbb{N}+3$ and $6\mathbb{N}+5$.

3.1 Penultimate terms

Since section 2 has established that the only cycle of length 1 an R-Cz sequence can have, is the cycle with value (1), it follows that if an R-Cz sequence converges, it must converge to 1. This implies that its penultimate term (the term preceding 1) is of the form $\frac{2^{2n}-1}{3}$, where $n \in \mathbb{N}^* \setminus \{1\}$.

Theorem 3.1. *If the R-Cz sequence $\{v_l^p\}_{l \in \mathcal{N}}$ converges to 1, then there exists $k \in \mathbb{N}$ such that $v_{k-1}^p = \frac{2^{2n}-1}{3}$, where $n \in \mathbb{N}^* \setminus \{1\}$, and such that $v_k^p = 1$.*

Proof. Let us suppose that the sequence $\{v_l^p\}$ converges to 1, and let k be the first index for which $v_k^p = 1$.

Given that $v_{k-1}^p \in 2\mathbb{N}+1$, we have:

$$v_k^p = \frac{3v_{k-1}^p + 1}{2^\alpha} = 1 \iff v_{k-1}^p = \frac{2^\alpha - 1}{3} \iff 2^\alpha - 1 \equiv 0 \pmod 3$$

For $\alpha = 1 \implies 2 - 1 \equiv 1 \pmod 3$ and for $\alpha = 2 \implies 4 - 1 \equiv 0 \pmod 3$. We prove by induction, assuming that for $\alpha = 2\alpha_0$ we have $2^{2\alpha_0} - 1 \equiv 0 \pmod 3$, where $\alpha_0 \in \mathbb{N}^*$, that:

$$2^{2(\alpha_0+1)} - 1 \equiv (3+1) \times 2^{2\alpha_0} - 1 \equiv 3 \times 2^{2\alpha_0} + (2^{2\alpha_0} - 1) \equiv 0 \pmod 3$$

And assuming that for $\alpha = 2\alpha_0 + 1$ we have $2^{2\alpha_0+1} - 1 \equiv 1 \pmod 3$, that:

$$2^{2(\alpha_0+1)+1} - 1 \equiv (3+1) \times 2^{2\alpha_0+1} - 1 \equiv 3 \times 2^{2\alpha_0+1} + (2^{2\alpha_0+1} - 1) \equiv 1 \pmod 3$$

Therefore, the penultimate term of the convergent R-Cz sequence $\{v_l^p\}$ is of the form $\frac{2^{2n}-1}{3}$, where $n \in \mathbb{N}^* \setminus \{1\}$. This sequence being arbitrary, it applies to all convergent R-Cz sequences. $\square$

This implies there are infinitely many distinct penultimate terms, each belonging to a distinct convergent R-Cz sequence. In all that follows, we will refer to $\boldsymbol{v_{b_p}^n}$ as the penultimate term of the convergent R-Cz sequence $\boldsymbol{\{v_l^p\}}$, where b_p is its index in the sequence and n is the variable n in $\frac{2^{2n}-1}{3}$.

3.2 Preceding terms

By going through the convergent R-Cz sequences, from their penultimate term $v_{b_p}^n$ to their second term v_0^p, we will study the preceding terms; first, those just preceding the penultimate terms, and then the other ones.

Let us recall that if a term, whether it is the penultimate term or any other term of a convergent R-Cz sequence, is immediately preceded by infinitely many terms, only one of them belongs to the same sequence.

3.2.1 Preceding terms of penultimate terms

The preceding terms of a penultimate term are those immediately before it, found in the convergent R-Cz sequences leading to this penultimate term. Each of these preceding terms belongs to a distinct convergent R-Cz sequence.

Theorem 3.2. *Let $v_{b_p}^n = \frac{2^{2n}-1}{3}$ be the penultimate term of the convergent R-Cz sequence $\{v_l^p\}_{l\in\mathcal{N}}$, and let $n = 3k + a$, where $k \in \mathbb{N}$, $a \in \{0, 1, 2\}$ and $n \geq 2$, therefore:*

- *if $a = 0$, $v_{b_p}^n \in 6\mathbb{N} + 3$ (class B) and has no preceding terms;*
- *if $a = 1$, $v_{b_p}^n \in 6\mathbb{N} + 1$ (class A) and has an infinite number of distinct preceding terms of the form $\frac{2^\alpha(2^{2(3k+1)}-1)-3}{9}$, where $\alpha \in 2\mathbb{N}^*$ and $k \in \mathbb{N}^*$;*
- *if $a = 2$, $v_{b_p}^n \in 6\mathbb{N} + 5$ (class C) and has an infinite number of distinct preceding terms of the form $\frac{2^\alpha(2^{2(3k+2)}-1)-3}{9}$, where $\alpha \in 2\mathbb{N} + 1$ and $k \in \mathbb{N}$.*

Proof. For the penultimate term $v_{b_p}^n$ we have:

$$v_{b_p}^n = \frac{3v_{b_p-1}^n + 1}{2^\alpha} \iff v_{b_p-1}^n = \frac{2^\alpha(2^{2n}-1)-3}{9} \iff 2^\alpha(2^{2n}-1) - 3 \equiv 0 \pmod 9$$

In other words, the term $v_{b_p-1}^n$ of the R-Cz sequence $\{v_l^p\}$ precedes the penultimate term if $2^\alpha(2^{2n} - 1) - 3$ is odd and divisible by 9. As the reader will note, to determine the preceding terms of a penultimate term, we will proceed by induction.

First case $a = 0$
For $(k, \alpha) = (1, 1)$, $2^1\,(2^{2\times3} - 1) - 3 \equiv 6 \pmod 9$.
And suppose that for $(k, \alpha) = (1, \alpha_0)$, where $\alpha_0 \in \mathbb{N}^* \setminus \{1\}$:

$$2^{\alpha_0}\left(2^{2\times3} - 1\right) - 3 \equiv 6 \pmod 9$$

We have for $(k, \alpha) = (1, \alpha_0 + 1)$:

$$2^{\alpha_0+1}\left(2^{2\times3} - 1\right) - 3 \equiv \left(2^{\alpha_0}\left(2^{2\times3} - 1\right) - 3\right) + 2\left(2^{2\times3} - 1\right) \equiv 6 \pmod 9$$

Then, suppose that for $(k, \alpha) = (k_0, \alpha)$, where $k_0 \in \mathbb{N}^* \setminus \{1\}$ and $\alpha \in \mathbb{N}^*$:

$$2^\alpha\left(2^{2\times3k_0} - 1\right) - 3 \equiv 6 \pmod 9$$

We have for $(k, \alpha) = (k_0 + 1, \alpha)$:

$$2^\alpha\left(2^{2\times3(k_0+1)} - 1\right) - 3 \equiv \left(2^\alpha\left(2^{2\times3k_0} - 1\right) - 3\right) + 63 \cdot 2^\alpha \cdot 2^{2\times3k_0} \equiv 6 \pmod 9.$$

Therefore, if $n \in 3\mathbb{N}^*$, the penultimate term $v_{b_p}^n$ has no preceding term. We demonstrate in Appendix 6.2 that in this case $v_{b_p}^n \in 6\mathbb{N} + 3$.

Second case $a = 1$
For $(k, \alpha) = (1, 1)$, $2^1\left(2^{2\times(3+1)} - 1\right) - 3 \equiv 3 \pmod 9$.
And for $(k, \alpha) = (1, 2)$,$2^2\left(2^{2\times(3+1)} - 1\right) - 3 \equiv 0 \pmod 9$.

Suppose that for $(k,\alpha) = (1, 2\alpha_0 + 1)$, where $\alpha_0 \in \mathbb{N}^* \setminus \{1\}$:

$$2^{2\alpha_0+1}\left(2^{2\times(3+1)} - 1\right) - 3 \equiv 3 \pmod 9$$

We have for $(k,\alpha) = (1, 2(\alpha_0 + 1) + 1)$:

$$2^{2(\alpha_0+1)+1}\left(2^{2\times(3+1)} - 1\right) - 3 \equiv \left(2^{2\alpha_0+1}\left(2^{2\times(3+1)} - 1\right) - 3\right) + 765 \cdot 2^{2\alpha_0+1} \equiv 3 \pmod 9.$$

And suppose that for $(k,\alpha) = (1, 2\alpha_0)$:

$$2^{2\alpha_0}\left(2^{2\times(3+1)} - 1\right) - 3 \equiv 0 \pmod 9$$

We have for $(k,\alpha) = (1, 2(\alpha_0 + 1))$:

$$2^{2(\alpha_0+1)}\left(2^{2\times(3+1)} - 1\right) - 3 \equiv \left(2^{2\alpha_0}\left(2^{2\times(3+1)} - 1\right) - 3\right) + 765 \cdot 2^{2\alpha_0} \equiv 0 \pmod 9$$

Then, suppose that for $(k,\alpha) = (k_0, \alpha)$, where $k_0 \in \mathbb{N}^* \setminus \{1\}$ and $\alpha \in 2\mathbb{N} + 1$:

$$2^{\alpha}\left(2^{2\times(3k_0+1)} - 1\right) - 3 \equiv 3 \pmod 9$$

We have for $(k,\alpha) = (k_0 + 1, \alpha)$:

$$2^{\alpha}\left(2^{2\times(3(k_0+1)+1)} - 1\right) - 3 \equiv \left(2^{\alpha}\left(2^{2\times(3k_0+1)} - 1\right) - 3\right) + 63 \cdot 2^{\alpha+2(3k_0+1)} \equiv 3 \pmod 9$$

Finally, suppose that for $(k,\alpha) = (k_0, \alpha)$, where $k_0 \in \mathbb{N}^* \setminus \{1\}$ and $\alpha \in 2\mathbb{N}^*$:

$$2^{\alpha}\left(2^{2\times(3k_0+1)} - 1\right) - 3 \equiv 0 \pmod 9$$

We have for $(k,\alpha) = (k_0 + 1, \alpha)$:

$$2^{\alpha}\left(2^{2\times(3(k_0+1)+1)} - 1\right) - 3 \equiv \left(2^{\alpha}\left(2^{2\times(3k_0+1)} - 1\right) - 3\right) + 63 \cdot 2^{\alpha+2(3k_0+1)} \equiv 0 \pmod 9$$

Therefore, if $n \in 3\mathbb{N}^* + 1$, the penultimate term $v^n_{b_p}$ is preceded an infinite number of distinct terms of the form $\frac{2^{\alpha}\left(2^{2n}-1\right)-3}{9}$, where $\alpha \in 2\mathbb{N}^*$. We demonstrate in Appendix 6.2 that in this case $v^n_{b_p} \in 6\mathbb{N} + 1$.

Third case $a = 2$
For $(k,\alpha) = (0,1), 2^1(2^2(0 + 2) - 1) - 3 \equiv 0 \pmod 9$.
And for $(k,\alpha) = (0,2), 2^2(2^2(0 + 2) - 1) - 3 \equiv 3 \pmod 9$.

Suppose that for $(k,\alpha) = (0, 2\alpha_0 + 1)$, where $\alpha_0 \in \mathbb{N}^* \setminus \{1\}$:

$$2^{2\alpha_0+1}(2^2(0+2)-1)-3 \equiv 0 \pmod 9$$

We have for $(k,\alpha) = (1, 2(\alpha_0+1)+1)$:

$$2^{2(\alpha_0+1)+1}(2^2(0+2)-1)-3 \equiv \big(2^{2\alpha_0+1}(2^2(0+2)-1)-3\big) + 45 \cdot 2^{2\alpha_0+1} \equiv 0 \pmod 9$$

And suppose that for $(k,\alpha) = (0, 2\alpha_0)$:

$$2^{2\alpha_0}(2^2(0+2)-1)-3 \equiv 3 \pmod 9$$

We have for $(k,\alpha) = (0, 2(\alpha_0+1))$:

$$2^{2(\alpha_0+1)}(2^2(0+2)-1)-3 \equiv \big(2^{2\alpha_0}(2^2(0+2)-1)-3\big) + 45 \cdot 2^{2\alpha_0} \equiv 3 \pmod 9$$

Then, suppose that for $(k,\alpha) = (k_0, \alpha)$, where $k_0 \in \mathbb{N}^*$ and $\alpha \in 2\mathbb{N}+1$:

$$2^{\alpha}(2^2(3k_0+2)-1)-3 \equiv 0 \pmod 9$$

We have for $(k,\alpha) = (k_0+1, \alpha)$:

$$2^{\alpha}(2^2(3(k_0+1)+2)-1)-3 \equiv \big(2^{\alpha}(2^2(3k_0+2)-1)-3\big)+63\cdot 2^{\alpha}\cdot 2^2(3k_0+2) \equiv 0 \pmod 9$$

Finally, suppose that for $(k,\alpha) = (k_0, \alpha)$, where $k_0 \in \mathbb{N}^*$ and $\alpha \in 2\mathbb{N}^*$:

$$2^{\alpha}(2^2(3k_0+2)-1)-3 \equiv 3 \pmod 9$$

We have for $(k,\alpha) = (k_0+1, \alpha)$:

$$2^{\alpha}(2^2(3(k_0+1)+2)-1)-3 \equiv \big(2^{\alpha}(2^2(3k_0+2)-1)-3\big)+63\cdot 2^{\alpha}\cdot 2^2(3k_0+2) \equiv 3 \pmod 9$$

Therefore, if $n \in 3\mathbb{N}+2$, the penultimate term $v^n_{b_p}$ is preceded by an infinite number of distinct terms of the form $\frac{2^{\alpha}(2^{2n}-1)-3}{9}$, where $\alpha \in 2\mathbb{N}+1$. We demonstrate in Appendix 6.2 that in this case $v^n_{b_p} \in 6\mathbb{N}+5$. □

To summarize, the penultimate terms fall into three classes, depending on the value of n modulo 3. Only those in $6\mathbb{N}+1$ or $6\mathbb{N}+5$ have preceding terms, as shown in the figure below.

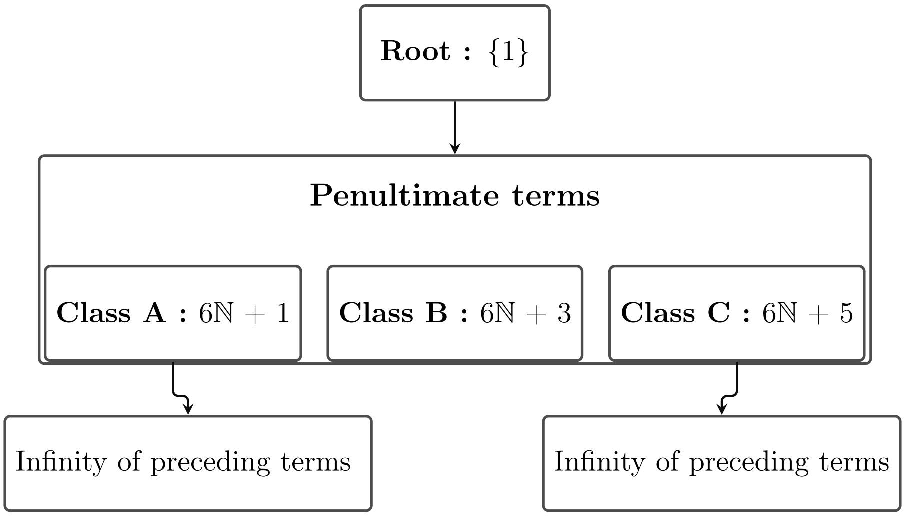

Figure 1. The 3 classes of penultimate terms

3.2.2 Other preceding terms

The other preceding terms are those that immediately precede a term of an R-Cz sequence, whether convergent or not. As in the case of the penultimate terms, each of these preceding terms belongs to a distinct R-Cz sequence.

Theorem 3.3. *Let $\{v_l^p\}_{l\in\mathcal{N}}$ be an R-Cz sequence, convergent or not, and let $v_i^p \in \{v_l^p\}$, where $i \in \mathbb{N}$ and $v_i^p \neq 1$, denoted as the parent term, then:*

- *if $v_i^p \in 6\mathbb{N}^* + 1$, it is preceded by an infinite number of distinct terms alternately in $6\mathbb{N}+1$, $6\mathbb{N}+3$ and $6\mathbb{N}+5$. These terms are of the form $\frac{2^{\alpha_i}.v_i^p-1}{3}$, where $\alpha_i \in 2\mathbb{N}^*$;*
- *if $v_i^p \in 6\mathbb{N}+3$, it has no preceding terms;*
- *if $v_i^p \in 6\mathbb{N}+5$, it is preceded by an infinite number of distinct terms alternately in $6\mathbb{N}+1$, $6\mathbb{N}+3$ and $6\mathbb{N}+5$. These terms are of the form $\frac{2^{\alpha_i}.v_i^p-1}{3}$, where $\alpha_i \in 2\mathbb{N}+1$.*

Proof. Let $v_i^p \in \{v_l^p\}$, where $\{v_l^p\}$ is an R-Cz sequence, $i \in \mathbb{N}$ and $v_i^p \neq 1$, we have:

$$v_i^p = \frac{3v_{i-1}^p + 1}{2^{\alpha_i}} \iff v_{i-1}^p = \frac{2^{\alpha_i}\ v_i^p - 1}{3} \iff 2^{\alpha_i}\ v_i^p - 1 \equiv 0 \pmod 3.$$

This implies that the previous terms of v_i^p are odd and divisible by 3, and one of them is v_{i-1}^p. As before, we will proceed by induction.

First case $v_i^p \in 6\mathbb{N}^* + 1$
If $v_i^p \in 6\mathbb{N}^* + 1$, then $v_i^p \geq 7$ and there exists $k \in \mathbb{N}^*$ such that $v_i^p = 6k+1$. In the sequel, we make use of the congruence $6k+1 \equiv 1 \pmod 3$.

For $(k, \alpha_i) = (k, 1)$, $2^1(6\times k+1)-1 \equiv 1 \pmod 3$ and for $(k, \alpha_i) = (k, 2)$, $2^2(6\times k+1)-1 \equiv 0 \pmod 3$. Suppose that for $(k, \alpha_i) = (k, 2\alpha_0+1)$, where $\alpha_0 \in \mathbb{N}^* \setminus \{1\}$:

$$2^{2\alpha_0+1}(6\times k+1) - 1 \equiv 1 \pmod 3$$

We have for $(k, \alpha_i) = (k, 2(\alpha_0 + 1) + 1)$:

$$2^{2(\alpha_0+1)+1}(6 \times k + 1) - 1 \equiv 2^{2\alpha_0+1}(6 \times k + 1) - 1 + 21 \cdot 2^{2\alpha_0+1} \equiv 1 \pmod 3.$$

And suppose that for $(k, \alpha_i) = (k, 2\alpha_0)$, where $\alpha_0 \in \mathbb{N}^* \setminus \{1\}$:

$$2^{2\alpha_0}(6 \times k + 1) - 1 \equiv 0 \pmod 3$$

We have for $(k, \alpha_i) = (k, 2(\alpha_0 + 1))$:

$$2^{2(\alpha_0+1)}(6 \times k + 1) - 1 \equiv \left(2^{2\alpha_0}(6 \times k + 1) - 1\right) + 21 \cdot 2^{2\alpha_0} \equiv 0 \pmod 3$$

Therefore, if $v_i^p \in 6\mathbb{N} + 1$, it is preceded by an infinite number of distinct terms of the form $\frac{2^{\alpha_i}(6k+1)-1}{3}$, where $\alpha_i \in 2\mathbb{N}^*$. We demonstrate in Appendix 6.3 that these preceding terms are alternately in $6\mathbb{N} + 1$, $6\mathbb{N} + 3$ and $6\mathbb{N} + 5$.

Second case $v_i^p \in 6\mathbb{N} + 3$
If $v_i^p \in 6\mathbb{N}+3$, then $v_i^p \geq 3$ and there exists $k \in \mathbb{N}^*$ such that $v_i^p = 6k+3$. Since $6k+3 \equiv 0 \pmod 3$, we have:

$$2^{\alpha_i}(6k + 3) - 1 \equiv 2 \pmod 3$$

Therefore, if $v_i^p \in 6\mathbb{N} + 3$, it has no preceding terms.

Third case $v_i^p \in 6\mathbb{N} + 5$
If $v_i^p \in 6\mathbb{N} + 5$, then $v_i^p \geq 5$ and there exists $k \in \mathbb{N}$ such that $v_i^p = 6k + 5$. In the sequel, we make use of the congruence $6k + 5 \equiv -1 \pmod 3$.

For $(k, \alpha_i) = (k, 1)$, $2^1(6\times k+5)-1 \equiv 0 \pmod 3$, and $(k, \alpha_i) = (k, 2)$, $2^2(6\times k+5)-1 \equiv 1 \pmod 3$. Suppose that for $(k, \alpha_i) = (k, 2\alpha_0 + 1)$, where $\alpha_0 \in \mathbb{N}^* \setminus \{1\}$:

$$2^{2\alpha_0+1}(6 \times k + 5) - 1 \equiv 0 \pmod 3$$

We have for $(k, \alpha_i) = (k, 2(\alpha_0 + 1) + 1)$:

$$2^{2(\alpha_0+1)+1}(6 \times k + 5) - 1 \equiv (2^{2\alpha_0+1}(6 \times k + 5) - 1) + 15 \cdot 2^{2\alpha_0+1} \equiv 0 \pmod 3$$

And suppose that for $(k, \alpha_i) = (k, 2\alpha_0)$, where $\alpha_0 \in \mathbb{N}^* \setminus \{1\}$:

$$2^{2\alpha_0}(6 \times k + 5) - 1 \equiv 1 \pmod 3$$

We have for $(k, \alpha_i) = (k, 2(\alpha_0 + 1))$:

$$2^{2(\alpha_0+1)}(6 \times k + 5) - 1 \equiv (2^{2\alpha_0}(6 \times k + 5) - 1) + 15 \cdot 2^{2\alpha_0} \equiv 1 \pmod 3$$

Therefore, if $v_i^p \in 6\mathbb{N}+5$, it is preceded by an infinite number of distinct terms of the form $\frac{2^{\alpha_i}(6k+5)-1}{3}$, where $\alpha_i \in 2\mathbb{N} + 1$. We demonstrate in Appendix 6.3 that these preceding terms are alternately in $6\mathbb{N} + 1$, $6\mathbb{N} + 3$ and $6\mathbb{N} + 5$. $\square$

Clearly, Theorem 3.3 also applies to the penultimate terms, and due to its recursive nature, it reveals an infinite tree structure, which is the subject of the following section.

3.3 Tree structure of the convergent R-Cz sequences

Thanks to what has been built and demonstrated so far, the convergent R-Cz sequences can be represented in the form of a tree structure as follows:

- **at level 0:** the root set containing the element 1;
- **at level 1:** three infinite subsets of $2\mathbb{N} + 1$. The first contains the penultimate terms in $6\mathbb{N} + 1$ (class A), the second those in $6\mathbb{N} + 3$ (class B), and the third those in $6\mathbb{N} + 5$ (class C). Their union contains the penultimate terms of all convergent R-Cz sequences;
- **at level 2:** each term of the class A or C is immediately preceded by an infinite number of distinct terms alternately in $6\mathbb{N} + 1$, $6\mathbb{N} + 3$ and $6\mathbb{N} + 5$, which form together the level 2. The terms in $6\mathbb{N} + 3$ have no preceding terms;
- **at level 3:** each term of the previous level is immediately preceded by an infinite number of distinct terms alternately in $6\mathbb{N} + 1$, $6\mathbb{N} + 3$ and $6\mathbb{N} + 5$, which form together the level 3. As in levels 1 and 2, the terms in $6\mathbb{N} + 3$ have no preceding terms;
- and so on *ad infinitum*.

From this description, a convergent R-Cz sequence is a path that starts from a node, or a leaf when its first term is in $6\mathbb{N} + 3$, and converges towards the root. The figure below shows this tree structure.

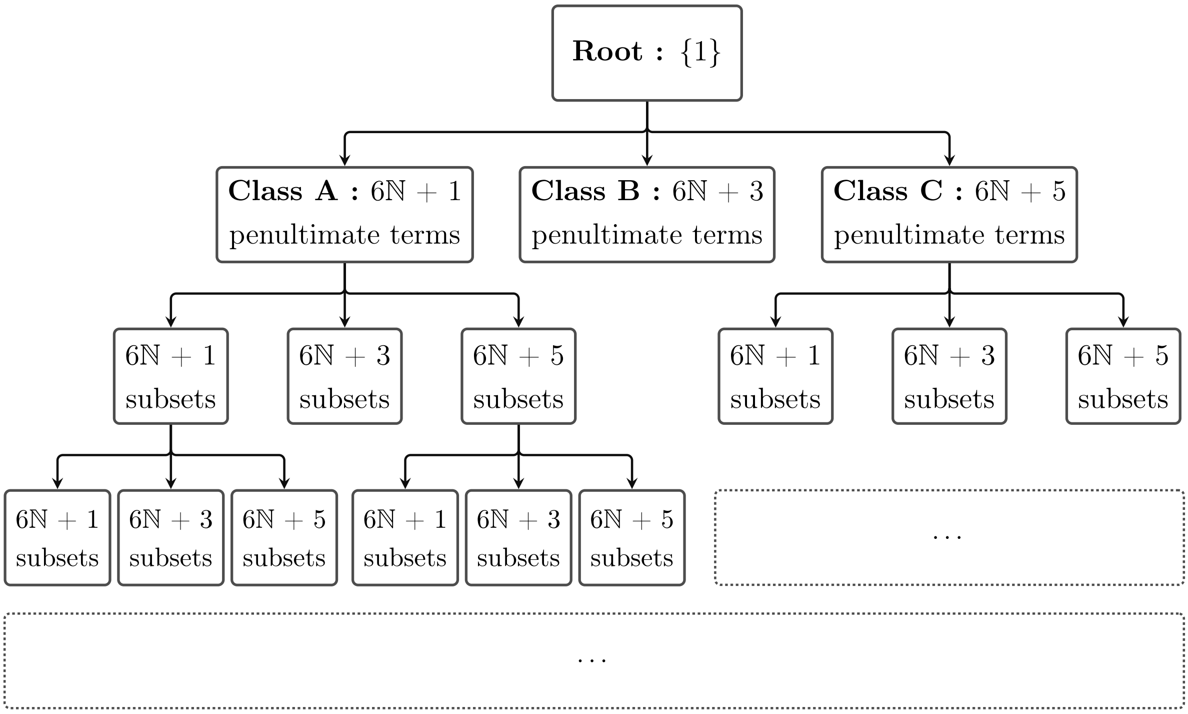

Figure 2. Tree structure of the convergent R-Cz sequences

Definition 3.1. *Let C_z be the union of the subsets of $6\mathbb{N}+1$, $6\mathbb{N}+3$ and $6\mathbb{N}+5$ contained in the first z levels of the tree structure (excluding level 0), defined as follows:*

$$C_z = \bigcup_{i=1}^{z} \bigcup_{j \in I_i} C_{i,j}$$

Where $z \in \mathbb{N}^*$, each $C_{i,j}$ is a subset of $6\mathbb{N}+1$, $6\mathbb{N}+3$ or $6\mathbb{N}+5$ that contains terms of the sequences, and I_i is the set indexing the subsets of level i of the tree structure. And let C be the complete union:

$$C = \bigcup_{i \geq 0} \bigcup_{j \in I_i} C_{i,j}$$

Theorem 3.4. *Let $(i_0, j_0) \in \mathbb{N} \times I_{i_0}$ and $(i_1, j_1) \in \mathbb{N} \times I_{i_1}$ such that $(i_0, j_0) \neq (i_1, j_1)$, then $C_{(i_0,j_0)} \cap C_{(i_1,j_1)} = \emptyset$.*

Proof. Let $(i_0, j_0) \in \mathbb{N} \times I_{i_0}$ and $(i_1, j_1) \in \mathbb{N} \times I_{i_1}$ be two distinct pairs. First of all, if C_{i_0,j_0} and C_{i_1,j_1} are not both subsets of $6\mathbb{N}+1$, $6\mathbb{N}+3$ or $6\mathbb{N}+5$, their intersection is clearly empty. Then, if there is a descending or ascending path in the tree structure from C_{i_0,j_0} to C_{i_1,j_1}, which corresponds to a partial R-Cz sequence, since Theorems 2.1 to 2.3 state that an R-Cz sequence has no cycle (except the cycle of length 1 and value (1)), the same term cannot appear in both C_{i_0,j_0} and C_{i_1,j_1}, hence, $C_{i_0,j_0} \cap C_{i_1,j_1} = \emptyset$.

Otherwise, if there is no descending or ascending path from C_{i_0,j_0} to C_{i_1,j_1}, suppose there exists $v_l^p \in C_{i_0,j_0} \cap C_{i_1,j_1}$, where v_l^p is a term of a convergent R-Cz sequence. This would imply that either a parent term has v_l^p as its preceding term twice, which cannot occur since the preceding terms of a parent term are distinct (see Theorems 3.2 and 3.3), or that two distinct parent terms have v_l^p as their preceding term, which also cannot occur since a term of an R-Cz sequence (in this case v_l^p) can only be followed by a single term. Therefore, in all cases, $C_{i_0,j_0} \cap C_{i_1,j_1} = \emptyset$. □

Lemma 3.1. *There are an infinite number of convergent R-Cz sequences.*

Proof. By Theorem 3.2, there are an infinite number of penultimate terms forming the level 1 of the tree structure, and each of them is preceded by an infinite number of distinct terms forming level 2.

In turn, by Theorem 3.3, each term of level 2 is preceded by an infinite number of distinct terms forming level 3, and so on.

Therefore, traversing the tree from the nodes, or the leaves (i.e., the terms in $6\mathbb{N}+3$ because they do not have preceding terms), to the root, there are infinitely many paths corresponding to as many convergent R-Cz sequences. □

Remark 3.1. *Since terms in* $6\mathbb{N}+3$ *have no preceding terms in any R-Cz sequence, only the first term of an R-Cz sequence, convergent or not, can be in* $6\mathbb{N}+3$.

3.4 Sequences of preceding terms

A sequence of preceding terms contains all preceding terms generated by a parent term[(i)], ordered in ascending order. In other words, it contains only the infinitely many children of that parent term, and therefore all terms of the sequence belong to the same level of the tree structure. In all that follows, this type of sequence will be referred to as a *Pt-sequence.*

3.4.1 Main objects

Definition 3.2. *Let* $\{b_i\}_{i\geq 0}$ *be the sequence of penultimate terms ordered in ascending order, such that for all* $i \in \mathbb{N}, b_i = \frac{2^{2(i+2)}-1}{3}$.

Thus $b_0 = 5, b_1 = 21, b_2 = 85, b_3 = 341, b_4 = 1365$, etc.

Definition 3.3. *Let* $\{s_{j,i}\}_{i\in\mathbb{N}}$ *be a Pt-sequence, according to Theorem 3.3, it is defined as follows:*

$$s_{j,i} = \begin{cases} \frac{2}{3}q_j \cdot 4^i - \frac{1}{3}, & \text{if } q_j \in 6\mathbb{N}+5 \\ \frac{4}{3}q_j \cdot 4^i - \frac{1}{3}, & \text{if } q_j \in 6\mathbb{N}+1 \end{cases}$$

(i) Except for 1, which is the parent term of the penultimate terms.

Where $j \in \mathbb{N}$ and q_j is the parent term of the Pt-sequence.

Let $\{\{s_{j,i}\}\}_{(j,i)\in I_{Pt}\times\mathbb{N}}$ be the family of the Pt-sequences sequences corresponding to all convergent R-Cz sequences, ordered by their first term in ascending order, where I_{Pt} is the set used for indexing the Pt-sequences with a single index.

Let $S = \bigcup_{j\in I_{Pt}}\{s_{j,i}\}$ be the union of terms of these Pt-sequences, then $C = S \cup \{b_i\} \cup \{1\}$.

In view of the above, the set C contains the terms of all convergent R-Cz sequences, consequently, according to Definition 1.2 and Theorem 1.1, C is in bijection with R^{c}_{Cz}.

Lemma 3.2. *For all $(j_0, j_1) \in I_{Pt} \times I_{Pt}$, with $j_0 \neq j_1$, we have $\{s_{j_0,i}\} \cap \{s_{j_1,i}\} = \emptyset$. In other words, no Pt-sequence shares terms with another.*

Proof. This lemma is a straightforward consequence of Theorems 3.3 and 3.4. Indeed, let $(j_0, j_1) \in I_{Pt} \times I_{Pt}$, with $j_0 \neq j_1$, and let q_{j_0} (respectively, q_{j_1}) be the parent term of the Pt-sequence $\{s_{j_0,i}\}$ (respectively, $\{s_{j_1,i}\}$).

It follows from Definition 3.3 that $\{s_{j_0,i}\} \cap \{s_{j_1,i}\} \neq \emptyset$, if there exists $\alpha_{j_0} \in 2\mathbb{N}^*$ or $2\mathbb{N}+1$, depending on whether $q_{j_0} \in 6\mathbb{N}^* + 1$ or $6\mathbb{N} + 5$, and if there exists $\alpha_{j_1} \in 2\mathbb{N}^*$ or $2\mathbb{N} + 1$, depending on whether $q_{j_1} \in 6\mathbb{N}^* + 1$ or $6\mathbb{N} + 5$, such that:

$$\frac{2^{\alpha_{j_0}} q_{j_0} - 1}{3} = \frac{2^{\alpha_{j_1}} q_{j_1} - 1}{3} \iff 2^{\alpha_{j_0}} q_{j_0} = 2^{\alpha_{j_1}} q_{j_1}$$

Without loss of generality, we can assume $\alpha_{j_0} > \alpha_{j_1}$, then $2^{\alpha_{j_0}-\alpha_{j_1}} q_{j_0} = q_{j_1}$. However, this leads to a contradiction since the left-hand side is even, while q_{j_1} is odd. The same contradiction arises if we assume $\alpha_{j_0} < \alpha_{j_1}$.

Finally, if $\alpha_{j_0} = \alpha_{j_1}$, then $\{s_{j_0,i}\} = \{s_{j_1,i}\}$ and therefore $j_0 = j_1$, contradicting our hypothesis. Hence, if $j_0 \neq j_1$, then $\{s_{j_0,i}\} \cap \{s_{j_1,i}\} = \emptyset$. □

Lemma 3.3. *The sets I_{Pt} and C are countable.*

Proof. According to Lemma 3.2, for all $(j_0, j_1) \in I_{Pt} \times I_{Pt}$, with $j_0 \neq j_1$, $\{s_{j_0,i}\} \cap \{s_{j_1,i}\} = \emptyset$, therefore, $\bigcup_{j\in I_{Pt}}\{s_{j,i}\}$ is a partition of S. And considering $S \subset 2\mathbb{N}+1$, I_{Pt} is countable. Then, given that $C = S \cup \{b_i\} \cup \{1\}$, as a union of countable sets, C is countable. □

3.4.2 Shifts

Lemma 3.4. *Let q_j be a term in the tree structure at level $k \geq 1$ (the parent term), and let $\{s_{j,i}\}_{i\geq 0}$ be the Pt-sequence it generates, shifted either to the left or to the right. Then, for the terms of the Pt-sequence we have:*

$$s_{j,i+1} - s_{j,i} = \begin{cases} 2^{2i+1}.q_j & \text{if } q_j \in 6\mathbb{N}+5 \\ 2^{2(i+1)}.q_j & \text{if } q_j \in 6\mathbb{N}+1 \end{cases} \tag{3.1}$$

$$\lim_{i\to\infty}\frac{s_{j,i+1}}{s_{j,i}}=4 \text{ and } 4<\frac{s_{j,i+1}}{s_{j,i}}\leq\frac{13}{3} \tag{3.2}$$

For the shift factor:

$$\beta^{q_j}=\frac{s_{j,0}}{q_j}=\begin{cases}\frac{2-\frac{1}{q_j}}{3}<\frac{2}{3} & \text{if } q_j\in 6\mathbb{N}+5 \ (s_{j,0}<q_j: \text{left shift})\\ \frac{4-\frac{1}{q_j}}{3}<\frac{4}{3} & \text{if } q_j\in 6\mathbb{N}+1 \ (s_{j,0}>q_j: \text{right shift})\end{cases} \tag{3.3}$$

And for the first term of the Pt-sequence:

$$s_{j,0}=\beta^{q_j}.q_j \tag{3.4}$$

Proof. For (3.1), if $q_j\in 6\mathbb{N}+5$ (respectively, $q_j\in 6\mathbb{N}+1$), Theorem 3.3 states that the term number i of $\{s_{j,i}\}$ is equal to $s_{j,i}=\frac{2^{2i+1}\ q_j-1}{3}$ (respectively, $s_{j,i}=\frac{2^{2(i+1)}\ q_j-1}{3}$), therefore:

$$s_{j,i+1}-s_{j,i}=\begin{cases}\frac{2^{2(i+1)+1}\ q_j-1}{3}-\frac{2^{2i+1}\ q_j-1}{3}=2^{2i+1}\ q_j & \text{if } q_j\in 6\mathbb{N}+5\\ \frac{2^{2(i+2)}\ q_j-1}{3}-\frac{2^{2(i+1)}\ q_j-1}{3}=2^{2(i+1)}\ q_j & \text{if } q_j\in 6\mathbb{N}+1\end{cases}$$

Concerning (3.2), if $q_j\in 6\mathbb{N}+1$:

$$\lim_{i\to+\infty}\frac{s_{j,i+1}}{s_{j,i}}=\frac{2^{2(i+2)}\ q_j-1}{2^{2(i+1)}\ q_j-1}=4 \tag{3.5}$$

For $(q_j,i)=(7,0)$, $4<\frac{s_{j,i+1}}{s_{j,i}}=\frac{16\times7-1}{4\times7-1}=\frac{111}{27}=\frac{13}{3}$, then, because of (3.5), for all $(q_j,i)\in(6\mathbb{N}+1)\times\mathbb{N}$, it follows that $4<\frac{s_{j,i+1}}{s_{j,i}}=\frac{2^{2(i+1)+1}\times7-1}{2^{2(i+1)}\times7-1}\leq\frac{13}{3}$. We get the same result for $q_j\in 6\mathbb{N}+5$.

Concerning (3.3), Theorem 3.3 states that $\frac{s_{j,0}}{q_j}=\frac{2-\frac{1}{q_j}}{3}<\frac{2}{3}$, when $q_j\in 6\mathbb{N}+5$, or that $\frac{s_{j,0}}{q_j}=\frac{4-\frac{1}{q_j}}{3}<\frac{4}{3}$, when $q_j\in 6\mathbb{N}+1$. And finally, (3.4) is clearly a result of (3.3). □

Definition 3.4. *Let $\beta^{(\mathrm{a})k}$ be the average shift factor of the first $k+1$ Pt-sequences, by which a Pt-sequence $\{s_{j,i}\}$, where $j\in I_{Pt}$ and $0\leq j\leq k$, is shifted from its parent term q_j, defined as follows:*

$$\beta^{(\mathrm{a})k}=\left(\prod_{j\in I_{Pt},\ 0\leq j\leq k}\beta^{q_j}\right)^{\frac{1}{k+1}}$$

Where $\beta^{q_j}=\begin{cases}\frac{2-\frac{1}{q_j}}{3} & \text{if } q_j\in 6\mathbb{N}+5\\ \frac{4-\frac{1}{q_j}}{3} & \text{if } q_j\in 6\mathbb{N}+1\end{cases}$.

Lemma 3.5. *If for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$, then the number of right-shifted Pt-sequences, among the first $k+1$, exceeds γ times the number of left-shifted ones, where $\gamma = \frac{\ln(\frac{3}{2})}{\ln(\frac{4}{3})}$.*

Proof. Suppose that for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$. Definition 3.4 states that for k sufficiently large:

$$\left(\prod_{\substack{j \in I_{Pt} \\ 0 \leq j \leq k}} \beta^{q_j} \right)^{\frac{1}{k+1}} \approx \left(\left(\frac{2}{3}\right)^l \left(\frac{4}{3}\right)^r \right)^{\frac{1}{k+1}}$$

Where $(l, r) \in \mathbb{N}^* \times \mathbb{N}^*$ and $l + r = k + 1$, hence:

$$\left(\left(\frac{2}{3}\right)^l \left(\frac{4}{3}\right)^r \right)^{\frac{1}{k+1}} > 1 \iff \left(\frac{2}{3}\right)^l \left(\frac{4}{3}\right)^r > 1 \iff \left(\frac{2}{3}\right) \left(\frac{4}{3}\right)^{\frac{r}{l}} > 1$$

Finally, this yields $r > \frac{\ln(\frac{3}{2})}{\ln(\frac{4}{3})} l \approx 1.4094\, l$, and we conclude that the first $k+1$ Pt-sequences are shifted to the right strictly more than $\gamma = \frac{\ln(\frac{3}{2})}{\ln(\frac{4}{3})}$ times as much as to the left. $\square$

Lemma 3.6. *If for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$, then the sequence $\{q_j\}_{j \in \mathbb{N}}$, ordered in ascending order and where q_j is the parent term of the Pt-sequence $\{s_{j,i}\}$, grows on average exponentially.*

Proof. Suppose that for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$. Let q_j be the parent term of the Pt-sequence $\{s_{j,i}\}$.

Since all Pt-sequences belonging to the tree of convergent R-Cz sequences originate from penultimate terms, there exists a path in the tree structure from b_{i_j} to q_j, such that:

$$b_{i_j} \to q_{j_1} \to q_{j_2} \to \cdots \to q_j = q_{j_{N_j - 1}}$$

Where b_{i_j} is the ancestor penultimate term of q_j, and N_j is the number of levels between b_{i_j} and q_j (excluding the level of b_{i_j} and including the level of q_j), so that q_j is located at level $N_j + 1$ in the tree structure.

According to Lemma 3.4, we then have:

$$q_j = \begin{cases} \delta_j + 4^{a_j} \beta^{b_{i_j}} b_{i_j} & \text{if } N_j = 1 \\ \delta_j + 4^{a_j} \beta^{b_{i_j}} b_{i_j} \prod_{n=1}^{N_j - 1} \beta^{q_{j_n}} & \text{if } N_j > 1 \end{cases}$$

Where:

- $\beta^{b_{i_j}}$ is the shift factor associated with the penultimate term b_{i_j};
- $\beta^{q_{j_n}}$ is the shift factor associated with the parent term q_{j_n};
- $a_j = \sum_{n=1}^{N_j} \text{pos}(q_{j_n})$, where $\text{pos}(q_{j_n})$ is the index[(i)] of q_{j_n} in its Pt-sequence;
- $\delta_j \in \mathbb{Q}_+$ depends on a_j.

Taking into account the average shift factor $\beta^{(\text{a})k}$, we have:

$$q_j \approx \begin{cases} \delta_j + 4^{a_j} \left(\beta^{(\text{a})k}\right) b_{i_j} & \text{if } N_j = 1 \\ \delta_j + 4^{a_j} \left(\beta^{(\text{a})k}\right)^{N_j} b_{i_j} & \text{if } N_j > 1 \end{cases}$$

$\left(\beta^{(\text{a})k}\right)^{N_j}$ is the average shift factor relative to the penultimate term b_{i_j}, and $\delta_j + 4^{a_j}$ indicates the horizontal position of q_j at level N_{j+1} in the tree structure.

Given that both $\left(\beta^{(\text{a})k}\right)^{N_j}$ and 4^{a_j} grow exponentially with N_j and a_j, it follows that the sequence $\{q_j\}$ grows on average exponentially. □

3.4.3 Central result

Building on what has been defined and demonstrated so far, we now introduce a central result that establishes a relationship between the distribution of the Pt-sequences in $2\mathbb{N}+1$ and the average shift factor.

Let D be the complement set of C in $2\mathbb{N} + 1$.

Definition 3.5. *Let $R_i^D = [a_i, b_i]$ be an interval of consecutive odd numbers in D, which corresponds to an empty space in C such that $R_i^D \cap C = \emptyset$, and let $R^D = \{R_i^D\}_{i \in E_s}$ be their set, where E_s is the set indexing all empty spaces of C.*

Theorem 3.5. *If the set C is countable, then there exists $K \in \mathbb{N}$ such that for all $k \geq K$, $\beta^{(\text{a})k} \leq 1$.*

Proof. We will proceed by contradiction.

Hypothesis
Suppose that C is countable and assume by contradiction that for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\text{a})k} > 1$.

As this applies to all K, let $\{(k_i, K_i)\}_{i \in \mathbb{N}}$ be a family of distinct pairs such that $k_i \geq K_i$ and $\beta^{(\text{a})k_i} > 1$, and let $\beta^{(\text{m})I} = \min\left(\{\beta^{(\text{a})k_i}\}_{0 \leq i \leq I}\right)$ be the minimum average shift factor on the interval $[0, k_I]$, where $I \in \mathbb{N}$.

(i) Starting from 0.

Pt-sequences shift
For the sake of the proof, we suppose that for all $(j_0, j_1) \in I_{Pt} \times I_{Pt}$, with $j_0 \neq j_1$, we have $\{s_{j_0,i}\} \cap \{s_{j_1,i}\} = \emptyset$, otherwise it would contradict the fact that no Pt-sequence shares terms with another, as stated in Lemma 3.2.

Then, given that for all $I \in \mathbb{N}$, $\beta^{(\mathrm{m})I} > 1$, the Pt-sequences are shifted to the right strictly more than γ times as much as to the left, as established by Lemma 3.5.

Thus, when $\beta^{(m)I}$ is close to 1, approximately 58.5% of the Pt-sequences are shifted to the right by a factor $\approx \frac{4}{3}$, and this percentage increases as $\beta^{(\mathrm{m})I}$ increases.

Empty spaces creation
According to Theorem 3.3, the creation of the Pt-sequences follows a tree-like pattern, as does the creation of the empty spaces resulting from these sequences.

Let T_{es} denote the tree of the empty spaces, we then have:

- at level 1, there are infinitely many empty spaces located between the penultimate terms, their size is unbounded as the penultimate terms progress exponentially (see Definition 3.2);

- at level 2, penultimate terms in $6\mathbb{N} + 1$ or $6\mathbb{N} + 5$ generate an infinite number of Pt-sequences, whose terms divide each empty spaces of level 1, into which they are inserted, giving rise to an infinite number of new empty spaces;

- at level 3, in turn terms of level 2 in $6\mathbb{N}+1$ or $6\mathbb{N}+5$ generate an infinite number of Pt-sequences, whose terms subdivide each empty spaces of level 2, into which they are inserted, giving rise once again to an infinite number of new empty spaces;

- and so on.

Owing to the hypothesis, the generated Pt-sequences are not only shifted predominantly to the right but do so exponentially.

Indeed, since for all $I \in \mathbb{N}$, $\beta^{(\mathrm{m})I} > 1$, Lemma 3.6 states that at level z the Pt-sequences are on average shifted to the right by a factor of at least $(\beta^{(\mathrm{m})I})^{z-1}$. To put it another way, the Pt-sequences become increasingly distant from one another.

Therefore, the gaps between successive terms, together with $\beta^{(\mathrm{a})k} > 1$, ensure that the spacing within and between sequences is sufficient to contain the Pt-sequences generated successively.

It follows that the creation of infinitely many new empty spaces at successive levels of the tree structure must necessarily continue endlessly (see Appendix 6.5).

As shown in the figure below, the creation process of empty spaces can be represented graphically by an infinite tree growing diagonally to the right.

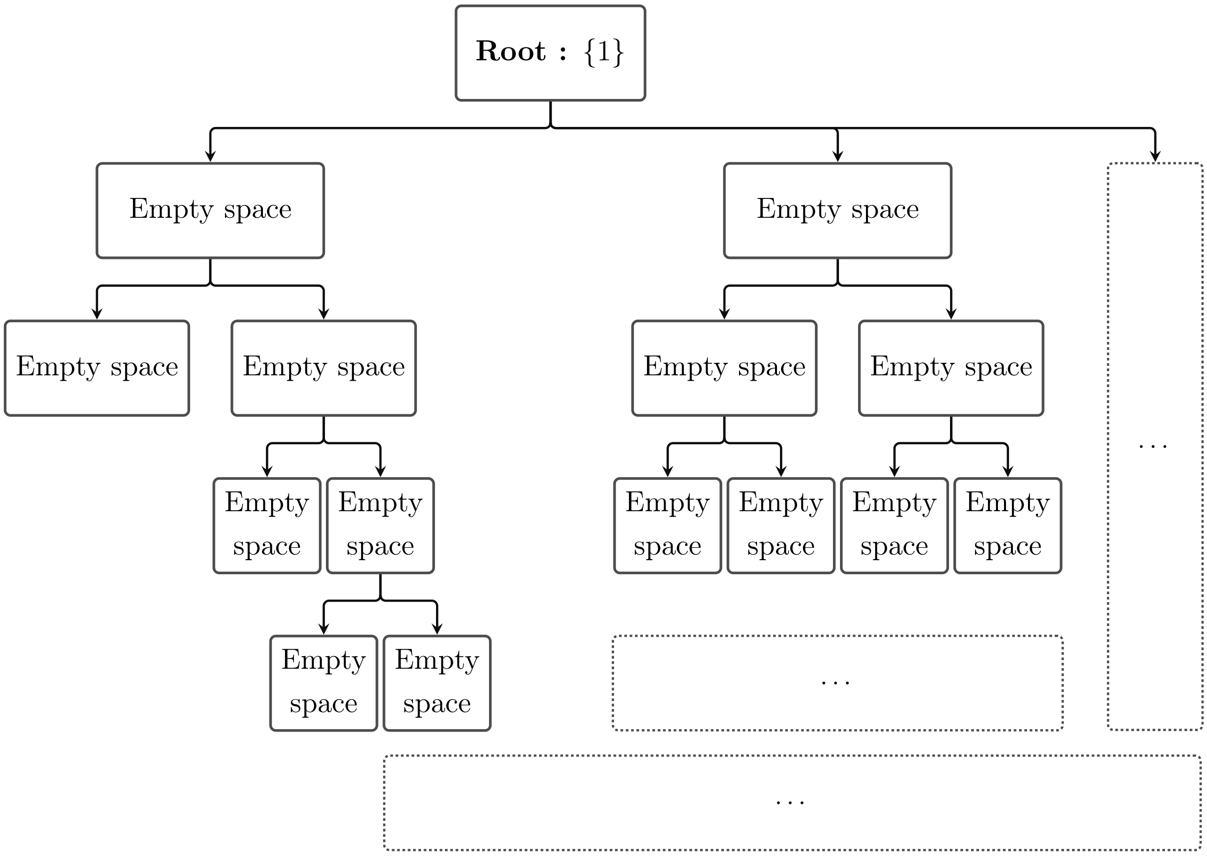

Figure 3. Tree structure of empty spaces

Uncountability

Insofar as it generates infinitely many levels, this creation process causes the tree to expand endlessly, producing an infinite number of unending paths, and thereby making the complete indexation of all empty spaces impossible.

Indeed, first it should be noted that although some branches terminate at certain levels, due to the predominantly rightward shift of the Pt-sequences, which causes some empty spaces to stop dividing, there remain infinitely many empty spaces at each level. Consequently, endless paths can still be constructed by passing through different empty spaces at the same level.

Then, let $p^i = (p^i_{n_0,a_0}, \ldots, p^i_{n_z,a_z}, \ldots) \in \mathbb{N}^{\mathbb{N}}$ be an endless path of T_{es}, where $i \in \mathbb{N}$ is the number of the path and $p^i_{n_z,a_z}$ the index of an empty space at level n_z.

We have $(n_0, a_0) = (1, 0)$ and for all $j \in \mathbb{N}, (n_{j+1}, a_{j+1}) = (n_j + 1, 0)$, except when a branch terminates at level n_j. In that case $(n_{j+1}, a_{j+1}) = (n_j, 1), p^i_{n_{j+1},a_{j+1}}$ is the index of the next empty space at level n_j that will be subdivided at the level $n_j + 1$, and we define $(n_{j+2}, a_{j+2}) = (n_j + 1, 0)$.

Finally, suppose there exists a bijective function $f : \mathbb{N} \to P$ defined as follows:

$$f(i) = (p^i_{n_0,a_0}, \ldots, p^i_{n_z,a_z}, \ldots)$$

Where $P = \{p^i\}_{i\in\mathbb{N}}$ is supposed to be the set of all endless paths of the tree structure, and supposed to be countable as it is in bijection with $\mathbb{N}$.

Since there are infinitely many empty spaces at every level of T_{es} (except at level 0), there exist indexes $p^x_{n_0,a_0} \neq p^i_{n_0,a_0}$, …, $p^x_{n_z,a_z} \neq p^i_{n_z,a_z}$, …, which form a new endless path $p^x = (p^x_{n_0,a_0}, \dots, p^x_{n_z,a_z}, \dots)$.

Although this path is valid, it does not belong to P, as it differs from every path $p^i \in P$ in at least one position, contradicting the assumption that P contains all endless paths of the tree structure, and thereby demonstrating that their set is uncountable.

Conclusion

As R^D is defined as containing the intervals in D corresponding to all empty spaces in C, and given that C is countable, it follows that the empty spaces of C are themselves countable.

However, it is impossible for the empty spaces to be countable in R^D while generating an uncountable number of distinct configurations in T_{es}. This contradiction invalidates the assumption that for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$. Therefore, the countability of C implies the existence of some $K \in \mathbb{N}$ such that for all $k \geq K$, $\beta^{(\mathrm{a})k} \leq 1$. This completes the proof. □

Remark 3.2. *In contrast to the paths of the tree of empty spaces, those of the tree of convergent R-Cz sequences, which correspond to a backward traversal of these sequences, are not endless. They all eventually terminate in a term belonging to* $6\mathbb{N} + 3$ *(which has no preceding terms), however this will not be demonstrated here, as it is not necessary.*

3.4.4 Distances

Definition 3.6. *Let* $\{s_{j_0,i_0}\}$ *and* $\{s_{j_1,i_1}\}$ *be two Pt-sequences such that* $s_{j_0,0} < s_{j_1,0}$*. Their distance is defined as the distance between their first term, as follows:*

$$d\left(\{s_{j_0,i_0}\}, \{s_{j_1,i_1}\}\right) = s_{j_1,0} - s_{j_1,0} - o_{j_0,j_1}$$

Where o_{j_0,j_1} *denotes twice the number of terms between* $s_{j_0,0}$ *and* $s_{j_1,0}$*, that belong to other Pt-sequences, excluding the first term of these latter. The average distance between two consecutive Pt-sequences, among the first* $k+1$ *ordered in ascending order, is thereby:*

$$d^{(\mathrm{a})k} = d\left(\{s_{j,i}\}_{0\leq j\leq k}\right) = \frac{1}{k}\left(\left(\sum_{j=0}^{k-1} s_{j+1,0} - s_{j,0}\right) - o_{0,k}\right) = \frac{s_{k,0} - s_{0,0} - o_{0,k}}{k}$$

Remark 3.3. *Insofar as the objective is to determine whether the Pt-sequences move away from each other depending on the value of* $\beta^{(\mathrm{a})k}$*, this distance only accounts for the* *possible empty spaces separating the first terms of the Pt sequences.*

Lemma 3.7. *If for all $K \in \mathbb{N}$ and for all $r > 1$, there exist $k_0, k_1 \in I_{Pt}$, with $k_0 \geq K$ and $k_0 < k_1$, such that $\frac{d^{(\mathrm{a})k_1}}{d^{(\mathrm{a})k_0}} > r$, then for all K, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$.*

Proof. By contraposition, suppose there exists $K \in \mathbb{N}$ such that for all $k \geq K$, $\beta^{(\mathrm{a})k} \leq 1$. Let k_0 and $k_1 \in I_{Pt}$, with $k_0 \geq K$ and $k_0 < k_1$, then:

$$\frac{d^{(\mathrm{a})k_1}}{d^{(\mathrm{a})k_0}} = \frac{k_0}{k_1} + \frac{s_{k_1,0} - s_{k_0,0} - o_{k_0,k_1}}{k_1 d^{(\mathrm{a})k_0}}$$

Since the distance between two consecutive Pt-sequences is at least $2^{(\mathrm{i})}$, and given that $k_0 \geq K$ and $k_1 \geq K$, implying that $\beta^{(\mathrm{a})k_0} \leq 1$ and $\beta^{(\mathrm{a})k_1} \leq 1$, it follows that there exists $c \geq 2$ (independently of k_0 and/or k_1), such that $s_{k_1,0} - s_{k_0,0} - o_{k_0,k_1} \leq c(k_1 - k_0)$ (see Appendix 6.4). Hence:

$$\frac{d^{(\mathrm{a})k_1}}{d^{(\mathrm{a})k_0}} \leq \frac{k_0}{k_1} + \frac{c}{d^{(\mathrm{a})k_0}} \left(1 - \frac{k_0}{k_1}\right)$$

Given that $k_0 < k_1$ and $d^{(a)k_0} \geq 2$, we have:

$$\frac{d^{(\mathrm{a})k_1}}{d^{(\mathrm{a})k_0}} < 1 + \frac{c}{2} \iff \frac{d^{(\mathrm{a})k_1}}{d^{(\mathrm{a})k_0}} \leq 1 + \frac{c}{2} - \varepsilon$$

We therefore conclude that there exist $K \in \mathbb{N}$ and $(r = 1 + \frac{c}{2} - \varepsilon) > 1$, with $\varepsilon \in]0, 1[$, such that for all k_0 and $k_1 \in I_{Pt}$, with $k_0 \geq K$ and $k_0 < k_1$, we have $\frac{d^{(\mathrm{a})k_1}}{d^{(\mathrm{a})k_0}} \leq r$. $\square$

3.5 Natural density of C

In this section, we will study the natural density of C relative to $2\mathbb{N}+1$ (see [1]). To this end, we will prove that the size of the eventual empty spaces is bounded.

Theorem 3.6. *If there exists $K \in \mathbb{N}$ such that for all $k \geq K$, $\beta^{(\mathrm{a})k} \leq 1$, then there exists $L \in \mathbb{N}^*$ such that for all $i \in \mathbb{N}$, $\mathrm{card}(R_i^D) \leq L$.*

Proof. By contraposition, suppose that for all $L \in \mathbb{N}^*$, there exists $i \in \mathbb{N}$ such that $\mathrm{card}(R_i^D) > L$.

This implies that among the empty spaces of C (i.e.,the intervals of D), infinitely many of them grow without bound, and consequently there exist $i_0 < i_1 < \cdots < i_n < \cdots$ such that:

$$\mathrm{card}(R_{i_0}^D) < \mathrm{card}(R_{i_1}^D) < \cdots < \mathrm{card}(R_{i_n}^D) < \cdots$$

Since these empty spaces have the effect of pushing the Pt-sequences further and further apart, it follows that the average distance between two consecutive Pt-sequences increases.

(i) Because $C \subset 2\mathbb{N} + 1$.

Then, for all $K \in \mathbb{N}$ and for all $r > 1$, there exist k_0 and $k_1 \in I_{Pt}$, with $k_0 \geq K$ and $k_0 < k_1$, such that $\frac{d^{(\mathrm{a})k_1}}{d^{(\mathrm{a})k_0}} > r$.

Therefore, by Lemma 3.7, for all K there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$, which demonstrates the theorem. □

Definition 3.7. *Let* $d_{(\mathrm{odd})} : \mathcal{P}(2\mathbb{N}+1) \to [0,1]$ *be the natural density relative to the set* $2\mathbb{N}+1$, *defined for all* $A \in \mathcal{P}(2\mathbb{N}+1)$ *as follows:*

$$d_{(\mathrm{odd})}(A) = \lim_{n\to+\infty} \frac{N_{(\mathrm{odd}),n}(A)}{N_{(\mathrm{odd}),n}(2\mathbb{N}+1)}$$

Where $N_{(\mathrm{odd}),n}(A) = \mathrm{card}(A \cap \{1,3,\ldots,2n+1\})$ *is the number of odd numbers in A, and* $N_{(\mathrm{odd}),n}(2\mathbb{N}+1) = \mathrm{card}(\{1,3,\ldots,2n+1\}) = n$*) in* $2\mathbb{N}+1$, *both restricted to numbers between* 1 *and* $2n+1$.

Theorem 3.7. *If there exists* $K \in \mathbb{N}$ *such that for all* $k \geq K$, $\beta^{(\mathrm{a})k} \leq 1$, *then* $d_{(\mathrm{odd})}(C) > 0$.

Proof. Once we have established, thanks to Theorem 3.5, that the countability of C implies there exists $K \in \mathbb{N}$ such that for all $k \geq K$, $\beta^{(\mathrm{a})k} \leq 1$, it follows that the Pt-sequences are close to one another, in the sense that the size of any eventual gaps between any group of them is bounded. This is precisely what Theorem 3.6 asserts.

If we were to suppose the opposite, it would mean that as we move through the Pt-sequences, the average distance between them becomes larger and larger, which, by Theorem 3.6, would imply that for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$, and in turn, by Theorem 3.5, would imply that the empty spaces are not countable.

Therefore, since Theorem 3.6 states that there exists $L \in \mathbb{N}^*$ such that the size of any empty space of C is at most L, it follows that in every interval of $L+1$ consecutive odd numbers, there is at least one element of C. Then, we have:

$$d_{(\mathrm{odd})}(C) = \lim_{n\to+\infty} \left(\frac{N_{(\mathrm{odd}),n}(C)}{N_{(\mathrm{odd}),n}(2\mathbb{N}+1)} \right) = \lim_{n\to+\infty} \left(\frac{\frac{n}{(L+1)}}{n} \right) = \frac{1}{L+1} > 0$$

This completes the proof. □

4 All R-Cz sequences converge to 1

Terence Tao proved in 2021 that divergent Collatz sequences are statistically negligible (see [7]). Considering the theorems proven so far, we have all necessary elements to demonstrate that all R-Cz sequences converge to 1.

4.1 Preliminaries

According to Definition 1.2, $R^{\mathrm{d}}_{C_z}$ contains all R-Cz sequences that do not converge, and as the complement of C in $2\mathbb{N}+1$, D contains all terms of these sequences. Therefore, as a corollary of Theorem 1.1, D is in bijection with $R^{\mathrm{d}}_{C_z}$.

Since the empty spaces of C are the R^D_i intervals of D (see Definition 3.5), D is defined as follows:

$$D = \bigcup_{i \in E_s} R^D_i$$

We know from Theorem 3.6 that there exists $L \in \mathbb{N}^*$ such that the size of any empty space of C is at most L. Therefore, the size of any R^D_i interval is at most L as well.

Lemma 4.1. *(I) Let $\{d^p_l\}_{l \in \mathcal{N}}$ be a divergent R-Cz sequence. Then, there exists a sequence of indices $i_0, i_1, \ldots, i_k, \ldots$, such that:*

$$d^p_{i_0} < d^p_{i_1} < \cdots < d^p_{i_k} < \cdots$$

We denote by $\{d^p_{i_k}\}_{k \in \mathbb{N}}$ this subsequence of $\{d^p_l\}$. (II) Only terms in $4\mathbb{N}+3$ can cause the sequence $\{d^p_l\}$ to increase.

Proof. Concerning (I), since the sequence $\{d^p_l\}$ diverges to infinity, an infinite number of its terms become larger and larger.

Therefore, there exists a sequence of indices $i_0, i_1, \ldots, i_k, \ldots$, such that:

$$d^p_{i_0} < d^p_{i_1} < \cdots < d^p_{i_k} < \cdots$$

Concerning (II), let d^p_i be a term of $\{d^p_l\}$. Since d^p_i is an odd number, it must be of the form $4n+1$ or $4n+3$, where $n \in \mathbb{N}$. If $d^p_i = 4n+1$:

$$d^p_{i+1} = \frac{3d^p_i + 1}{2^\alpha} = \frac{3(4n+1)+1}{2^\alpha} = \frac{12n+4}{2^\alpha} = \frac{6n+2}{2^{\alpha-1}} = \frac{3n+1}{2^{\alpha-2}}$$

Hence, $\alpha \geq 2$ and $d^p_{i+1} < d^p_i$. Otherwise, if $d^p_i = 4n+3$:

$$d^p_{i+1} = \frac{3d^p_i + 1}{2^\alpha} = \frac{3(4n+3)+1}{2^\alpha} = \frac{12n+10}{2^\alpha} = \frac{6n+5}{2^{\alpha-1}}$$

Noting that $6n$ is even and 5 is odd, $\alpha = 1$ and then $(d^p_{i+1} = 6n+5) > (d^p_i = 4n+3)$. $\square$

Theorem 4.1. *If D is supposed to be non-empty, then all R-Cz sequences of $R^{\mathrm{d}}_{C_z}$ diverge to infinity.*

Proof. Suppose D is non-empty. Then, R^{d}_{Cz} contains at least one R-Cz sequence that does not converge.

On the one hand, Theorems 2.1 to 2.3 state that no R-Cz sequence has a cycle other than the trivial cycle of length 1 and value (1), implying in view of Definition 1.1 that all its terms are distinct (except when the sequence reaches 1). On the other hand, D is discrete.

It follows that if an R-Cz sequence does not converge to 1, then it must diverge to infinity. Therefore, all R-Cz sequences of R^{d}_{Cz} diverge to infinity. □

4.2 Case 1: D is a finite union of intervals

Although we have not proved that D is composed of a finite number of R_i^D intervals, we will address this case without recourse to theorems or lemmas.

Suppose there exists a divergent R-Cz sequence $\{d_l^p\}$, from which we extract its subsequence $\{d_{i_k}^p\}$. Since $\{d_{i_k}^p\}$ diverges to infinity, as stated in Lemma 4.1, we have:

$$d_{i_0}^p < d_{i_1}^p < \cdots < d_{i_k}^p < \cdots$$

However, since the R_i^D intervals are finite in number, there exists $k > 0$ such that for all $i \geq 0$, we have $d_{i_k}^p \notin R_i^D$. Therefore, there cannot not exist any divergent R-Cz sequence, and D is empty.

4.3 Case 2: D is an infinite union of intervals

Definition 4.1. *Let $\{v_l^p\}_{l\in\mathcal{N}}$ be an R-Cz sequence and v_i^p be one of its terms. Then, $T^{v_i^p}$ is the tree structure associated with the term v_i^p, which contains all its predecing terms at all levels of descent (all its descendants), v_i^p being the root of the tree. That is, all terms that precede v_i^p within R-Cz sequences, arranged in a tree structure (see Section 3.3).*

Remark 4.1. *If the first term v_0^p of an R-Cz sequence is in $6\mathbb{N}+3$, then, since terms in $6\mathbb{N}+3$ have no preceding terms, $T^{v_0^p} = v_0^p$.*

Theorem 4.2. *Let v_i^p be a term of the R-Cz sequence $\{v_l^p\}$ and $v_{i'}^{p'}$ be a term of the R-Cz sequence $\{v_{l'}^{p'}\}$, the two sequences may be the same but $v_i^p \neq v_{i'}^{p'}$.*

Then, both $T^{v_i^p}$ and $T^{v_{i'}^{p'}}$ contain an infinite number of terms, and if there is no ascending or descending path from v_i^p to $v_{i'}^{p'}$, we have $T^{v_i^p} \cap T^{v_{i'}^{p'}} = \emptyset$.

Otherwise, if there is an ascending path from v_i^p to $v_{i'}^{p'}$ (respectively, from $v_{i'}^{p'}$ to v_i^p), then $T^{v_i^p} \subsetneq T^{v_{i'}^{p'}}$ (respectively, $T^{v_{i'}^{p'}} \subsetneq T^{v_i^p}$), and the set $T^{v_{i'}^{p'}} \setminus T^{v_i^p}$ (respectively, $T^{v_i^p} \setminus T^{v_{i'}^{p'}}$) contains infinitely many terms of R-Cz sequences.

Proof. First, according to Theorem 3.3, both $T^{v_i^p}$ and $T^{v_{i'}^{p'}}$ are infinite sets, and if there is no ascending or descending path from v_i^p to $v_{i'}^{p'}$ (i.e., no R-Cz sequence connecting one term to the other), then Theorem 3.4 states that $T^{v_i^p} \cap T^{v_{i'}^{p'}} = \emptyset$.

Otherwise, if there is an ascending path from v_i^p to $v_{i'}^{p'}$ (respectively, from $v_{i'}^{p'}$ to v_i^p), then, considering v_i^p as a descendant of $v_{i'}^{p'}$ (respectively, $v_{i'}^{p'}$ as a descendant of v_i^p), we have $T^{v_i^p} \subsetneq T^{v_{i'}^{p'}}$ (respectively, $T^{v_{i'}^{p'}} \subsetneq T^{v_i^p}$).

It also follows, since Theorem 3.3 states that $v_{i'}^{p'}$ (respectively, v_i^p) is the parent term of infinitely many preceding terms, that $T^{v_{i'}^{p'}} \setminus T^{v_i^p}$ (respectively, $T^{v_i^p} \setminus T^{v_{i'}^{p'}}$) contains infinitely many terms of R-Cz sequences. $\square$

Proposition 4.1. *If D is supposed to be non-empty, then R^{d}_{Cz} contains infinitely many divergent R-Cz sequences, arranged in a tree structure whose root lies at infinity*[(i)].

Proof. Suppose D is non-empty. Then, by Theorem 4.1 R^{d}_{Cz} contains at least one divergent R-Cz sequence.

Let $\{d_l^p\}$ be this sequence and let d_k^p be one of its terms (with $k \geq 0$ to avoid d_k^p belonging to $6\mathbb{N}+3$), and let $T^{d_k^p}$ be the its associated tree (see Definition 4.1).

As stated by Theorem 3.3, d_k^p has infinitely many distinct preceding terms that constitute level 1 of $T^{d_k^p}$, in turn each term of level 1 has infinitely many distinct preceding terms that constitute level 2 of $T^{d_k^p}$, and so on (d_k^p being the root).

Given that each such term lies in a distinct R-Cz sequence and that these sequences merge at the term d_k^p, it follows that they must be divergent as well.

By the same argument, the term d_{k+1}^p is associated with the tree $T^{d_{k+1}^p}$, and like d_k^p it has infinitely many distinct preceding terms that constitute level 1 of $T^{d_{k+1}^p}$ (among which is d_k^p), etc. And as with d_k^p, all these terms belong to infinitely many divergent R-Cz sequences.

The same holds for the subsequent terms of $\{d_l^p\}$. Therefore, we have (with $n > 1$):

$$T^{d_k^p} \subsetneq T^{d_{k+1}^p} \subsetneq \cdots \subsetneq T^{d_{k+n}^p} \subsetneq \cdots$$

As the sequence $\{d_l^p\}$ diverges to infinity, this endless chain of strict inclusions yields infinitely many divergent R-Cz sequences, arranged in a tree structure growing upward without bound, whose root must therefore lie at infinity. $\square$

Thus, the divergent R-Cz sequences correspond to the paths of a tree structure that start from a node, or a leaf when their first term is in $6\mathbb{N}+3$, and ascend endlessly toward the

[(i)] We speak of the "root at infinity" as a metaphor, meaning simply that one can climb indefinitely up the tree structure toward the root, without ever reaching it.

root. The figure below illustrates this tree structure.

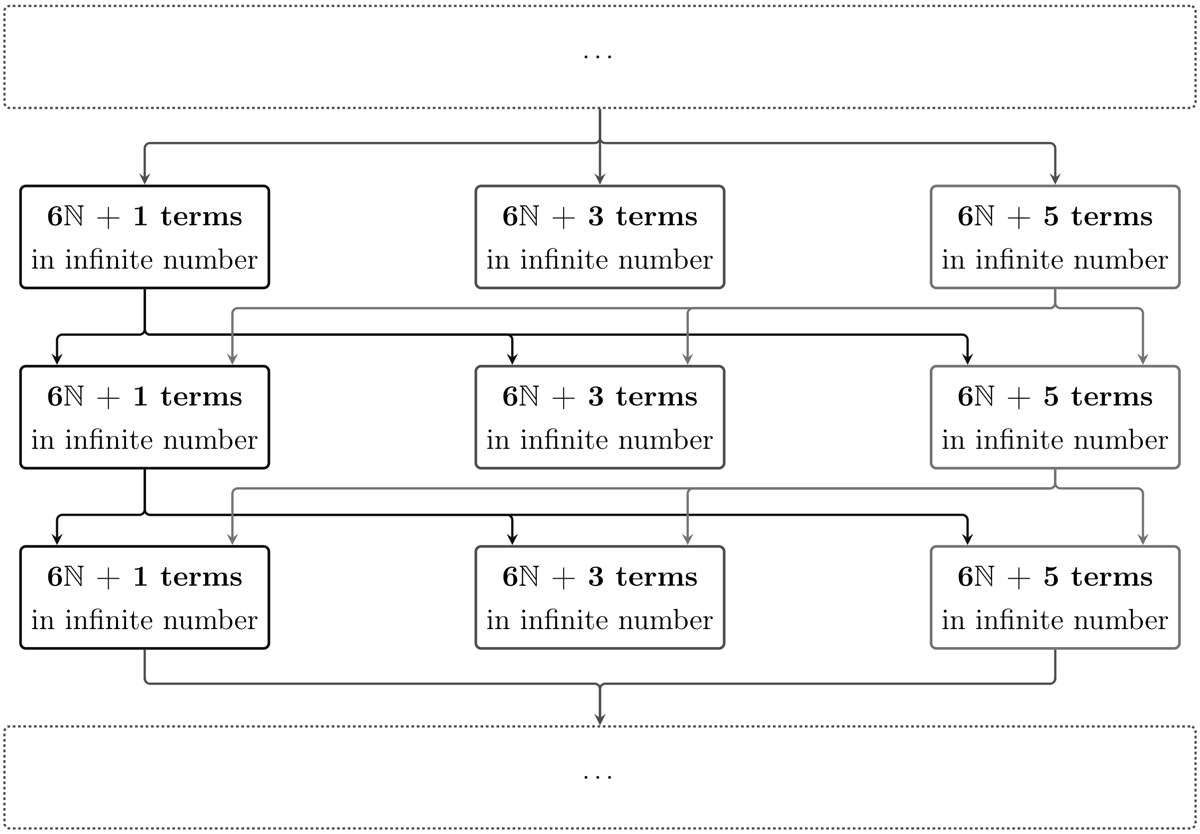

Figure 4. Tree structure of the divergent R-Cz sequences

Proposition 4.2. *The set R^{d}_{Cz} cannot contain infinitely many divergent R-Cz sequences, arranged in a tree structure whose root lies at infinity.*

First proof (set-theoretic). We will proceed by direct proof.

Hypothesis

Let $\{d_l^{p_x}\}$ be a divergent R-Cz sequence and let $d_{-1}^{p_x}$ be its first term, then according to Proposition 4.1:

$$T^{\mathrm{d}} = \bigcup_{l=-1}^{+\infty} T^{d_l^{p_x}}$$

is the union of the associated trees with all terms of $\{d_l^{p_x}\}$, such that:

$$T^{d_{-1}^{p_x}} \subsetneq T^{d_0^{p_x}} \subsetneq \cdots \subsetneq T^{d_k^{p_x}} \subsetneq \cdots$$

Uncountability

The tree T^{d} is an infinitely branching tree whose root lies at infinity. As in the case of the tree of the empty spaces in the proof of Theorem 3.5, the paths of T^{d}, which correspond to distinct divergent R-Cz sequences, are endless, but here, they never end when traced back toward the root.

Indeed, let p^n be such a path starting from a node n. Since the tree extends infinitely in both directions, from the root to the nodes and from the nodes to the root, level counting starts at 0 at the starting node and increases as one moves toward the root. Thus, p^n is of the form $(p_0^n, \ldots, p_z^n, \ldots) \in \mathbb{N}^{\mathbb{N}}$, where p_i^n is the local index of the i-th node (at level i) along the path, with $i \in \mathbb{N}$. However, a complete enumeration of all paths is impossible, as demonstrated below.

It must first be noted that Theorems 3.3 and 3.4, and Lemma 3.2 apply to all R-Cz sequences, including the divergent ones. Then, let $f : P \to \mathbb{N}^{\mathbb{N}}$ be a function defined as follows:

$$f(p^i) = (p_0^i, \ldots, p_z^i, \ldots)$$

where $P = \{p^i = (p_0^i, \ldots, p_z^i, \ldots)\}_{i \in \mathbb{N}}$ is the set of the paths from all nodes of T_{div} to the root. Let $(p^i, p^j) \in P^2$, with $p^i \neq p^j$, then there exists k such that $p_k^i \neq p_k^j$, and we have:

$$f(p^i) = (p_0^i, \ldots, p_k^i, \ldots) \neq (p_0^j, \ldots, p_k^j, \ldots) = f(p^j)$$

Hence, f is injective. Regarding surjectivity, let $(n_0, \ldots, n_k, \ldots) \in \mathbb{N}^{\mathbb{N}}$.

Because the root of T^{d} lies at infinity, and due to the recursive nature of Theorem 3.3, which states that each parent term in $6\mathbb{N}+1$ or $6\mathbb{N}+5$ generates infinitely many children terms[(i)], which in turn generate infinitely many others, and so on *ad infinitum*, as a result there are infinitely many terms (i.e., nodes) at each level of the tree, all distinct as stated in Lemma 3.2.

Consequently, we can construct for all $k \in \mathbb{N}$, a reverse path $(\ldots, p_k^y, \ldots, p_0^y)$, from the root to the node y, where for all $j > 0$, the node p_j^y is the parent of the node p_{j-1}^y, such that reversing it yields the path $p^y = (p_0^y, \ldots, p_k^y, \ldots)$, and such that:

$$f(p^y) = (p_0^y, \ldots, p_k^y, \ldots) = (n_0, \ldots, n_k, \ldots)$$

Hence, f is bijective. Since $\mathbb{N}^{\mathbb{N}}$ is uncountable and f is a bijection from P to $\mathbb{N}^{\mathbb{N}}$, it follows that P is also uncountable, and the complete enumeration of all paths is therefore impossible.

Conclusion

Since Theorem 1.1 states that R_{Cz}^{d} is in bijection with the countable set D, it follows that R_{Cz}^{d} is also countable. However, insofar as each path in T^{d} corresponds to a divergent R-Cz sequence, the uncountability of these paths highlights a contradiction, as an infinite number of paths, composed of infinite sequences of nodes, cannot be reduced to a countable set of finite indexes, and thus cannot be contained in a countable set.

(i) We recall that they lie alternately in $6\mathbb{N}+1$, $6\mathbb{N}+3$ and $6\mathbb{N}+5$.

This leads us to conclude that R^{d}_{Cz} cannot contain infinitely many divergent R-Cz sequences, arranged in a tree structure whose root lies at infinity. $\square$

Second proof (combinatorial). As Theorem 3.6 states that the size of the R_i^D intervals are bounded, there exists $L \in \mathbb{N}^*$ such that for all $i \geq 0, R_i^D \leq L$.

It follows that there exists $n_d \leq L$, such that at least n_d divergent R-Cz sequences can lie within the R_i^D intervals. Now, since the root lies at infinity and there are infinitely many distinct nodes at each level of the tree, there exists a subfamily $R^{\mathrm{dd}}_{Cz} = \{\{d_l^{p_j}\}_{j\in\mathbb{N}} \subset R^{\mathrm{d}}_{Cz}$ such that for all $X \in \mathbb{N}$ and for all $(j_0, j_1) \in \mathbb{N}^2$:

$$\left\{d_{l_0}^{p_{j_0}}\right\}_{0\leq l_0 \leq X} \cap \left\{d_{l_1}^{p_{j_1}}\right\}_{0\leq l_1 \leq X} = \emptyset.$$

In other words, there exist infinitely many pairwise disjoint divergent R-Cz sequences. We will refer to these sequences as *DR-Cz sequences*. Then, let:

$$I_k = \bigcup_{n=0}^{k} \left\{R_i^D \in R^D : \exists j \geq 0 \text{ such that } d_n^{p_j} \in R_i^D\right\}$$

be the set of R_i^D intervals which contain the k first terms of all DR-Cz sequences, ordered in ascending order.

For $k = 0$, since the size of the R_i^D intervals is bounded, I_0 contains infinitely many of these intervals, and some other R_i^D intervals can lie between the intervals of I_0.

For $k \geq 1$, since the term $d_k^{p_j}$ of the sequence $\{d_k^{p_j}\}$ is at a determined distance from the term $d_{k-1}^{p_j}$[(i)], it must lie either in the same interval as the term $d_{k-1}^{p_j}$, or in one of the intervals that precedes or follows it.

Insofar as it applies for all j, the terms $d_k^{p_j}$ of the DR-Cz sequences can lie in infinitely many new R_i^D intervals, and we have the following chain of inclusion:

$$I_0 \subseteq I_1 \subseteq \cdots \subseteq I_{k-1} \subseteq I_k$$

Finally, considering (i) that the length of the R_i^D intervals is at most L, and (ii) that there is at most a finite number of R_i^D intervals between each interval of I_0, there exists a rank $K \geq 1$ in the DR-Cz sequences, from which all R_i^D intervals will completely be filled by terms of the DR-Cz sequences, and we will have:

$$I_K = I_{K+1} = I_{K+2} = \cdots$$

[(i)] if $d_{k-1}^{p_j} < d_k^{p_j}$ then $d_k^{p_j} = \frac{3d_{k-1}^{p_j}+1}{2}$.

Therefore, we conclude that R^{d}_{Cz} cannot contain infinitely many divergent R-Cz sequences, arranged in a tree structure whose root lies at infinity[(i)]. $\square$

Remark 4.2. *The reader will note that a single divergent R-Cz sequence is sufficient to generate a tree structure, giving rise to an uncountable number of divergent R–Cz sequences.*

Theorem 4.3. *If D is supposed to be non-empty, the contradiction that arises between Propositions 4.1 and 4.2, leads to the conclusion that D must be empty. Consequently, all R-Cz sequences converge to* 1.

Proof. Suppose D is non-empty. Then Proposition 4.1 states that $R^{\mathrm{d}}_{C_z}$ contains infinitely many divergent R-Cz sequences, whereas Proposition 4.2 states that $R^{\mathrm{d}}_{C_z}$ cannot contain such a number of divergent R-Cz sequences. These two propositions cannot be true simultaneously.

The only way to resolve this contradiction is to conclude that D must be empty. Indeed, if D is empty, there is no longer any possibility of an infinite number of divergent R-Cz sequences arranged in a tree structure, and there is no longer a contradiction between Propositions 4.1 and 4.2; both propositions are true. There is no other possible resolution.

As D is empty, it follows that $C = 2\mathbb{N} + 1$, which allows us to conclude that all R-Cz sequences converge to 1. $\square$

Remark 4.3. *The proof of Theorem 3.5 and the set-theoretic proof of Proposition 4.2 together constitute a two-fold cardinality argument, indeed:*

- *The proof of Theorem 3.5 demonstrates that even purely convergent R-Cz sequences, if spaced, give rise to an uncountable tree of empty spaces;*
- *The set-theoretic proof of Proposition 4.2 demonstrates that any hypothetical divergent R-Cz sequence would likewise generate an uncountable tree of divergent R-Cz sequences.*

In either scenario, spacing or divergence, the same cardinality contradiction arises, reinforcing the consistency of the overall approach.

It is also worth noting that even in the absence of Theorem 3.5, and thus without considering the structure of D, the set-theoretic proof of Proposition 4.2 alone suffices to establish the convergence of all R-Cz sequences.

[(i)] Combinatorially speaking, we could say that $\aleph_0 \times \aleph_0 \gg \aleph_0 \times L$.

Ultimately, beyond the R-Cz sequences, this paper reveals a hierarchical relationship between all odd numbers through the recursive formula which characterizes the Pt-sequences:

$$s_{j,i} = \begin{cases} \frac{2}{3} q_j \cdot 4^i - \frac{1}{3}, & \text{if } q_j \in 6\mathbb{N} + 5 \\ \frac{4}{3} q_j \cdot 4^i - \frac{1}{3}, & \text{if } q_j \in 6\mathbb{N} + 1 \end{cases}$$

Where $(j, i) \in \mathbb{N}^2$ and the numbers $s_{j,i}$ are the children of the parent number q_j. The figure below illustrates these relationships, with the number 1 as the ancestor of all odd numbers.

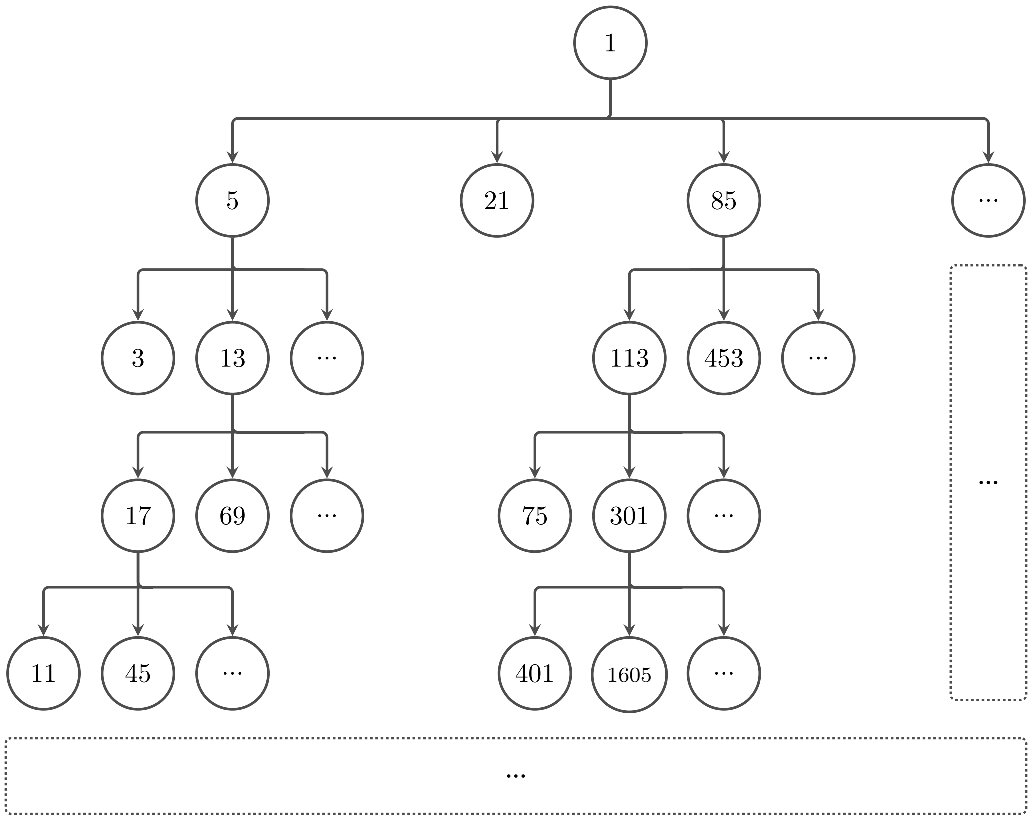

Figure 5. Tree of odd numbers

Since $s_{0,0} = \frac{4}{3} \times 1 \times 4^0 - \frac{1}{3} = 1$, the number 1 is its own parent, and numbers such as $21, 3, 453, 69, 75, \ldots,$ since they are in $6\mathbb{N} + 3$, do not have any child numbers.

5 Conclusion

As the first step, we defined the R-Cz sequences, which are the Collatz sequences without their even terms. As the first essential result, we established that the R-Cz sequences do not exhibit any cycles other than the cycle of length 1 and value (1). We then studied their properties in the case where they converge, particularly highlighting their tree structure, which can be viewed either as a collection of sequences converging to the penultimate terms and then to the root, or as the descendants generated through the Pt-sequences.

This led us to establish Theorem 3.5, which constitutes the central result of Section 3. Then, the consequences that can be inferred from this theorem, when combined with the fact that no sequence has any cycle other than the trivial one, gave rise to a contradiction regarding the set R^{d}_{Cz}, which is supposed to contain all divergent R-Cz sequences.

Finally, this contradiction could only be resolved through the realization that D is empty, leading to the conclusion that all R-Cz sequences converge to 1. The main steps of the proof can be represented graphically as follows, where the black arrows indicate the mathematical implications.

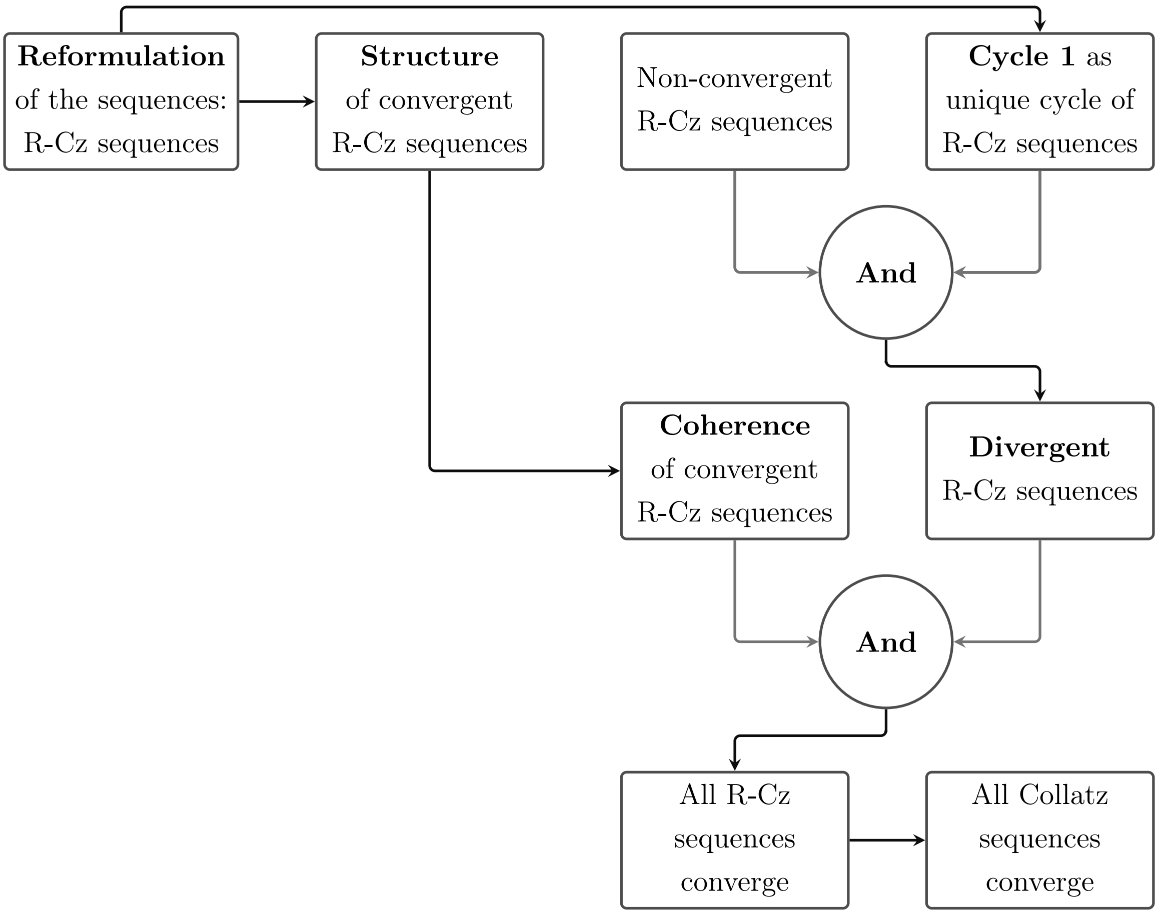

Figure 6. Diagram of the proof

Since an R-Cz sequence is derived from a Collatz sequence by removing its even terms, the fact that all R-Cz sequences converge to 1 demonstrates that all Collatz sequences eventually reach the cycle $(1, 4, 2)$. This completes the proof of the Collatz conjecture.

6 Appendix

6.1 General term of the sequence $\{v_l^p\}$

We know that the formula is valid for the third term in the sequence. Suppose that for $l < L$, where $L \geq 2$:

$$v_l^p = \frac{3^l(3p+1) + \sum_{i=0}^{l-1}\left(3^{l-1-i} \cdot 2^{\sum_{j=0}^{i}\alpha_j}\right)}{2^{\sum_{i=0}^{l}\alpha_i}}$$

And calculate v_{l+1}^p as a function of v_l^p:

$$v_{l+1} = \frac{3v_l^p + 1}{2^{\alpha_{l+1}}} = \frac{3\left(\frac{3^l(3p+1)+\sum_{i=0}^{l-1}\left(3^{l-1-i}.2^{\sum_{j=0}^{i}\alpha_j}\right)}{2^{\sum_{i=0}^{l}\alpha_i}}\right)+1}{2^{\alpha_{l+1}}}$$

And we get:

$$v_{l+1} = \frac{3^{l+1}(3p+1) + \sum_{i=0}^{l}\left(3^{l-i} \cdot 2^{\sum_{j=0}^{i}\alpha_j}\right)}{2^{\sum_{i=0}^{l+1}\alpha_i}}$$

Therefore, for all $l \in \mathbb{N}^*$:

$$v_l^p = \frac{3^l(3p+1) + \sum_{i=0}^{l-1}\left(3^{l-1-i} \cdot 2^{\sum_{j=0}^{i}\alpha_j}\right)}{2^{\sum_{i=0}^{l}\alpha_i}}$$

$\square$

6.2 Classes of penultimate terms

The fact that the penultimate term $v_{b_p}^n = \frac{2^{2n}-1}{3}$, where $n \in \mathbb{N}^*$, $n = 3k + a$ and $k \in \mathbb{N}$, lies in $6\mathbb{N}+3$, $6\mathbb{N}+1$ or $6\mathbb{N}+5$ depends on whether $a = 0$, $a = 1$ or $a = 2$.

First case $a = 0$

The penultimate term $\left(v_{b_p}^n = \frac{2^{2(3k)}-1}{3}\right) \in 6\mathbb{N}+3$, if there exists $k' \in \mathbb{N}$ such that $\frac{2^{6k}-1}{3} = 3 + 6k'$, which is equivalent to:

$$2^{6k-1} - 5 \equiv 0 \pmod 9$$

For $k = 1$, $2^5 - 5 = 27 \equiv 0 \pmod 9$, and suppose that for $k = k_0$, where $k_0 \in \mathbb{N}^* \setminus \{1\}$, we have $2^{6k_0-1} - 5 \equiv 0 \pmod 9$.

Then, for $k = k_0 + 1$, we have:

$$2^{6(k_0+1)-1} - 5 \equiv \left(2^{6k_0-1} - 5\right) + 63 \cdot 2^{6k_0-1} \equiv 0 \pmod 9$$

Therefore, for all $k \in \mathbb{N}^*$, $\frac{2^{2(3k)}-1}{3} \in 6\mathbb{N} + 3$.

Second case $a = 1$
The penultimate term $\left(v^n_{b_p} = \frac{2^{2(3k+1)}-1}{3}\right) \in 6\mathbb{N}+1$, if there exists $k' \in \mathbb{N}$ such that $\frac{2^{6k+2}-1}{3} = 1 + 6k'$, which is equivalent to:

$$2^{6k+1} - 2 \equiv 0 \pmod 9$$

For $k = 0$, $2 - 2 \equiv 0 \pmod 9$, for $k = 1$, $2^7 - 2 \equiv 126 \equiv 0 \pmod 9$, and suppose that for $k = k_0$, where $k_0 \in \mathbb{N}^* \setminus \{1\}$, $2^{6k_0+1} - 2 \equiv 0 \pmod 9$.

Then, for $k = k_0 + 1$ we have:

$$2^{6(k_0+1)+1} - 2 \equiv (2^{6k_0+1} - 2) + 63 \cdot 2^{6k_0+1} \equiv 0 \pmod 9$$

Therefore, for all $k \in \mathbb{N}$, $\frac{2^{2(3k+1)}-1}{3} \in 6\mathbb{N} + 1$.

Third case $a = 2$
The penultimate term $\left(v^n_{b_p} = \frac{2^{2(3k+2)}-1}{3}\right) \in 6\mathbb{N}+5$, if there exists $k' \in \mathbb{N}$ such that $\frac{2^{6k+4}-1}{3} = 5 + 6k'$, which is equivalent to:

$$2^{6k+3} - 8 \equiv 0 \pmod 9$$

For $k = 0$, $8 - 8 \equiv 0 \pmod 9$, for $k = 1$, $2^9 - 8 \equiv 504 \equiv 0 \pmod 9$, and suppose that for $k = k_0$, where $k_0 \in \mathbb{N}^* \setminus \{1\}$, $2^{6k_0+3} - 8 \equiv 0 \pmod 9$.

Then, for $k = k_0 + 1$ we have:

$$2^{6(k_0+1)+3} - 8 \equiv (2^{6k_0+3} - 8) + 63 \cdot 2^{6k_0+3} \equiv 0 \pmod 9$$

Therefore, for all $k \in \mathbb{N}$, $\frac{2^{2(3k+2)}-1}{3} \in 6\mathbb{N} + 5$. $\square$

6.3 Classes of preceding terms

We are going to demonstrate that the term v^p_i of the R-Cz sequence $\{v^p_l\}$, when it is preceded by terms (i.e., when it is in $6\mathbb{N} + 1$ or $6\mathbb{N} + 5$), they are alternately in $6\mathbb{N} + 1$, $6\mathbb{N} + 3$ and $6\mathbb{N} + 5$.

First case $v^p_i \in 6\mathbb{N}^* + 1$
Since $v^p_i \in 6\mathbb{N}^* + 1$, its preceding terms are of the form $\frac{2^{\alpha_i}(6k+1)-1}{3}$, it follows that:

- if $k \equiv 0 \pmod 3$, its preceding terms are alternately $6\mathbb{N}+1$, $6\mathbb{N}+5$ and $6\mathbb{N}+3$, depending on whether $\alpha_i \in 6\mathbb{N}+2$, $6\mathbb{N}+4$ or $6\mathbb{N}^*$;
- if $k \equiv 1 \pmod 3$, its preceding terms are alternately $6\mathbb{N}+3$, $6\mathbb{N}+1$ and $6\mathbb{N}+5$, depending on whether $\alpha_i \in 6\mathbb{N}+2$, $6\mathbb{N}+4$ or $6\mathbb{N}^*$;
- and finally, when $k \equiv 2 \pmod 3$, its preceding terms are alternately $6\mathbb{N}+5$, $6\mathbb{N}+3$ and $6\mathbb{N}+1$, depending on whether $\alpha_i \in 6\mathbb{N}+2$, $6\mathbb{N}+4$ or $6\mathbb{N}^*$.

Because the demonstration is the same for $k \equiv 0 \pmod 3$, $k \equiv 1 \pmod 3$, and $k \equiv 2 \pmod 3$, we will only demonstrate the case $k \equiv 1 \pmod 3$.

For $(k, \alpha_i) \in \{(7,2),(7,4),(7,6)\}$, we have respectively:

$$\frac{2^2 \times 7 - 1}{3} = 9 \in 6\mathbb{N}+3,\ \alpha_i \in 6\mathbb{N}+2$$

$$\frac{2^4 \times 7 - 1}{3} = 37 \in 6\mathbb{N}+1,\ \alpha_i \in 6\mathbb{N}+4$$

$$\frac{2^6 \times 7 - 1}{3} = 149 \in 6\mathbb{N}+5,\ \alpha_i \in 6\mathbb{N}^*$$

Suppose that for $(k, \alpha_i) \in \{(7, 6\alpha_0+2),(7,6\alpha_0+4),(7,6\alpha_0)\}$, where $\alpha_0 \in \mathbb{N}^*$, we have respectively:

$$\frac{2^{6\alpha_0+2} \times 7 - 1}{3} \in 6\mathbb{N}+3 \iff 2^{6\alpha_0+1} \times 7 - 5 \equiv 0 \pmod 9$$

$$\frac{2^{6\alpha_0+4} \times 7 - 1}{3} \in 6\mathbb{N}+1 \iff 2^{6\alpha_0+3} \times 7 - 2 \equiv 0 \pmod 9$$

$$\frac{2^{6\alpha_0} \times 7 - 1}{3} \in 6\mathbb{N}+5 \iff 2^{6\alpha_0-1} \times 7 - 8 \equiv 0 \pmod 9$$

For $(k, \alpha_i) \in \{(7, 6(\alpha_0+1)+2),(7,6(\alpha_0+1)+4),(7,6(\alpha_0+1))\}$, we have respectively:

$$2^{6(\alpha_0+1)+1} \times 7 - 5 \equiv (2^{6\alpha_0+1} \times 7 - 5) + 63 \times 7 \equiv 0 \pmod 9$$

$$2^{6(\alpha_0+1)+3} \times 7 - 2 \equiv (2^{6\alpha_0+3} \times 7 - 2) + 63 \times 7 \equiv 0 \pmod 9$$

$$2^{6(\alpha_0+1)-1} \times 7 - 8 \equiv (2^{6\alpha_0-1} \times 7 - 8) + 63 \times 7 \equiv 0 \pmod 9$$

Suppose that for $(k, \alpha_i) \in \{(6k_0+1, \alpha_1),(6k_0+1,\alpha_2),(6k_0+1,\alpha_3)\}$, where $k_0 \in 3\mathbb{N}^*+1$ and $(\alpha_1,\alpha_2,\alpha_3) \in (6\mathbb{N}+2) \times (6\mathbb{N}+4) \times 6\mathbb{N}^*$, we have respectively:

$$2^{\alpha_1}(6k_0+1) - 5 \equiv 0 \pmod 9$$

$$2^{\alpha_2}(6k_0+1) - 2 \equiv 0 \pmod 9$$

$$2^{\alpha_3}(6k_0+1)-8 \equiv 0 \pmod 9$$

For $(k, \alpha_i) \in \{(6(k_0+3)+1, \alpha_1), (6(k_0+3)+1, \alpha_2), (6(k_0+3)+1, \alpha_3)\}$, we have respectively:

$$2^{\alpha_1}(6(k_0+3)+1)-5 \equiv (2^{\alpha_1}(6k_0+1)-5)+18\times 2^{\alpha_1} \equiv 0 \pmod 9$$

$$2^{\alpha_2}(6(k_0+3)+1)-2 \equiv (2^{\alpha_2}(6k_0+1)-2)+18\times 2^{\alpha_2} \equiv 0 \pmod 9$$

$$2^{\alpha_3}(6(k_0+3)+1)-8 \equiv (2^{\alpha_3}(6k_0+1)-8)+18\times 2^{\alpha_3} \equiv 0 \pmod 9$$

Therefore, when v_i^p is of the form $\frac{2^{\alpha_i}(6k+1)-1}{3}$, with $k \equiv 1 \pmod 3$, its preceding terms are alternately in $6\mathbb{N}+1$, $6\mathbb{N}+3$ and $6\mathbb{N}+5$, depending on whether $\alpha_i \in 6\mathbb{N}+4$, $6\mathbb{N}+2$ or $6\mathbb{N}^*$.

Second case $v_i^p \in 6\mathbb{N}+5$
Since $v_i^p \in 6\mathbb{N}+5$, its preceding terms are of the form $\frac{2^{\alpha_i}(6k+5)-1}{3}$, it follows that:

- if $k \equiv 0 \pmod 3$, its preceding terms are alternately in $6\mathbb{N}+3$, $6\mathbb{N}+1$ and $6\mathbb{N}+5$, depending on whether $\alpha_i \in 6\mathbb{N}+1$, $6\mathbb{N}+3$ or $6\mathbb{N}+5$;
- if $k \equiv 1 \pmod 3$, its preceding terms are alternately in $6\mathbb{N}+1$, $6\mathbb{N}+5$ and $6\mathbb{N}+3$, depending on whether $\alpha_i \in 6\mathbb{N}+1$, $6\mathbb{N}+3$ or $6\mathbb{N}+5$;
- and finally, when $k \equiv 2 \pmod 3$, its preceding terms are alternately in $6\mathbb{N}+5$, $6\mathbb{N}+3$ and $6\mathbb{N}+1$, depending on whether $\alpha_i \in 6\mathbb{N}+1$, $6\mathbb{N}+3$ or $6\mathbb{N}+5$.

Because the demonstration is the same for $k \equiv 0[3]$, $k \equiv 1[3]$, and $k \equiv 2[3]$, we will only demonstrate the case $k \equiv 0[3]$.

For $(k, \alpha_i) \in \{(5,1), (5,3), (7,5)\}$, we have respectively:

$$\frac{2^1\times 5-1}{3} = 3 \in 6\mathbb{N}+3, \quad \alpha_i \in 6\mathbb{N}+1$$

$$\frac{2^3\times 5-1}{3} = 13 \in 6\mathbb{N}+1, \quad \alpha_i \in 6\mathbb{N}+3$$

$$\frac{2^5\times 5-1}{3} = 53 \in 6\mathbb{N}+5, \quad \alpha_i \in 6\mathbb{N}+5$$

Suppose that for $(k, \alpha_i) \in \{(5, 6\alpha_0+1), (7, 6\alpha_0+3), (7, 6\alpha_0+5)\}$, where $\alpha_0 \in \mathbb{N}^*$, we have respectively:

$$\frac{2^{6\alpha_0+1}\times 5-1}{3} \in 6\mathbb{N}+3 \iff 2^{6\alpha_0}\times 5-5 \equiv 0 \pmod 9$$

$$\frac{2^{6\alpha_0+3}\times 5-1}{3} \in 6\mathbb{N}+1 \iff 2^{6\alpha_0+2}\times 5-2 \equiv 0 \pmod 9$$

$$\frac{2^{6\alpha_0+5}\times 5-1}{3}\in 6\mathbb{N}+5 \iff 2^{6\alpha_0+4}\times 5-8\equiv 0 \pmod 9$$

For $(k,\alpha_i)\in\{(7,6(\alpha_0+1)+1),(7,6(\alpha_0+1)+3),(7,6(\alpha_0+1)+5)\}$, we have respectively:

$$2^{6(\alpha_0+1)}\times 5-5\equiv(2^{6\alpha_0}\times 5-5)+63\times 5\equiv 0 \pmod 9$$

$$2^{6(\alpha_0+1)+2}\times 5-2\equiv(2^{6\alpha_0+2}\times 5-2)+63\times 5\equiv 0 \pmod 9$$
$$2^{6(\alpha_0+1)+4}\times 5-8\equiv(2^{6\alpha_0+4}\times 5-8)+63\times 5\equiv 0 \pmod 9$$

Suppose that for $(k,\alpha_i)\in\{(6k_0+5,\alpha_1),(6k_0+5,\alpha_2),(6k_0+5,\alpha_3)\}$, where $k_0\in 3\mathbb{N}^*+1$ and $(\alpha_1,\alpha_2,\alpha_3)\in(6\mathbb{N}+1)\times(6\mathbb{N}+3)\times(6\mathbb{N}+5)$, we have respectively:

$$2^{\alpha_1}(6k_0+5)-5\equiv 0 \pmod 9$$

$$2^{\alpha_2}(6k_0+5)-2\equiv 0 \pmod 9$$
$$2^{\alpha_3}(6k_0+5)-8\equiv 0 \pmod 9$$

Then for $(k,\alpha_i)\in\{(6(k_0+3)+5,\alpha_1),(6(k_0+3)+5,\alpha_2),(6(k_0+3)+5,\alpha_3)\}$, we have respectively:

$$2^{\alpha_1}(6(k_0+3)+5)-5\equiv(2^{\alpha_1}(6k_0+5)-5)+18\cdot 2^{\alpha_1}\equiv 0 \pmod 9$$

$$2^{\alpha_2}(6(k_0+3)+5)-2\equiv(2^{\alpha_2}(6k_0+5)-2)+18\cdot 2^{\alpha_2}\equiv 0 \pmod 9$$
$$2^{\alpha_3}(6(k_0+3)+5)-8\equiv(2^{\alpha_3}(6k_0+5)-8)+18\cdot 2^{\alpha_3}\equiv 0 \pmod 9$$

Therefore, when v_i^p is of the form $\frac{2^{\alpha_i}(6k+5)-1}{3}$, with $k\equiv 0 \pmod 3$, its preceding terms are alternately in $6\mathbb{N}+3$, $6\mathbb{N}+1$ and $6\mathbb{N}+5$, depending on whether $\alpha_i\in 6\mathbb{N}+1$, $6\mathbb{N}+3$ or $6\mathbb{N}+5$. □

6.4 Existence of c in Lemma 3.7

For all $(k_0,k_1)\in I_{Pt}\times I_{Pt}$, we have:

$$s_{k_1,0}-s_{k_0,0}=\sum_{j=k_0}^{k_1-1} q_{j+1}\beta^{q_{j+1}}-q_j\beta^{q_j}=\varepsilon_{k_1}+\sum_{j=k_0}^{k_1-1}\beta^{(\mathrm{a})k}(q_{j+1}-q_j)$$

Where $\varepsilon_{k_1}\in\mathbb{R}$ is a correction factor. Now, considering that for all $k\geq K$, $\beta^{(\mathrm{a})k}\leq 1$ and the Pt-sequences do not intersect or overlap, we obtain, after subtracting the term o_{k_0,k_1}

from the previous equation:

$$\varepsilon_{k_1} - o_{k_0,k_1} + \sum_{j=k_0}^{k_1-1} \beta^{(\mathrm{a})k}(q_{j+1} - q_j) \leq c_{k_1}(k_1 - k_0)$$

Where $c_{k_1} \geq 2$. If for all $L \in \mathbb{N}^*$, there existed k_1 such that $c_{k_1} > L$, this would contradict our hypothesis that for all $k \geq K$, $\beta^{(\mathrm{a})k} \leq 1$, which implies that the distance of Definition 3.6 between two successive Pt-sequences is necessarily bounded. Therefore, there exists $c \geq 2$ such that for all k_1, $c_{k_1} \leq c$. □

6.5 Empty spaces creation in the proof of Theorem 3.5

Under the hypothesis of the proof of Theorem 3.5, we will demonstrate that, regardless of the number of Pt-sequences generated, the creation of empty spaces is an endless process.

Unbounded empty spaces
Considering that C is discrete, and that new empty spaces are formed by subdividing previous ones into which terms are inserted, the creation of new empty spaces could only cease if their size became bounded. However, such a scenario cannot occur. Indeed, first by Lemma 3.4, for all $(j,i) \in \mathbb{N}^2$ (to simplify, here $I_{Pt} = \mathbb{N}$), we have:

$$s_{j,i} = \begin{cases} \frac{2}{3}q_j \cdot 4^i - \frac{1}{3} & \text{if } q_j \in 6\mathbb{N}+5 \\ \frac{4}{3}q_j \cdot 4^i - \frac{1}{3} & \text{if } q_j \in 6\mathbb{N}+1 \end{cases}$$

Where q_j is the parent term of the Pt-sequence $\{s_{j,i}\}$. Let $f_j = \beta q_j$, where $\beta = \frac{2}{3}$ or $\frac{4}{3}$ depending on whether $q_j \in 6\mathbb{N}+5$ or $6\mathbb{N}+1$.

Then, as we have assumed that for all $K \in \mathbb{N}$, there exists $k \geq K$ such that $\beta^{(\mathrm{a})k} > 1$, according to Lemma 3.6, the sequence $\{f_j\}_{j\in\mathbb{N}}$ grows on average exponentially.

Let $\{s_{j,i}\}_{0\leq j\leq J}$, with $J > 0$, be the family of the first J Pt-sequences ordered by their first term in ascending order. Then we have:

$$s_{0,0} < s_{1,0} < \cdots < s_{J,0}$$

Given that no Pt-sequence shares terms with another, there exists a sequence $\{(l_n, k_n)\}_{n\in\mathbb{N}}$ such that:

$$\{s_{l_0,k_0}, s_{l_1,k_1}, \ldots, s_{l_n,k_n}, \ldots\} = \bigcup_{j=0}^{J}\{s_{j,i}\}$$

And such that:

$$s_{l_0,k_0} < s_{l_1,k_1} < \cdots < s_{l_n,k_n} < \cdots$$

Finally, since for all $(j,i) \in \{0,\dots,J-1\} \times \mathbb{N}, s_{j+1,i} - s_{j,i} = 4^i(f_{j+1} - f_j)$, and for all $(j,i) \in \{0,\dots,J\} \times \mathbb{N}$, $s_{j,i+1} = 4s_{j,i} + 1$, it follows that the sequence $\{s_{l_n,k_n}\}_{n\in\mathbb{N}}$ grows on average exponentially.

Hence, for all $L \in \mathbb{N}^*$, there exist $(x,y) \in \mathbb{N}^2$ with $x+1 = y$, such that $\left]s_{l_x,k_x},\ s_{l_y,k_y}\right[\cap C = \emptyset$, and such that:

$$s_{l_y,k_y} - s_{l_x,k_x} - 1 > L$$

Therefore, as J can be taken arbitrarily large, no matter how many Pt-sequences are generated, there will always be empty spaces, at every level of the tree structure, whose size remains unbounded.

Cardinality

Let E_z be the set of empty spaces created from level 1 to z. At level 1, E_1 contains an infinite number of empty spaces, located between the penultimate terms. From level 2, and for subsequent levels, starting from level 2, two phenomena occur:

- the predominantly rightward shift of the Pt-sequences may prevent further subdivision of some empty spaces;
- as demonstrated above, there are infinitely many empty spaces of unbounded size.

Consequently, from level 2, depending on the Pt-sequences generated, each empty space can undergo one of four possible outcomes: either it ceases to be subdivided and has no descendants, or its size is reduced because terms of Pt-sequences are inserted at its edges, or it is subdivided into at least two parts, or finally, it remains unchanged and becomes its own descendant (i.e., it is present at the next level).

For all $z \geq 2$, let:

- E_z^{ns} be the set of empty spaces that cease to be subdivided at level z, due to the rightward shift of the Pt-sequences;
- E_z^{sub} be the set of empty spaces resulting from subdivision of empty spaces at level $z-1$;
- and finally let E_z^{fix} and E_z^{red} be the sets of empty spaces that are respectively unchanged or reduced relative to level $z-1$.

Then we have:

$$E_z = \bigcup_{i=2}^{z-1} E_i^{\text{ns}} \cup E_z^{\text{sub}} \cup E_z^{\text{fix}} \cup E_z^{\text{red}}$$

Considering that at each level $z \geq 2$, infinitely many Pt-sequences are generated, whose terms grow exponentially, and there are infinitely many empty spaces of arbitrary size, it follows that at each level z, infinitely many of these empty spaces will be subdivided, at least into two parts, giving rise to infinitely many new empty spaces.

Therefore, without taking into account the other sets, we conclude, under the hypothesis of the proof of Theorem 3.5, that the creation of empty spaces is an endless process, ensuring that the branching structure of the tree keeps growing indefinitely. $\square$

Remark 6.1. *If we were to suppose that the size of the empty spaces becomes bounded, then the Pt-sequences would cease to move further apart from each other. As a result, there would exist $K \in \mathbb{N}$ such that for all $k \geq K, \beta^{(\mathrm{a})k} \leq 1$, which would contradict the hypothesis itself and render Theorem 3.5 trivially true.*

6.6 Interdependencies between proofs of mathematical statements

The figure bellow shows the interdependencies between proofs of lemmas, theorems and propositions.

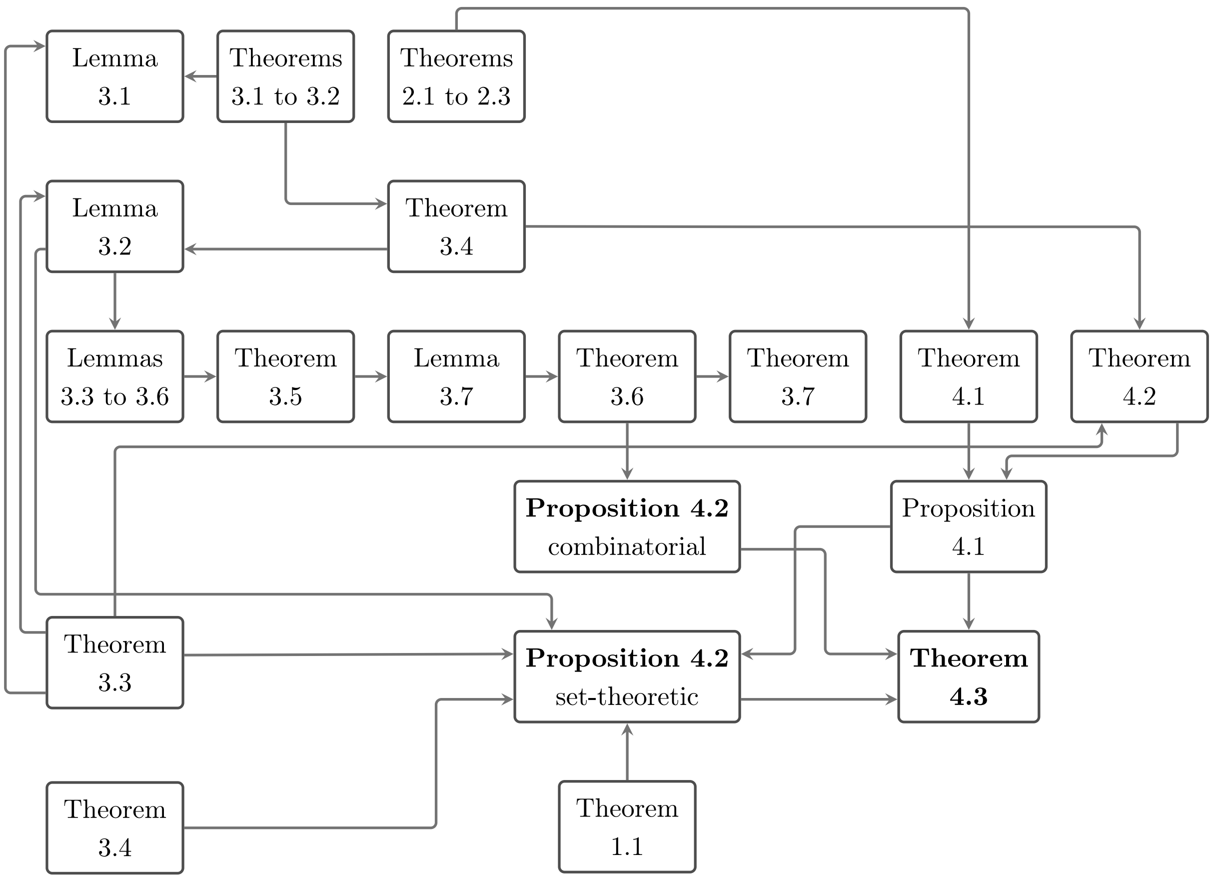

Figure 7. Interdependencies between proofs of mathematical statements

7 Glossary

$\mathcal{N}$: $\mathbb{N} \cup \{-1\}$.

$\{v_l^p\}_{l\in\mathcal{N}}$: R-Cz sequence whose first term is $v_{-1}^p = p$.

α_i: 2 raised to the power of α_i, such that $v_i^p = \dfrac{3v_{i-1}^p + 1}{2^{\alpha_i}}$ is odd.

$R_{Cz}, R_{Cz}^{\mathrm{c}}, R_{Cz}^{\mathrm{d}}$: sets containing all R-Cz sequences, all convergent R-Cz sequences, and all divergent R-Cz sequences, respectively.

C: set of the terms of all convergent R-Cz sequences.

D: complement of C in $2\mathbb{N}+1$.

$\{b_i\}_{i\geq 0}$: sequence of the penultimate terms, ordered in ascending order.

$\{s_{j,i}\}_{(j,i)\in I_{\mathrm{Pt}}\times\mathbb{N}}$: family of the Pt-sequences corresponding to all convergent R-Cz sequences. Each sequence contains the preceding terms of the parent term q_j.

β^{q_j}: shift factor of the parent term q_j.

$\beta^{(\mathrm{a})k}$: average shift factor of the first $k+1$ Pt-sequences.

R_i^D: interval $[a_i, b_i]$ in D.

R^D: set of R_i^D intervals.

$d^{(\mathrm{a})k}$: average distance between two consecutive Pt-sequences, among the first $k+1$ Pt-sequences (see Definition 3.6).

$\{d_l^p\}_{l\in\mathbb{N}}$: divergent R-Cz sequence.

$\{d_{i_k}^p\}_{k\in\mathbb{N}}$: strictly increasing subsequence of the divergent R-Cz sequence $\{d_l^p\}$.

T_{es}: tree of the empty spaces of the proof of Theorem 3.5.

$T^{d_k^p}$: tree associated with the term d_k^p of the divergent R-Cz sequence $\{d_l^p\}$.

T^{d}: tree of the divergent R-Cz sequences.

8 Declarations

As the sole author of this paper, I declare that I have no conflict of interest and have received no support from any organization for the submitted work.

References

[1] Tom M. Apostol, *Introduction to analytic number theory*, Springer, 1976

[2] Manon Bischoff, *The simplest math problem could be unsolvable*, Scientific American, 12 March 2024.

[3] Lothar Collatz,*On the motivation and origin of the* $3n+1$ *problem*, J. Chinese Math. Soc., 2, 1–25, 1963

[4] Richard Crandall, *On the* $3x+1$ *problem*, Math of computation, vol. 32, no. 144, pp. 1281-1291, october 1978

[5] Ilia Krasikov & Jeffrey C. Lagarias, *Bounds for the 3x + 1 problem using difference inequalities*, Atca Arithmetica, 30 april 2022

[6] Jeffrey C. Lagarias, *The* $3x+1$ *problem: an overview*, American Mathematical Society, Providence, RI, 2010.

[7] Terence Tao, *Almost all orbits of the Collatz map attain almost bounded values*, arXiv:1909.03562v2 [math.PR],13 September 2019.