Collatz Map Basics

Ralf Stephan AI Modelz

March 31, 2026

Abstract

This is a collection of autoformalizations of definitions and elementary lemmas associated with the Collatz Conjecture. The focus is first on facts needed to build up recent arguments by Rozier and Terracol[5]. Secondly, we attempt to be more complete regarding elementary lemmas to allow shorter proofs of more advanced theorems. The work is motivated by an urge to understand current research better, but also by the idea of a specialized library of autoformalizations, should the need arise for it.

Chapter 1

The Collatz Conjecture

1.1 Elementary Number Theory

Lemma 1.

$$2^k \pmod 3 = \begin{cases} 1 & \textit{if } k \textit{ is even,} \\ 2 & \textit{if } k \textit{ is odd.} \end{cases} \tag{1.1}$$

Proof. Trivial by induction on k. □

Lemma 2. *For any even $x \in \mathbb{N}$, the odd part of $x/2$ is equal to the odd part of x.*

Proof. Since x is even, $x = 2 \cdot (x/2)$. The dyadic valuation of $2 \cdot (x/2)$ is one more than that of $x/2$ (if $x/2 > 0$), but the odd part remains the same. □

Lemma 3. *For any $k, x \in \mathbb{N}$, the odd part of $2^k \cdot x$ is equal to the odd part of x.*

Proof. This is a standard property of the odd part (dyadic valuation). Multiplying by a power of 2 only changes the valuation, not the odd part. □

Lemma 4. *For any $k \in \mathbb{N}$ and any odd $m \in \mathbb{N}$, the odd part of $2^k \cdot m$ is m.*

Proof. By Lemma 3, the odd part of $2^k \cdot m$ is the odd part of m. Since m is odd, it is its own odd part. □

Lemma 5. *Let y be an odd natural number such that $y \pmod 3 \neq 0$. Then there exist natural numbers x and k such that x is odd and*

$$3x + 1 = 2^k y. \tag{1.2}$$

Proof. Since $y \pmod 3 \neq 0$, we have either $y \equiv 1 \pmod 3$ or $y \equiv 2 \pmod 3$. We consider these two cases:

If $y \equiv 1 \pmod 3$, we choose $x = (4y - 1)/3$ and $k = 2$. If $y \equiv 2 \pmod 3$, we choose $x = (2y - 1)/3$ and $k = 1$.

In both cases, it is straightforward to verify that x is an odd natural number and $3x + 1 = 2^k y$. □

Lemma 6. *Let y be a natural number such that $y \equiv 0 \pmod 3$. Then for all natural numbers n and k,*

$$3n + 1 \neq 2^k y. \tag{1.3}$$

Proof. Suppose for contradiction that $3n + 1 = 2^k y$ for some natural numbers n and k. Taking both sides modulo 3, the left-hand side reduces to $1 \pmod 3$. Since $y \equiv 0 \pmod 3$, the right-hand side is a multiple of 3, which reduces to $0 \pmod 3$. This yields $1 \equiv 0 \pmod 3$, a contradiction. □

Lemma 7. *Let $y > 1$ be a natural number such that $y \equiv 1 \pmod 6$. Then there exists an odd natural number $x > 1$ such that*

$$3x + 1 = 4y. \tag{1.4}$$

Proof. We choose $x = (4y - 1)/3$. Since $y \equiv 1 \pmod 6$, it follows that $4y - 1$ is an odd multiple of 3, so x is an odd integer. Furthermore, since $y > 1$, we have $x > 1$. It is straightforward to verify that $3x + 1 = 4y = 2^2 y$. □

Lemma 8. *Let y be a natural number such that $y \equiv 5 \pmod 6$. Then there exists an odd natural number $x > 1$ such that*

$$3x + 1 = 2y. \tag{1.5}$$

Proof. We choose $x = (2y - 1)/3$. Since $y \equiv 5 \pmod 6$, it follows that $2y - 1$ is an odd multiple of 3, so x is an odd integer. Furthermore, since $y \equiv 5 \pmod 6$, we have $y \geq 5$, which implies $x \geq 3 > 1$. It is straightforward to verify that $3x + 1 = 2y$. □

Lemma 9. *For any $k, m, n \in \mathbb{N}$, if $m \equiv n \pmod{2^{k+1}}$, then $m \equiv n \pmod 2$.*

Proof. This follows from the fact that 2 divides 2^{k+1}. □

Lemma 10. *For any $a, b, c \in \mathbb{N}$, if $a \equiv b \pmod c$, then $c \mid (a - b)$ in the integers.*

Proof. Standard property of modular arithmetic. □

Lemma 11. *For any $a, b, c \in \mathbb{N}$, if $c \mid (a - b)$ in the integers, then $a \equiv b \pmod c$ in the natural numbers.*

Proof. Standard property of modular arithmetic. □

Lemma 12. *For any $s, k \in \mathbb{N}$, 3^s and 2^k are coprime.*

Proof. The only prime factor of 3^s is 3, and the only prime factor of 2^k is 2. Since 2 and 3 are distinct primes, any power of 3 and any power of 2 are coprime. □

Definition 13. We define the indicator function $X : \mathbb{N} \to \mathbb{N}$ such that $X(n) = 0$ if n is even and $X(n) = 1$ if n is odd. Formally,

$$X(n) = \frac{1 - (-1)^n}{2}. \tag{1.6}$$

Lemma 14. *If $n \equiv 0 \pmod 2$, then $X(n) = 0$.*

Proof. If n is even, $(-1)^n = 1$, so $X(n) = (1 - 1)/2 = 0$. □

Lemma 15. *If $n \equiv 1 \pmod 2$, then $X(n) = 1$.*

Proof. If n is odd, $(-1)^n = -1$, so $X(n) = (1 - (-1))/2 = 1$. □

Lemma 16. *For all $n \in \mathbb{N}$, $X(n) = n \pmod 2$.*

Proof. By cases on the parity of n. □

Lemma 17. *If $m \equiv n \pmod 2$, then $X(m) = X(n)$.*

Proof. This follows from Lemma 16. □

1.2 The Collatz Map

Definition 18. The Collatz map $C : \mathbb{N} \to \mathbb{N}$ is defined by

$$C(n) = \begin{cases} n/2 & \text{if } n \text{ is even,} \\ 3n+1 & \text{if } n \text{ is odd.} \end{cases} \tag{1.7}$$

Lagarias reviewed the history of the map and the conjecture in [4].

Definition 19. For any $k \in \mathbb{N}$, the k-fold iteration of the Collatz map $C^k : \mathbb{N} \to \mathbb{N}$ is defined recursively by

$$C^0(n) = n \tag{1.8}$$

and

$$C^{k+1}(n) = C^k(C(n)). \tag{1.9}$$

Lemma 20. *For all $k, n \in \mathbb{N}$, we have $C^{k+1}(n) = C(C^k(n))$.*

Proof. We proceed by induction on $k \in \mathbb{N}$. For the base case $k = 0$, $C^1(n) = C(C^0(n)) = C(n)$ by definition. For the inductive step $k+1$, we have:

$$\begin{aligned} C^{k+2}(n) &= C^{k+1}(C(n)) && \text{(by definition)} \\ &= C(C^k(C(n))) && \text{(by induction hypothesis applied to } C(n)) \\ &= C(C^{k+1}(n)) && \text{(by definition.)} \end{aligned}$$

□

Lemma 21. $C(0) = 0$.

Proof. Trivial by definition. □

Lemma 22. *For all $k \in \mathbb{N}$, $C^k(0) = 0$.*

Proof. By induction on k. □

Lemma 23. *If $n \geq 1$, then $C(n) \geq 1$.*

Proof. If n is even, then $n \geq 2$, so $n/2 \geq 1$. If n is odd, then $3n+1 \geq 4 \geq 1$. □

Lemma 24. *If $n \geq 1$, then for all $k \in \mathbb{N}$, $C^k(n) \geq 1$.*

Proof. By induction on k, using Lemma 23. □

Lemma 25. *For all $k \geq 1$, $C(2^k) = 2^{k-1}$.*

Proof. Since 2^k is even for $k \geq 1$, $C(2^k) = 2^k/2 = 2^{k-1}$. □

Lemma 26. *For any $k, m \in \mathbb{N}$, $C^k(2^k \cdot m) = m$.*

Proof. We proceed by induction on $k \in \mathbb{N}$. The base case $k = 0$ is trivial as $C^0(m) = m$. For the inductive step $k+1$, we have

$$\begin{aligned} C^{k+1}(2^{k+1} \cdot m) &= C^k(C(2 \cdot 2^k \cdot m)) \\ &= C^k(2^k \cdot m) && \text{(since } 2 \cdot 2^k \cdot m \text{ is even)} \\ &= m && \text{(by induction hypothesis.)} \end{aligned}$$

□

Lemma 27. *For all $l \in \mathbb{N}$, $C^l(2^l) = 1$.*

Proof. By induction on l, using Lemma 25. □

Lemma 28. *For all $k < l$, $C^k(2^l) \neq 1$.*

Proof. By induction on k, using Lemma 25. □

Lemma 29. *For all $l \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that $C^l(n) = 1$ and for all $k < l$, $C^k(n) \neq 1$.*

Proof. Choose $n = 2^l$ and apply Lemmas 27 and 28. □

Lemma 30. *For all $a, b, n \in \mathbb{N}$, $C^{a+b}(n) = C^a(C^b(n))$.*

Proof. By induction on b. □

Lemma 31. *If $C^k(n) = n$, then for all $m \in \mathbb{N}$, $C^{m \cdot k}(n) = n$.*

Proof. By induction on m, using Lemma 30. □

Lemma 32. *If $n \in \{1, 2, 4\}$, then for all $i \in \mathbb{N}$, $C^i(n) \in \{1, 2, 4\}$.*

Proof. By induction on i, checking the cases $n = 1, 2, 4$ for $C(n)$. □

Lemma 33. *For all $i \in \mathbb{N}$, $C^i(1) \leq 4$.*

Proof. Since $1 \in \{1, 2, 4\}$, we have $C^i(1) \in \{1, 2, 4\}$ by Lemma 32, so $C^i(1) \leq 4$. □

Lemma 34. *The Collatz map $C(n)$ has the closed form*

$$C(n) = \frac{(7n + 2) - (5n + 2)(-1)^n}{4} \tag{1.10}$$

for all $n \in \mathbb{N}$.

Proof. We split into cases based on the parity of n. If n is even, then $(-1)^n = 1$, so

$$\frac{(7n + 2) - (5n + 2)(1)}{4} = \frac{2n}{4} = \frac{n}{2} = C(n). \tag{1.11}$$

If n is odd, then $(-1)^n = -1$, so

$$\frac{(7n + 2) - (5n + 2)(-1)}{4} = \frac{12n + 4}{4} = 3n + 1 = C(n). \tag{1.12}$$

In both cases, the formula yields $C(n)$. □

1.3 The Collatz Conjecture

Theorem 35. *For every natural number n, either $n = 0$ or there exists a natural number k such that $C^k(n) = 1$.*

Proof. No proof exists. □

Lemma 36. *If there exists a natural number $n > 4$ and $k \geq 1$ such that $C^k(n) = n$, then the Collatz conjecture is false* **(No-cycle condition)**.

Proof. Suppose for the sake of contradiction that the Collatz conjecture is true, which would imply every positive natural number eventually reaches 1. Since $n > 4 \geq 1$, there exists some $j \geq 0$ such that $C^j(n) = 1$. Because n is part of a cycle of length $k \geq 1$, we have $C^{j \cdot k}(n) = n$. Since $j \cdot k \geq j$, we can write $j \cdot k = (j \cdot k - j) + j$, allowing us to decompose the iteration: $n = C^{j \cdot k}(n) = C^{j \cdot k - j}(C^j(n)) = C^{j \cdot k - j}(1)$. However, it is a known property of the Collatz map that iterating starting from 1 only yields values in the set $\{1, 2, 4\}$. Thus, $C^{j \cdot k - j}(1) \leq 4$, which implies $n \leq 4$. This contradicts our initial assumption that $n > 4$, so the conjecture must be false. □

Lemma 37. *If there exists a natural number n such that its orbit under C is unbounded, then the Collatz conjecture is false* **(No-unbounded-orbit condition)**.

Proof. Suppose for the sake of contradiction that the Collatz conjecture is true, meaning every positive integer reaches 1. For our starting value $n \geq 1$, there must exist some index $j \geq 0$ such that $C^j(n) = 1$. Let M be the maximum value reached in the first j steps of the orbit, that is, $M = \max_{0 \leq i \leq j} C^i(n)$. Because the orbit of n is unbounded, there exists some index k such that $C^k(n) > M + 4$. If we had $k \leq j$, then $C^k(n) \leq M < M + 4$, which is a contradiction. Therefore, we must have $k > j$. We can then decompose the iteration as $C^k(n) = C^{k-j}(C^j(n)) = C^{k-j}(1)$. Since iterating the Collatz map from 1 only produces values in $\{1, 2, 4\}$, we have $C^{k-j}(1) \leq 4$. This implies $C^k(n) \leq 4$, which contradicts $C^k(n) > M + 4 \geq 4$. Thus, the assumption that the conjecture is true must be false. □

Lemma 38. *If no natural number $n > 4$ lies on a nontrivial cycle and every orbit under C is bounded, then the Collatz conjecture is true.*

$$(\forall nk, n > 4 \to k \geq 1 \to C^k(n) \neq n) \to (\forall n, \exists B, \forall k, C^k(n) \leq B) \to \forall m, m = 0 \lor \exists j, C^j(m) = 1$$

Proof. To prove that the Collatz conjecture is true under these assumptions, we must show that for an arbitrary target number m (distinct from the hypothesis parameter n), either $m = 0$ or there exists some index j such that $C^j(m) = 1$. If $m = 0$, this condition trivially holds. Assume $m \geq 1$. By the boundedness assumption, there exists some bound B such that $C^k(m) \leq B$ for all $k \geq 0$. Because the infinite sequence $(C^k(m))_{k=0}^{\infty}$ only takes values in the finite set $\{1, 2, \ldots, B\}$, the Pigeonhole Principle implies that there must be a collision. That is, there exist indices $i < j$ such that $C^i(m) = C^j(m)$. Let $c = C^i(m)$. Then $C^{j-i}(c) = C^j(m) = C^i(m) = c$, meaning c is part of a cycle of length $j - i \geq 1$. By our no-cycle assumption, no number strictly greater than 4 can be part of a cycle, so we must have $c \leq 4$. Since $m \geq 1$, all of its iterates are strictly positive, giving $c \in \{1, 2, 3, 4\}$. We can verify that every number in this set eventually reaches 1: $C^0(1) = 1$, $C^1(2) = 1$, $C^7(3) = 1$ (since $3 \mapsto 10 \mapsto 5 \mapsto 16 \mapsto 8 \mapsto 4 \mapsto 2 \mapsto 1$), and $C^2(4) = 1$. In all cases, there exists some j' such that $C^{j'}(c) = 1$. Therefore, starting from m, after i steps we reach c, and after j' more steps we reach 1, yielding $C^{i+j'}(m) = 1$. Thus, the Collatz conjecture is true. □

1.4 The Collatz Graph

Definition 39. The Collatz graph $\mathcal{G}$ is a directed graph on $\mathbb{N}$ with an edge from n to $T(n)$ for all $n \in \mathbb{N}$. We consider its associated simple undirected graph $\mathcal{G}'$.

Lemma 40. *Two distinct vertices a, b are adjacent in $\mathcal{G}'$ if and only if $b = T(a)$ or $a = T(b)$.*

Proof. Immediate from the definition of the simple graph associated with a directed graph. □

Lemma 41. *For any $k, n \in \mathbb{N}$, $T^k(n)$ is reachable from n in $\mathcal{G}'$.*

Proof. By induction on k, following the edges $(T^i(n), T^{i+1}(n))$. □

Lemma 42. *If $T^i(a) = T^j(b)$ and b is adjacent to c in $\mathcal{G}'$, then there exist i', j' such that $T^{i'}(a) = T^{j'}(c)$.*

Proof. If $c = T(b)$, then $T^i(a) = T^j(b)$, so $T^{i+1}(a) = T(T^i(a)) = T(T^j(b)) = T^{j+1}(b) = T^j(c)$. If $b = T(c)$, then $T^i(a) = T^j(b) = T^j(T(c)) = T^{j+1}(c)$. □

Lemma 43. *If a and b are reachable in $\mathcal{G}'$, then they have a common descendant under T, i.e., there exist i, j such that $T^i(a) = T^j(b)$.*

Proof. By induction on the path length between a and b, using Lemma 42. □

Theorem 44. *The Collatz graph $\mathcal{G}$ restricted to positive integers is weakly connected (i.e., any two $a, b \geq 1$ are reachable in $\mathcal{G}'$) if and only if the Collatz conjecture holds.*

Proof. ($\Rightarrow$) Assume connectivity. For any $n \geq 1$, n is reachable from 1. By Lemma 43, there exist i, j such that $T^i(n) = T^j(1)$. Since 1 is in a cycle $\{1, 2\}$, $T^j(1) \in \{1, 2\}$, so $T^i(n)$ eventually hits 1. By Lemma 115, n eventually hits 1 under C. ($\Leftarrow$) Assume the Collatz conjecture. For any $n \geq 1$, n eventually hits 1 under C, and by Lemma 116, it also hits 1 under T. By Lemma 41, n is reachable from 1 in $\mathcal{G}'$. Thus any $a, b \geq 1$ are both reachable from 1, and hence reachable from each other. □

Chapter 2

The Reduced Collatz Map

2.1 Definitions

Definition 45. The reduced Collatz map $R : \mathbb{N} \to \mathbb{N}$ is defined by

$$R(n) = \frac{3n+1}{2^{v_2(3n+1)}} \tag{2.1}$$

where $v_2(m)$ is the dyadic valuation of m. In short $R(n) =$ odd part of $3n+1$.

This map seems to have been introduced by Crandall[1]. One of its advantages is that multiplications by 3 are constant for each step, regardless if n is even or odd.

Definition 46. For any $k \in \mathbb{N}$, the k-fold iteration of the reduced Collatz map $R^k : \mathbb{N} \to \mathbb{N}$ is defined recursively by

$$R^0(n) = n \tag{2.2}$$

and

$$R^{k+1}(n) = R^k(R(n)). \tag{2.3}$$

Lemma 47. $R(0) = 1$.

Proof. By definition, $R(0) =$ odd part of $(3 \cdot 0 + 1) =$ odd part of $1 = 1$. □

Lemma 48. $R(1) = 1$.

Proof. By definition, $R(1) =$ odd part of $(3 \cdot 1 + 1) =$ odd part of $4 = 1$. □

Lemma 49. *For all $n \in \mathbb{N}$, $R(n) \geq 1$.*

Proof. Since $3n+1 \geq 1$ for all n, the odd part of $3n+1$ is a positive divisor of a positive number, hence $R(n) \geq 1$. □

Lemma 50. *For all $n, i \in \mathbb{N}$ with $n \geq 1$, we have $R^i(n) \geq 1$.*

Proof. We proceed by induction on i. The base case $i = 0$ is immediate since $R^0(n) = n \geq 1$. For the inductive step, $R^{i+1}(n) = R^i(R(n))$. By Lemma 49, $R(n) \geq 1$, so the induction hypothesis gives $R^i(R(n)) \geq 1$. □

Lemma 51. *For all $n \in \mathbb{N}$, $R(n)$ is odd, i.e. $R(n) \equiv 1 \pmod 2$.*

Proof. By definition, $R(n)$ is the odd part of $3n + 1$. $\square$

Lemma 52. *For all $n \in \mathbb{N}$ and $i \geq 1$, $R^i(n)$ is odd, i.e. $R^i(n) \equiv 1 \pmod 2$.*

Proof. We proceed by induction on i. The base case $i = 1$ follows directly from Lemma 51, since $R^1(n) = R(n)$ is odd. For the inductive step, $R^{i+1}(n) = R^i(R(n))$. If $i \geq 1$, the induction hypothesis gives that $R^i(R(n))$ is odd. If $i = 0$, then $R^{i+1}(n) = R^1(n) = R(n)$, which is odd by Lemma 51. $\square$

Lemma 53. *If $m \equiv 0 \pmod 3$, then there is no $n \in \mathbb{N}$ such that $R(n) = m$. In other words, multiples of 3 are not in the image of R.*

Proof. Suppose for contradiction that $R(n) = m$ for some n, with $3 \mid m$. Since $R(n)$ is the odd part of $3n + 1$, we have $R(n) \mid 3n + 1$. $\square$

Lemma 54. *Let $m > 0$ be odd with $m \not\equiv 0 \pmod 3$. Then there exists an odd $n > 1$ such that $R(n) = m$.*

Proof. If $m = 1$, take $n = 5$; one checks $R(5) = 1$ by computation.

If $m > 1$, we split on the residue of m modulo 3:

- If $m \equiv 1 \pmod 3$: set $n = (4m-1)/3$. Then $3 \mid (4m-1)$ and $3n+1 = 4m = 2^2 \cdot m$. Since m is odd, the odd part of $2^2 \cdot m$ is m, so $R(n) = m$. One verifies n is odd and $n > 1$.

- If $m \equiv 2 \pmod 3$: set $n = (2m-1)/3$. Then $3 \mid (2m-1)$ and $3n+1 = 2m = 2^1 \cdot m$. Since m is odd, the odd part of $2m$ is m, so $R(n) = m$. One verifies n is odd and $n > 1$.

$\square$

Lemma 55. *For any odd $n > 0$, $R(n) > n$ if and only if $n \equiv 3 \pmod 4$.*

Proof. ($\Rightarrow$) Suppose $n \equiv 1 \pmod 4$, say $n = 4k + 1$. Then $3n + 1 = 12k + 4 = 2^2(3k + 1)$. The reduced step is $R(n) =$ odd part of $(3k+1) \leq 3k+1$. Since $3k+1 < 4k+1$ (for $k \geq 0$), we have $R(n) < n$, a contradiction. Thus we must have $n \equiv 3 \pmod 4$.

($\Leftarrow$) Suppose $n \equiv 3 \pmod 4$, say $n = 4k + 3$. Then $3n + 1 = 12k + 10 = 2(6k + 5)$. Since $6k + 5$ is odd, $R(n) = 6k + 5$. Since $k \geq 0$, we have $6k + 5 > 4k + 3$, so $R(n) > n$. $\square$

Lemma 56. *If $n > 0$ is odd and $R(n) = m$, then there exists $i \geq 1$ such that $C^i(n) = m$.*

Proof. By definition, $R(n)$ is the odd part of $3n+1$, so $3n+1 = 2^k \cdot m$ for some $k \in \mathbb{N}$. Since n is odd, $3n+1$ is even, so $k \geq 1$. One iteration of the Collatz map on n gives $C^1(n) = 3n+1 = 2^k \cdot m$. By Lemma 26, applying the Collatz map k more times to $2^k \cdot m$ yields $C^k(2^k \cdot m) = m$. Thus $C^{k+1}(n) = C^k(C(n)) = C^k(2^k \cdot m) = m$. $\square$

Lemma 57. *If n is odd and $R^k(n) = m$, then there exists $i \geq k$ such that $C^i(n) = m$.*

Proof. We proceed by induction on k. The base case $k = 0$ is trivial as $R^0(n) = n = C^0(n)$. For the inductive step $k + 1$, let $n_1 = R(n)$. By the induction hypothesis, since $R^k(n_1) = m$, there exists $i \geq k$ such that $C^i(n_1) = m$. By Lemma 56, there exists $j \geq 1$ such that $C^j(n) = n_1$. Then $C^{i+j}(n) = C^i(C^j(n)) = C^i(n_1) = m$. Since $i \geq k$ and $j \geq 1$, we have $i + j \geq k + 1$. $\square$

Definition 58. For $d, n \in \mathbb{N}$, the number of odd steps in the first d iterations of the standard Collatz map starting from n is defined by

$$\text{countOdds}(d, n) = \sum_{j=0}^{d-1} [C^j(n) \text{ is odd}] \tag{2.4}$$

where $[P]$ is the Iverson bracket.

Lemma 59. *Let $n \in \mathbb{N}$ be odd. For any $d \in \mathbb{N}$, the odd part of $C^d(n)$ is equal to $R^k(n)$, where $k = \text{countOdds}(d, n)$. That is,*

$$\textit{odd part of } C^d(n) = R^k(n). \tag{2.5}$$

Proof. We proceed by induction on d. The base case $d = 0$ is trivial as $C^0(n) = n$ is odd, its odd part is n, and count_odds$(0, n) = 0$, with $R^0(n) = n$.

For the inductive step $d+1$, let $C^d(n) = m$. By induction, the odd part of m is $R^{\text{countOdds}(d,n)}(n)$. If m is even, then $C^{d+1}(n) = m/2$. The odd part of $m/2$ is the same as the odd part of m. Since m is even, $\text{countOdds}(d + 1, n) = \text{countOdds}(d, n)$, and the equality holds. If m is odd, then $C^{d+1}(n) = 3m + 1$. Since m is odd, it equals its own odd part, $m = R^{\text{countOdds}(d,n)}(n)$. The odd part of $3m + 1$ is $R(m) = R(R^{\text{countOdds}(d,n)}(n)) = R^{\text{countOdds}(d,n)+1}(n)$. Since m is odd, $\text{countOdds}(d + 1, n) = \text{countOdds}(d, n) + 1$, confirming the result. □

Lemma 60. *Let $n \in \mathbb{N}$ be odd. There exists a cycle in the standard Collatz map starting from n if and only if there exists a cycle in the reduced Collatz map starting from n.*

Proof. ($\Rightarrow$) Suppose there exists $k > 0$ such that $C^k(n) = n$. By Lemma 59, the odd part of $C^k(n)$ is $R^j(n)$ where $j = \text{countOdds}(k, n)$. Since n is odd, its odd part is n, so $R^j(n) = n$. Since n is odd and $k > 0$, at least one standard Collatz step must have occurred, and since $n \geq 1$, we must have $j > 0$. Thus n is part of a reduced cycle.

($\Leftarrow$) Suppose there exists $j > 0$ such that $R^j(n) = n$. By Lemma 57, there exists $k \geq j$ such that $C^k(n) = n$. Since $j > 0$, we have $k > 0$, so n is part of a standard cycle. □

Lemma 61. *Let $n \in \mathbb{N}$ be odd. The orbit of n under the standard Collatz map is bounded if and only if its orbit under the reduced Collatz map is bounded.*

Proof. ($\Rightarrow$) Suppose there exists B such that $C^k(n) \leq B$ for all $k \geq 0$. By Lemma 57, every reduced iterate $R^j(n)$ is equal to some $C^k(n)$, so $R^j(n) \leq B$ for all $j \geq 0$.

($\Leftarrow$) Suppose there exists B such that $R^j(n) \leq B$ for all $j \geq 0$. We claim that $C^k(n) \leq \max(n, 3B + 1)$ for all $k \geq 0$. We proceed by induction on k. The base case $k = 0$ is $C^0(n) = n$, which is clearly bounded. For the inductive step, let $C^k(n) = m$. If m is even, $C^{k+1}(n) = m/2 \leq m$, which is bounded by the induction hypothesis. If m is odd, then by Lemma 59, m is the odd part of $C^k(n)$, which is $R^j(n)$ for $j = \text{countOdds}(k, n)$. Thus $m \leq B$. Then $C^{k+1}(n) = 3m + 1 \leq 3B + 1$, which is within the bound. □

2.2 Primitives of The Reduced Map

Lemma 62. *For all $n > 1$, $R(4n + 1) = R(n)$.*

Proof. We compute $3(4n+1)+1 = 12n+4 = 4(3n+1) = 2^2 \cdot (3n+1)$. Since multiplying by 2^2 does not change the odd part, we have $R(4n + 1) = \text{odd part of } 2^2(3n + 1) = \text{odd part of } (3n + 1) = R(n)$. □

Lemma 63. *For all $n > 1$ and $i > 0$, $R^i(4n+1) = R^i(n)$.*

Proof. For $i = k+1$, we have $R^{k+1}(4n+1) = R^k(R(4n+1)) = R^k(R(n)) = R^{k+1}(n)$, where the middle equality follows from Lemma 62. □

Lemma 64. *For any odd $n \in \mathbb{N}$, $2R(n) \leq 3n+1$.*

Proof. By definition, $R(n)$ is the odd part of $3n+1$, meaning $3n+1 = 2^k \cdot R(n)$ for some $k \in \mathbb{N}$. Since n is odd, $3n+1$ is even, which implies $k \geq 1$. Thus $2^k \geq 2$, and we have $3n+1 \geq 2R(n)$. □

Definition 65. An odd number $n > 1$ is called a *primitive* at level $i > 1$ if

1. $R^i(n) = 1$,
2. there is no odd k with $4k+1 = n$ and $R^i(k) = 1$.

Lemma 66. *Let $n > 1$ be odd with $n \neq 5$, let $i > 1$, and suppose $R^i(n) = 1$. Then n is a primitive at level i if and only if $n \not\equiv 5 \pmod 8$.*

Proof. ($\Rightarrow$) Suppose n is primitive and $n \equiv 5 \pmod 8$. Write $n = 8k+5 = 4(2k+1)+1$. Then $2k+1$ is odd and $2k+1 > 1$ (since $n \neq 5$ implies $k \geq 1$). By Lemma 63, $R^i(2k+1) = R^i(4(2k+1)+1) = R^i(n) = 1$, so $2k+1$ is an odd predecessor of n via the $4k+1$ map with the same step count, contradicting primitivity.

($\Leftarrow$) Suppose $n \not\equiv 5 \pmod 8$ and there exists an odd k with $4k+1 = n$ and $R^i(k) = 1$. Since k is odd, write $k = 2m+1$; then $n = 4(2m+1)+1 = 8m+5 \equiv 5 \pmod 8$, a contradiction. □

Lemma 67. *Let $x > 1$ be odd with $R(x) \neq 1$. Then there exists an odd $n > 1$ with $R(n) = R(x)$ and $n \not\equiv 5 \pmod 8$.*

Proof. We proceed by strong induction on x. If $x \not\equiv 5 \pmod 8$, take $n = x$.

If $x \equiv 5 \pmod 8$, write $x = 8k+5 = 4(2k+1)+1$. We must have $2k+1 > 1$, since $k = 0$ gives $x = 5$ and $R(5) = 1$, contradicting $R(x) \neq 1$. By Lemma 62, $R(x) = R(2k+1)$, so $R(2k+1) \neq 1$. Since $2k+1 < x$, the induction hypothesis yields an odd $n > 1$ with $R(n) = R(2k+1) = R(x)$ and $n \not\equiv 5 \pmod 8$. □

Lemma 68. *Let $y > 1$ be odd with $y \not\equiv 0 \pmod 3$ and $R^i(y) = 1$. Then there exists a primitive n at level $i+1$.*

Proof. Since y is odd, positive, and not divisible by 3, there exists an odd $x > 1$ with $R(x) = y$ (by the surjectivity of R onto odd numbers not divisible by 3). In particular $R(x) \neq 1$ since $y > 1$.

By Lemma 67, there exists an odd $n > 1$ with $R(n) = R(x) = y$ and $n \not\equiv 5 \pmod 8$. Then $R^{i+1}(n) = R^i(R(n)) = R^i(y) = 1$. Since $n \neq 5$ (as $5 \equiv 5 \pmod 8$), Lemma 66 gives that n is a primitive at level $i+1$. □

Definition 69. The function reduce4x1 : $\mathbb{N} \to \mathbb{N}$ is defined recursively by

$$\text{reduce4x1}(n) = \begin{cases} \text{reduce4x1}((n-1)/4) & \text{if } n \equiv 5 \pmod 8, \\ n & \text{otherwise.} \end{cases}$$

Lemma 70. *For any $n \in \mathbb{N}$, $R(\text{reduce4x1}(n)) = R(n)$.*

Proof. We proceed by strong induction on n. If $n \not\equiv 5 \pmod 8$, then $\mathrm{reduce4x1}(n) = n$ and the result is trivial. If $n \equiv 5 \pmod 8$, then $n = 4k + 1$ for some odd $k = (n-1)/4$. By the definition of reduce4x1, $\mathrm{reduce4x1}(n) = \mathrm{reduce4x1}(k)$. By induction, $R(\mathrm{reduce4x1}(k)) = R(k)$. By Lemma 62, $R(n) = R(4k+1) = R(k)$. Thus $R(\mathrm{reduce4x1}(n)) = R(n)$. □

Lemma 71. *Let $x_1, x_2 \in \mathbb{N}$ and y_1, y_2 be their respective images under R, with $y_1 \neq y_2$. Then reduce4x1$(x_1) \neq$ reduce4x1(x_2).*

Proof. Suppose for contradiction that $\mathrm{reduce4x1}(x_1) = \mathrm{reduce4x1}(x_2)$. Then by Lemma 70, $y_1 = R(x_1) = R(\mathrm{reduce4x1}(x_1)) = R(\mathrm{reduce4x1}(x_2)) = R(x_2) = y_2$, contradicting $y_1 \neq y_2$. □

Lemma 72. *Given a level $m \geq 1$, a seed $y_0 > 1$ that is odd and satisfies $R^m(y_0) = 1$, and $B \in \mathbb{N}$, there exists an odd $y > B$ with $y > 1$, $y \not\equiv 0 \pmod 3$, and $R^m(y) = 1$.*

Proof. We first show by induction on B that there exists an odd $y > B$ with $y > 1$ and $R^m(y) = 1$. The base case $B = 0$ is given by y_0. For the inductive step, if $y > B$ satisfies the conditions, let $y' = 4y + 1$. Then $y' > 4B + 1 > B + 1$, y' is odd, $y' > 1$, and $R^m(y') = R^m(y) = 1$ by Lemma 63.

Now, fixed such a $y > B$ with $R^m(y) = 1$ and y odd, $y > 1$. If $y \not\equiv 0 \pmod 3$ we are done. Otherwise, if $y \equiv 0 \pmod 3$, consider $y' = 4y + 1$. Then $y' > y > B$, y' is odd, $y' > 1$, and $R^m(y') = 1$. Finally, $4y + 1 \equiv 4(0) + 1 \equiv 1 \pmod 3$, so $y' \not\equiv 0 \pmod 3$. □

Lemma 73. *For every level $m \geq 2$, there are infinitely many primitive numbers at level m.*

Proof. We proceed by induction on $m \geq 2$. Base case $m = 2$: Let $B \in \mathbb{N}$. We apply Lemma 72 starting from the seed $y_0 = 5$ at level 1 (with $R(5) = 1$) to find an odd $y > 2B + 2$ with $y \not\equiv 0 \pmod 3$ and $R(y) = 1$. By Lemma 68, there exists a primitive n at level 2 such that $R(n) = y$. Since $2R(n) \leq 3n + 1$ (by Lemma 64), $n \geq (2y-1)/3 > \frac{4B+3}{3} > B$. Thus $n > B$ is the required primitive.

Inductive step $m \to m + 1$: Suppose there is a primitive p at level m. By Lemma 72, there exists an odd $y > 2B + 2$ with $y > 1$, $y \not\equiv 0 \pmod 3$, and $R^m(y) = 1$. By Lemma 68, there is a primitive n at level $m+1$ with $R(n) = y$. As in the base case, the bound $2R(n) \leq 3n+1$ ensures $n > B$. □

Chapter 3

The Compact Map

3.1 Definitions

Definition 74. The compact Collatz map $T : \mathbb{N} \to \mathbb{N}$ is defined by

$$T(n) = \frac{n \cdot 3^{X(n)} + X(n)}{2} \tag{3.1}$$

where $X(n) = n \pmod 2$.

$T(n)$ was independently introduced by Terras (1976)[6] and Everett (in an equivalent form, 1977)[3]. It is also called the **Syracuse map**.

One of the advantages of using T instead of C or R is that the number of divisions by 2 per step is the same, regardless if n is even or odd. This leads to formulae for $T^i(n)$ that can be handled better, although they are still not closed in the usual sense.

Lemma 75. *If n is even, $T(n) = n/2$.*

Proof. If n is even, $X(n) = 0$, so $T(n) = (n \cdot 3^0 + 0)/2 = n/2$. □

Lemma 76. *If n is odd, $T(n) = (3n + 1)/2$.*

Proof. If n is odd, $X(n) = 1$, so $T(n) = (n \cdot 3^1 + 1)/2 = (3n + 1)/2$. □

Definition 77. For any $k \in \mathbb{N}$, the k-fold iteration of the map T is defined recursively by $T^0(n) = n$ and $T^{k+1}(n) = T(T^k(n))$.

Lemma 78. *If $n \geq 1$, then $T(n) \geq 1$.*

Proof. If n is even, $n \geq 2$, so $n/2 \geq 1$. If n is odd, $3n + 1 \geq 4$, so $(3n + 1)/2 \geq 2$. □

Lemma 79. *If $n \geq 1$, then $T^k(n) \geq 1$ for all $k \in \mathbb{N}$.*

Proof. By induction on k, using Lemma 78. □

Lemma 80. *For any $k, m, n \in \mathbb{N}$, if $m \equiv n \pmod{2^{k+1}}$, then $T(m) \equiv T(n) \pmod{2^k}$.*

Proof. Let $m \equiv n \pmod{2^{k+1}}$. This implies $m \equiv n \pmod 2$, so $X(m) = X(n)$. We have $2T(m) = 3^{X(m)}m + X(m)$ and $2T(n) = 3^{X(n)}n + X(n)$. Subtracting these yields $2(T(m) - T(n)) = 3^{X(m)}(m - n)$. Since $2^{k+1} \mid (m - n)$, we have $2^k \mid (T(m) - T(n))$. □

Definition 81. The number of odd steps in the first k iterations of T starting from n is defined by

$$Q(k,n) = \sum_{i=0}^{k-1} X(T^i(n)). \tag{3.2}$$

Definition 82. The stopping time $\sigma(n)$ of n under T is the smallest $k \geq 1$ such that $T^k(n) < n$. If no such k exists, $\sigma(n) = \infty$.

Definition 83. The total stopping time $\sigma_\infty(n)$ of n under T is the smallest $k \geq 1$ such that $T^k(n) = 1$. If no such k exists, $\sigma_\infty(n) = \infty$.

3.2 Parity Vectors

Definition 84. A parity vector is a finite sequence of bits (0 and 1). In the formalization, it is represented as a list of boolean values, where `false` corresponds to 0 and `true` corresponds to 1.

Definition 85. For a parity vector v, $q(v)$ is the number of ones (true entries) in v.

Definition 86. The size of a parity vector v, denoted $|v|$, is its length.

Definition 87. The ones-ratio of a parity vector v is the rational number $q(v)/|v|$. If the vector is empty, the ratio is defined to be 0.

Definition 88. For $j, n \in \mathbb{N}$, the parity vector $V(j,n)$ of length j for the compact Collatz sequence starting at n is defined as

$$V(j,n) = (X(n), X(T(n)), \dots, X(T^{j-1}(n))).$$

Lemma 89. *The length of $V(j,n)$ is j.*

Proof. Immediate from the definition. □

Lemma 90. *For any $i < j$, the i-th entry of $V(j,n)$ is $X(T^i(n))$.*

Proof. Immediate from the definition. □

Definition 91. The elementary precedence relation $\prec_e$ is defined by swapping a 01 subword with 10:

$$w_1 01 w_2 \prec_e w_1 10 w_2. \tag{3.3}$$

The precedence relation $\preceq$ is the reflexive transitive closure of $\prec_e$.

Definition 92. For $k, n \in \mathbb{N}$, the vector $E(k,n)$ is the function from $\{0, \dots, k-1\}$ to $\{0, 1\}$ defined by

$$E(k,n)(i) = X(T^i(n)). \tag{3.4}$$

Lemma 93. *Each entry of $E(k,n)$ is at most 1.*

Proof. Immediate from the definition of $X(n)$. □

Lemma 94. *The number of odd steps $Q(k,n)$ is equal to the sum of the entries of $E(k,n)$:*

$$Q(k,n) = \sum_{i=0}^{k-1} E(k,n)(i). \tag{3.5}$$

Proof. By definition of $Q(k, n)$ and $E(k, n)$. ☐

Lemma 95. *If $E(k + 1, m) = E(k + 1, n)$, then $E(k, m) = E(k, n)$.*

Proof. The agreement on the first $k + 1$ entries implies agreement on the first k entries. ☐

Lemma 96. *If $E(k, m) = E(k, n)$, then $Q(k, m) = Q(k, n)$.*

Proof. Directly from Lemma 94. ☐

3.3 Linear Decomposition

Definition 97. For $k, n \in \mathbb{N}$, the accumulated correction term $Q(k, n)$ of the Terras map is defined recursively by $Q(0, n) = 0$ and

$$Q(k + 1, n) = 3^{X(T^k(n))}Q(k, n) + 2^k X(T^k(n)). \tag{3.6}$$

Lemma 98. *If $j \leq k$, then $Q(j, n) \leq Q(k, n)$.*

Proof. By definition of $Q(k, n)$ as a sum of non-negative terms. ☐

Lemma 99. $Q(k + 1, n) = Q(k, n) + X(T^k(n))$.

Proof. Immediate from the definition of $Q(k, n)$. ☐

Lemma 100. *For any $k, n \in \mathbb{N}$, the k-fold iteration of the Terras map $T^k(n)$ satisfies the linear identity*

$$2^k T^k(n) = 3^{Q(k,n)}n + Q(k, n) \tag{3.7}$$

where $Q(k, n)$ is the number of odd iterates as defined in Definition 81.

Proof. By induction on k, using the identity $2T(m) = 3^{X(m)}m + X(m)$. ☐

This appeared first in Terras (1976)[6]. See Lemma 109 below for a variant of this formula.

Definition 101. The correction term $Q(k, n)$ has the closed-form expression

$$Q(k, n) = \sum_{j=0}^{k-1} X(T^j(n))2^j 3^{Q(k,n)-Q(j+1,n)}. \tag{3.8}$$

Lemma 102. *The recursive definition of $Q(k, n)$ is equivalent to the closed-form sum.*

Proof. By induction on k. ☐

Definition 103. The decomposition coefficient $C(k, n)$ is the rational number

$$C(k, n) = \frac{3^{Q(k,n)}}{2^k}. \tag{3.9}$$

Definition 104. The correction ratio $E(j, n)$ is the rational number

$$E(j, n) = \frac{Q(j, n)}{2^j}. \tag{3.10}$$

Lemma 105. $E(0, n) = 0$.

Proof. Trivial by definition. □

Lemma 106. *The ratio E satisfies the recurrence*

$$E(k+1,n) = \frac{3^{X(T^k(n))}}{2} E(k,n) + \frac{X(T^k(n))}{2}. \tag{3.11}$$

Proof. By diving the recurrence for $Q(k,n)$ by 2^{k+1}. □

Lemma 107. *For a fixed parity $x \in \{0,1\}$, the step map $a \mapsto \frac{3^x}{2}a + \frac{x}{2}$ is strictly monotone.*

Proof. Since $3^x/2 > 0$ for $x \in \{0,1\}$. □

Lemma 108. *If $E(k,m) = E(k,n)$, then $Q(k,m) = Q(k,n)$.*

Proof. By induction on k. □

Lemma 109. *Another form of the linear identity is given by:*

$$T^j(n) = C(j,n) \cdot n + E(j,n). \tag{3.12}$$

Proof. By substituting the definitions $C(j,n) = 3^{q(j,n)}/2^j$ and $E(j,n) = Q(j,n)/2^j$ into the linear identity $2^j T^j(n) = 3^{q(j,n)} n + Q(j,n)$. □

Definition 110. The coefficient stopping time $\tau(n)$ is the smallest $j \geq 1$ such that $C(j,n) < 1$. If no such j exists, $\tau(n) = \infty$.

3.4 Periodicity

Lemma 111. *For any $k \in \mathbb{N}$ and $m, n \in \mathbb{N}$, if $m \equiv n \pmod{2^k}$, then $E(k,m) = E(k,n)$.*

Proof. By induction on k. The base case $k = 0$ is trivial as $E(0,n)$ is the empty function. For the inductive step $k+1$, if $m \equiv n \pmod{2^{k+1}}$, then $m \equiv n \pmod 2$, so $X(m) = X(n)$. By the congruence properties of T (Lemma 80), $T(m) \equiv T(n) \pmod{2^k}$. By the induction hypothesis, $E(k, T(m)) = E(k, T(n))$, which implies agreement on the remaining k entries of the parity vector. □

Lemma 112. *For $m, n \geq 1$, if $E(k,m) = E(k,n)$, then $m \equiv n \pmod{2^k}$.*

Proof. Assume $E(k,m) = E(k,n)$. Then $Q(k,m) = Q(k,n) = S$ and $Q(k,m) = Q(k,n) = Q$. From the linear decomposition (Lemma 100), we have:

$$2^k T^k(m) = 3^S m + Q$$
$$2^k T^k(n) = 3^S n + Q$$

Subtracting these yields $2^k(T^k(m) - T^k(n)) = 3^S(m-n)$. Since 2^k is coprime to 3^S (Lemma 12), 2^k must divide $m-n$, so $m \equiv n \pmod{2^k}$. □

Theorem 113. *(Theorem 1.2 in Terras (1976)[6]) Two positive integers $m, n \geq 1$ have the same parity vector of length k if and only if they are congruent modulo 2^k:*

$$E(k,m) = E(k,n) \iff m \equiv n \pmod{2^k}. \tag{3.13}$$

Proof. This follows directly from Lemma 111 and Lemma 112. □

3.5 Correspondence to the Original Map

Lemma 114. *For any $n \in \mathbb{N}$, one step of the compact Collatz map T can be simulated by one or two steps of the original Collatz map C. Specifically, there exists $j \in \{1, 2\}$ such that $C^j(n) = T(n)$.*

Proof. We split into cases based on the parity of n. If n is even, $T(n) = n/2 = C(n)$, so $j = 1$. If n is odd, $T(n) = (3n + 1)/2 = C(3n + 1)/2 = C(C(n))$, so $j = 2$ (since $3n + 1$ is even). □

Lemma 115. *For any $k, n \in \mathbb{N}$, there exists $j \geq k \in \mathbb{N}$ such that $C^j(n) = T^k(n)$.*

Proof. By induction on k, using Lemma 114. □

Lemma 116. *If $n \geq 1$ and $C^j(n) = 1$ for some $j \in \mathbb{N}$, then there exists $k \in \mathbb{N}$ such that $T^k(n) = 1$.*

Proof. By induction on j. In the base case $j = 0$, $n = 1 = T^0(n)$. For the inductive step $j+1$, if n is even, then $C(n) = n/2 = T(n)$, so we apply the induction hypothesis to $T(n)$. If n is odd, then $C(n) = 3n + 1$, and $C(3n + 1) = (3n + 1)/2 = T(n)$. Since $C^j(3n + 1) = 1$, we apply the induction hypothesis to $3n + 1$, which reaches 1 under T. Since $T(n)$ is a step in that sequence, n also reaches 1 under T. □

Lemma 117. *For any odd $n \in \mathbb{N}$, n belongs to a cycle under the original Collatz map C if and only if it belongs to a cycle under the compact map T.*

Proof. ($\Rightarrow$) If n is on a cycle of length $k > 0$ under C, we can count the number of even steps in the cycle to determine the corresponding number of steps under T. The "halving" steps in the original sequence correspond exactly to the steps in the compact sequence. Since the sequence is periodic, it must eventually return to n under T. ($\Leftarrow$) If $T^k(n) = n$ for some $k > 0$, since every step of T can be simulated by one or two steps of C (Lemma 114), it follows that $C^j(n) = n$ for some $j \geq k > 0$. □

Lemma 118. *For any $n \in \mathbb{N}$, the orbit of n under C is bounded if and only if the orbit of n under T is bounded.*

Proof. ($\Rightarrow$) Since every value in the T-orbit is also present in the C-orbit (Lemma 115), if the C-orbit is bounded by B, then the T-orbit is also bounded by B. ($\Leftarrow$) If the T-orbit is bounded by B, then any value $C^d(n)$ in the C-orbit is either a value $T^k(n) \leq B$, or it is an "intermediate" odd step $m = C^d(n)$ such that $C(m) = 3m+1$ is even and $C(3m+1) = (3m+1)/2 = T^k(n) \leq B$. In the latter case, $3m + 1 \leq 2B$, so $m \leq (2B - 1)/3 < B$. Thus the entire C-orbit is bounded by $2B + 1$. □

Chapter 4

Rozier & Terracol (2025)

4.1 Rozier–Terracol Lemma 2.1

Lemma 119. *Let $k \leq j$. If $E_k(n) < E_k(m)$ and the parity bits of n and m agree for steps $i \in \{k, \dots, j-1\}$ (i.e., $X(T^i(n)) = X(T^i(m))$ for $k \leq i < j$), then $E_j(n) < E_j(m)$.*

Proof. We proceed by induction on j. The base case $j = k$ follows from the hypothesis. For the inductive step $j \to j+1$, we assume $E_j(n) < E_j(m)$ and $X(T^j(n)) = X(T^j(m)) = x$. Then:

$$E_{j+1}(n) = \frac{3^x}{2}E_j(n) + \frac{x}{2} < \frac{3^x}{2}E_j(m) + \frac{x}{2} = E_{j+1}(m) \tag{4.1}$$

since $3^x/2 > 0$. ☐

Lemma 120. *For any $i < j$, the i-th entry of the parity vector $V_j(n)$ is true if and only if $X(T^i(n)) = 1$. It is false if and only if $X(T^i(n)) = 0$.*

Proof. Immediate from the definition of $V_j(n)$. ☐

Lemma 121. *If the parity vectors $V_j(m)$ and $V_j(n)$ agree on their first k entries ($k \leq j$), then $E_k(m) = E_k(n)$.*

Proof. Since $E_k(n)$ depends only on the parity bits $X(T^0(n)), \dots, X(T^{k-1}(n))$, agreement of the first k entries of the parity vector implies equality of these bits, and thus equality of the correction terms and E_k. ☐

Lemma 122. *If $V_j(m)$ and $V_j(n)$ differ by a single elementary swap $01 \to 10$ (i.e., $V_j(m) \prec_{el} V_j(n)$), then $E_j(n) < E_j(m)$.*

Proof. Let the swap occur at positions k and $k+1$. Before position k, the parity vectors agree, so $E_k(n) = E_k(m)$. In $V_j(m)$, we have bits $(0,1)$ at $(k, k+1)$, while in $V_j(n)$ we have $(1,0)$. Calculating the two-step increase:

$$E_{k+2}(m) = \frac{3^1}{2}\left(\frac{3^0}{2}E_k(m) + \frac{0}{2}\right) + \frac{1}{2} = \frac{3}{4}E_k(m) + \frac{1}{2}$$
$$E_{k+2}(n) = \frac{3^0}{2}\left(\frac{3^1}{2}E_k(n) + \frac{1}{2}\right) + \frac{0}{2} = \frac{3}{4}E_k(n) + \frac{1}{4}$$

Thus $E_{k+2}(n) < E_{k+2}(m)$. Since the bits agree for all positions $i > k+1$, Lemma 119 implies $E_j(n) < E_j(m)$. ☐

Lemma 123. *The correction ratio E is strictly monotonic with respect to the transitive closure of the elementary precedence relation on parity vectors.*

Proof. This follows by induction on the number of elementary swaps, applying Lemma 122 at each step. □

Theorem 124 (Rozier–Terracol Lemma 2.1)**.** *If $V_j(m)$ strictly precedes $V_j(n)$ in the partial order generated by elementary swaps (i.e., $V_j(m) \prec V_j(n)$), then $E_j(n) < E_j(m)$.*

Proof. This is exactly the statement that the precedence relation $V_j(m) \prec V_j(n)$ implies $E_j(n) < E_j(m)$, which follows from Lemma 123 by identifying the sequence-based E_j with the vector-based calculation. □

4.2 Rozier–Terracol Theorem 2.2

Definition 125. The lower bound sequence $L_j(q)$ is defined for $j, q \in \mathbb{N}$ by:

$$L_j(q) = \frac{3^q - 2^q}{2^j}. \tag{4.2}$$

Definition 126. The upper bound $R(q)$ is defined for $q \in \mathbb{N}$ by:

$$R(q) = \frac{3^q - 2^q}{2^q} = \left(\frac{3}{2}\right)^q - 1. \tag{4.3}$$

Theorem 127 (Rozier–Terracol Theorem 2.2)**.** *For every positive integer j and any $n \in \mathbb{N}$, let q be the number of odd steps in the first j iterations of the compact Collatz map on n. Then the correction ratio $E_j(n)$ satisfies:*

$$L_j(q) \le E_j(n) \le R(q). \tag{4.4}$$

Proof. We proceed by induction on j. The base case $j = 1$ is trivial. For the inductive step $j+1$, let q be the number of odd steps in the first j iterations and $x \in \{0, 1\}$ be the $(j+1)$-th parity bit. Then the number of odd steps in $j+1$ iterations is $q+x$. By the recursive definition, $E_{j+1}(n) = \frac{3^x}{2}E_j(n) + \frac{x}{2}$. If $x = 0$, $E_{j+1}(n) = E_j(n)/2$. Since $L_{j+1}(q) = L_j(q)/2$ and $R(q)/2 \le R(q)$, the bounds hold. If $x = 1$, $E_{j+1}(n) = (3E_j(n) + 1)/2$. Since $L_{j+1}(q+1) = (3L_j(q) + 2^q/2^j)/2$ and $2^q/2^j \le 1$ (as $q \le j$), and $R(q+1) = (3R(q) + 1)/2$, the bounds follow from the induction hypothesis. □

Theorem 128. *For $j > 0$, the upper bound $E_j(n) = R(q)$ is reached if and only if all q odd steps occur at the end of the first j iterations. That is, the parity vector $V_j(n)$ consists of $j - q$ zeros followed by q ones.*

Proof. By induction on j. The condition $E_{j+1}(n) = R(q + x)$ requires $E_j(n) = R(q)$ and $x = 1$ (if $q > 0$), or $E_j(n) = 0$ and $q = 0$ and $x = 0$. This forces the bits to align at the end. □

Theorem 129. *For $j > 0$, the lower bound $E_j(n) = L_j(q)$ is reached if and only if all q odd steps occur at the beginning of the first j iterations. That is, the parity vector $V_j(n)$ consists of q ones followed by $j - q$ zeros.*

Proof. Similar to the upper bound, by induction on j. Equality $E_{j+1}(n) = L_{j+1}(q + x)$ forces the bits to be "front-loaded." □

4.3 Rozier–Terracol Lemma 2.3

Lemma 130. *For any $j, m \in \mathbb{N}$, the parity of the iterates at m and $m + 2^j$ are opposite:*

$$X(T^j(m)) + X(T^j(m + 2^j)) = 1. \tag{4.5}$$

Proof. By the linear decomposition $2^j T^j(n) = 3^q n + \mathcal{Q}_j(n)$, and the fact that $V_j(m+2^j) = V_j(m)$, we have $T^j(m + 2^j) = T^j(m) + 3^q$. Since 3^q is always odd, $T^j(m + 2^j)$ and $T^j(m)$ have opposite parity. □

Lemma 131. *The correction ratio $E_j(n)$ is periodic with period 2^j:*

$$E_j(m + 2^j) = E_j(m). \tag{4.6}$$

Proof. Since $E_j(n)$ depends only on the parity vector $V_j(n)$, and $V_j(n)$ is periodic with period 2^j, the result follows. □

Lemma 132. *For any $j, m \in \mathbb{N}$, we have the identity:*

$$E_{j+1}(m) + E_{j+1}(m + 2^j) = 2E_j(m) + \frac{1}{2}. \tag{4.7}$$

Proof. Using the recurrence $E_{j+1}(n) = \frac{3^{X(T^j(n))}}{2} E_j(n) + \frac{X(T^j(n))}{2}$ and the shift properties:

$$\begin{aligned}
E_{j+1}(m) + E_{j+1}(m + 2^j) &= \frac{3^{X(T^j(m))} + 3^{X(T^j(m+2^j))}}{2} E_j(m) + \frac{X(T^j(m)) + X(T^j(m + 2^j))}{2} \\
&= \frac{3^0 + 3^1}{2} E_j(m) + \frac{1}{2} \\
&= 2E_j(m) + \frac{1}{2}.
\end{aligned}$$

□

Theorem 133 (Rozier–Terracol Lemma 2.3)**.** *For every positive integer j, the average value of the correction ratio $E_j(n)$ over a full period is $j/4$:*

$$\frac{1}{2^j} \sum_{n=1}^{2^j} E_j(n) = \frac{j}{4}. \tag{4.8}$$

Proof. By induction on j. For $j = 1$, the average is $(E_1(1) + E_1(2))/2 = (1/2 + 0)/2 = 1/4$. For the inductive step $j + 1$, we split the sum over $\{1, \dots, 2^{j+1}\}$ into two halves and use Lemma 132:

$$\begin{aligned}
\sum_{n=1}^{2^{j+1}} E_{j+1}(n) &= \sum_{n=1}^{2^j} (E_{j+1}(n) + E_{j+1}(n + 2^j)) \\
&= \sum_{n=1}^{2^j} \left(2E_j(n) + \frac{1}{2}\right) \\
&= 2\left(\frac{j}{4} \cdot 2^j\right) + \frac{2^j}{2} \\
&= \frac{j \cdot 2^j + 2^j}{2} = \frac{(j + 1)2^{j+1}}{4}.
\end{aligned}$$

Dividing by 2^{j+1} yields $(j + 1)/4$. □

4.4 Diophantine Approximations

Definition 134. The irrational constant ξ is defined as:

$$\xi = \frac{\log 2}{\log 3}. \tag{4.9}$$

Lemma 135. *The constant ξ is irrational.*

Proof. Suppose for contradiction that $\xi = a/b$ for some integers a, b with $b \neq 0$. Then $b \log 2 = a \log 3$, which implies $\log(2^b) = \log(3^a)$. Since the exponential function is injective, we must have $2^b = 3^a$. However, an integer equation of the form $2^b = 3^a$ holds if and only if $a = b = 0$, contradicting the assumption that $b \neq 0$ (or checking via parity for nonzero a, b). Therefore, ξ cannot be written as a fraction of integers. □

Definition 136. For any $\varepsilon \in \mathbb{R}$, we define the bound $\delta(\varepsilon)$ as:

$$\delta(\varepsilon) = \frac{-\log(1-\varepsilon)}{\log 3}. \tag{4.10}$$

Lemma 137. *For any positive integers $a, b > 0$ and any $\varepsilon < 1$, the following equivalence holds:*

$$1 - \varepsilon < \frac{3^a}{2^b} < 1 \iff 0 < \xi - \frac{a}{b} < \frac{\delta(\varepsilon)}{b}. \tag{4.11}$$

Proof. We apply the natural logarithm to the inequality $1 - \varepsilon < \frac{3^a}{2^b} < 1$. The left-hand side gives $\log(1-\varepsilon) < a \log 3 - b \log 2$, which rearranges to $b \log 2 - a \log 3 < -\log(1-\varepsilon)$. Dividing by $b \log 3$ (which is positive since $b > 0$), we obtain $\xi - a/b < \delta(\varepsilon)/b$. The right-hand side gives $a \log 3 - b \log 2 < 0$, which is equivalent to $a \log 3 < b \log 2$. Dividing by $b \log 3$ yields $a/b < \xi$, or $0 < \xi - a/b$. These steps are reversible. □

Lemma 138. *There are infinitely many rational numbers $q = n/d < \xi$ such that*

$$|\xi - q| < \frac{1}{d^2}. \tag{4.12}$$

Proof. This is a standard consequence of the theory of continued fractions. The continued fraction convergents of an irrational number ξ alternate in being strictly less than and strictly greater than ξ. The even convergents $q_k = p_{2k}/q_{2k}$ provide an infinite sequence of rational approximations strictly smaller than ξ, satisfying the well-known approximation bound $|\xi - q_k| < 1/(q_k)^2$. Since ξ is irrational, the set of such convergents is infinite. □

Theorem 139. *(Lemma 3.1 in Rozier-Terracol 2025) For any $\varepsilon \in (0,1)$, there exist infinitely many pairs of positive integers (a, b) such that:*

$$1 - \varepsilon < \frac{3^a}{2^b} < 1. \tag{4.13}$$

Proof. Since $\varepsilon \in (0,1)$, the value $\delta(\varepsilon)$ is strictly positive. By Lemma 138, there are infinitely many rational approximations $q = a/b < \xi$ such that $|\xi - a/b| < 1/b^2$. Because there are infinitely many such approximations, we can choose ones where the denominator b is arbitrarily large. In particular, we restrict to the infinite subset where b is large enough such that $b \geq 2$ and $1/b \leq \delta(\varepsilon)$ (or more formally, $b \geq \lceil 1/\delta(\varepsilon) \rceil + 1$). For such an approximation, we have:

$$\xi - \frac{a}{b} = \left|\xi - \frac{a}{b}\right| < \frac{1}{b^2} \leq \frac{\delta(\varepsilon)}{b} \cdot \frac{1}{b} \leq \frac{\delta(\varepsilon)}{b}.$$

Since $q < \xi$, we have $0 < \xi - a/b$. Thus, $0 < \xi - a/b < \delta(\varepsilon)/b$. We ensure a and b are positive (a trivial consequence since $\xi > 0$ and b is large). By Lemma 137, this implies $1-\varepsilon < \frac{3^a}{2^b} < 1$. Since the subset constructed in this way is infinite and each fraction maps to a unique pair, there are infinitely many such pairs (a, b). □

4.5 Paradoxical Sequences

Definition 140. A pair (j, n) consisting of a positive iteration count j and a starting integer n is called *paradoxical* if:

$$T^j(n) \geq n \quad \text{and} \quad C_j(n) < 1. \tag{4.14}$$

That is, the j-th iterate of the Terras map is at least as large as the starting value, and the decomposition coefficient at j is smaller than 1. The sequence of values

$$n, T(n), T^2(n), \cdots, T^j(n)$$

is called a paradoxical sequence.

Lemma 141. *For any $k, j, n \in \mathbb{N}$, the correction ratio behaves under powers of two as:*

$$E_{k+j}(2^k n) = E_j(n). \tag{4.15}$$

Proof. This follows from the shift properties of the decomposition formula: the number of odd steps $q(k+j, 2^k n) = q(j, n)$ and the decomposition correction $\mathcal{Q}_{k+j}(2^k n) = 2^k \mathcal{Q}_j(n)$. Substituting these into the definition of E clears the 2^k factor. □

Lemma 142. *If $n \geq 1$ has an infinite stopping time, then the number of odd steps $q(j, n)$ is unbounded as $j \to \infty$.*

Proof. Suppose $q(j, n)$ were bounded by some M. Since q is non-decreasing, it must eventually stay constant at some value q_s. This would imply that all iterates $T^j(n)$ for $j \geq J$ are even, so they are repeatedly divided by 2. This contradicts the infinite stopping time, as the sequence would eventually drop below n (or reach 1 and then cycle). □

Lemma 143. *If $n \geq 1$ has an infinite stopping time, then for every $a \in \mathbb{N}$, there exists an iteration count j such that $q(j, n) = a$.*

Proof. Since $q(0, n) = 0$, $q(j, n)$ is unbounded, and $q(j + 1, n) - q(j, n) \in \{0, 1\}$, the discrete intermediate value theorem implies that q must hit every natural number. □

Theorem 144 (Rozier–Terracol Theorem 3.2)**.** *If the integer n has an infinite stopping time, then there exist infinitely many paradoxical sequences starting from integers of the form $2^k n$ with $k \geq 0$.*

Proof. We construct an infinite set of paradoxical pairs (j', n') where n' is of the form $2^k n$. Consider the set of Diophantine approximations $S = \{(a, b) \in \mathbb{N}^2 \mid 1 - \frac{1}{4n} < \frac{3^a}{2^b} < 1\}$. By Theorem 139, this set is infinite. For each $(a, b) \in S$, we use Lemma 143 to find an iteration j such that $q(j, n) = a$. Let $j_{of}(a)$ be the smallest such j. We distinguish two cases for each $(a, b) \in S$:

Case 1: $j_{of}(a) \geq b$**.** We claim that $(j_{of}(a), n)$ is paradoxical. First, $T^{j_{of}(a)}(n) \geq n$ because n has infinite stopping time. Second, $C_{j_{of}(a)}(n) = \frac{3^a n + \mathcal{Q}_j(n)}{2^{j_{of}(a)}} \approx \frac{3^a}{2^{j_{of}(a)}} n$. Since $j_{of}(a) \geq b$ and

$3^a/2^b < 1$, we have $3^a/2^{j_{of}(a)} < 1$. More precisely, $C_j(n) = \frac{3^a}{2^j}n + E_j(n)$. Using the bounds from Theorem 127, $E_j(n) \le R(a) = \frac{3^a - 2^a}{2^a}$. Then:

$$C_j(n) \le \frac{3^a}{2^b}n + \left(\frac{3}{2}\right)^a - 1. \tag{4.16}$$

The condition $3^a/2^b > 1 - 1/(4n)$ ensures that $C_j(n) < 1$.

Case 2: $j_{of}(a) < b$**.** Let $k = b - j_{of}(a)$. We claim $(b, 2^k n)$ is paradoxical. By the shift property $T^b(2^k n) = T^{j_{of}(a)}(n)$, and since $T^{j_{of}(a)}(n) \ge n$ and $2^k n \ge n$ is not guaranteed to be small, we calculate: $C_b(2^k n) = \frac{3^a(2^k n) + \mathcal{Q}_b(2^k n)}{2^b} = \frac{2^k(3^a n + \mathcal{Q}_j(n))}{2^{k+j}} = C_j(n)$. Again, the Diophantine condition $3^a/2^b > 1 - 1/(4n)$ implies $C_b(2^k n) < 1$, and $T^b(2^k n) \ge 2^k n$ follows from the linear decomposition:

$$2^j T^j(n) = 3^a n + \mathcal{Q}_j(n). \tag{4.17}$$

Substituting $T^b(2^k n) = T^j(n)$, we need $T^j(n) \ge 2^k n$. Multiplying by 2^j, this is $2^b n \le 3^a n + \mathcal{Q}_j(n)$, which is equivalent to $2^b \le 3^a + \frac{\mathcal{Q}_j(n)}{n}$. The condition $3^a/2^b > 1 - 1/(4n)$ provides exactly the separation needed to satisfy this inequality when n is large or when $\mathcal{Q}_j$ is at least $3^a - 2^a$.

Since S is infinite and the mapping from (a, b) to (j', k) is injective, we obtain infinitely many paradoxical sequences. □

Corollary 145 (Rozier–Terracol Corollary 3.3)**.** *If the set of paradoxical sequences (j, m) with $m > 2$ is finite, then the Collatz conjecture is true.*

Proof. Assume the set of paradoxical sequences with $m > 2$ is finite. Let n be an arbitrary natural number. If $n = 0$, the conjecture is vacuously true. Suppose $n > 0$ and n never reaches 1 under Collatz iteration.

Let $\mathcal{O}(n)$ be the orbit of n under the Terras map T. Let $m = \min \mathcal{O}(n)$.

1. $m \ge 1$ since $n \ge 1$.
2. $m \ne 1$ because if $1 \in \mathcal{O}(n)$, then n would reach 1.
3. $m \ne 2$ because if $2 \in \mathcal{O}(n)$, its next iterate $T(2) = 1$ would also be in $\mathcal{O}(n)$.
4. Thus, $m > 2$.
5. Since m is the minimal element of the orbit, for any $k \ge 1$, $T^k(m) \ge m$. This means the stopping time of m is infinite.

By Theorem 144, there exist infinitely many paradoxical sequences starting from integers of the form $2^k m$. Since $m > 2$, all such integers $2^k m$ are also strictly greater than 2. This demonstrates the existence of infinitely many paradoxical sequences with starting value greater than 2, which contradicts our initial hypothesis. Therefore, every $n > 0$ must eventually reach 1. □

4.6 More Paradoxical Sequences

Lemma 146. *Let $j, n \in \mathbb{N}$ with $n > 0$. If the pair (j, n) is paradoxical, then:*

$$0 < 1 - C(j, n) \le \frac{E(j, n)}{n}. \tag{4.18}$$

Proof. From the paradoxical condition $C(j,n) < 1$, we get $0 < 1 - C(j,n)$. By Lemma 109, $T^j(n) = C(j,n) \cdot n + E(j,n)$. The paradoxical condition $T^j(n) \geq n$ therefore gives $C(j,n) \cdot n + E(j,n) \geq n$, hence $E(j,n) \geq (1 - C(j,n)) \cdot n$, and dividing by $n > 0$ yields the result. ☐

Lemma 147. *Let $j, n, m \in \mathbb{N}$ with $m > 0$. Suppose that every odd iterate $T^k(n)$ with $k < j$ satisfies $T^k(n) \geq m$. Let $q = q(j,n)$ be the number of odd steps in the first j iterations. Then:*

$$T^j(n) \leq \frac{n \cdot \left(3 + \frac{1}{m}\right)^q}{2^j}. \tag{4.19}$$

Proof. By induction on j. The base case $j = 0$ is trivial.

For the inductive step, assume the bound holds for j. We consider two cases depending on the parity of $T^j(n)$:

Odd case $(X(T^j(n)) = 1)$: By definition,

$$T^{j+1}(n) = T(T^j(n)) \leq T^j(n) \cdot \frac{3 + 1/m}{2}. \tag{4.20}$$

Combining with the inductive hypothesis $T^j(n) \leq n(3 + 1/m)^q/2^j$ and using $q_{\text{new}} = q + 1$:

$$T^{j+1}(n) \leq \frac{n(3 + 1/m)^q}{2^j} \cdot \frac{3 + 1/m}{2} = \frac{n(3 + 1/m)^{q+1}}{2^{j+1}}. \tag{4.21}$$

Even case $(X(T^j(n)) = 0)$: By definition, $T^{j+1}(n) = T^j(n)/2$. Since q stays the same,

$$T^{j+1}(n) \leq \frac{n(3 + 1/m)^q}{2^j \cdot 2} = \frac{n(3 + 1/m)^q}{2^{j+1}}. \quad ☐ \tag{4.22}$$

Lemma 148. *Under the same hypotheses as Lemma 147, we have:*

$$\frac{E(j,n)}{n} \leq \frac{\left(3 + \frac{1}{m}\right)^q - 3^q}{2^j}. \tag{4.23}$$

Proof. By Lemma 147:

$$T^j(n) \leq \frac{n(3 + 1/m)^q}{2^j}. \tag{4.24}$$

Substituting the factored form $T^j(n) = C(j,n) \cdot n + E(j,n)$ from Lemma 109 and dividing through by $n > 0$:

$$C(j,n) + \frac{E(j,n)}{n} \leq \frac{(3 + 1/m)^q}{2^j}. \tag{4.25}$$

Since $C(j,n) = 3^q/2^j$, rearranging gives the result. ☐

Since the harmonic mean h is not less than the minimum m of the odd iterates of T, Lemma 148 is a weaker version of Theorem 4.1 in Rozier–Terracol.

Lemma 149. *Let $j, q, m \in \mathbb{N}$ with $j > 0$ and $m > 0$. Suppose the following inequalities hold for the rationals $3^q/2^j$ and $(3 + 1/m)^q/2^j$:*

$$1 - \frac{3^q}{2^j} > 0 \quad \textit{and} \quad 1 - \frac{3^q}{2^j} \leq \frac{(3 + 1/m)^q - 3^q}{2^j}. \tag{4.26}$$

Then the ratio q/j satisfies:

$$\frac{\log 2}{\log(3 + 1/m)} \leq \frac{q}{j} < \frac{\log 2}{\log 3}. \tag{4.27}$$

Proof. For the upper bound, $1 - 3^q/2^j > 0$ implies $3^q < 2^j$. Taking logs gives $q \log 3 < j \log 2$, so $q/j < \log 2/\log 3$. For the lower bound, the second inequality implies $2^j - 3^q \leq (3+1/m)^q - 3^q$, which simplifies to $2^j \leq (3+1/m)^q$. Taking logs gives $j \log 2 \leq q \log(3+1/m)$, yielding the final result. □

Corollary 150. *Let (j, n) be a paradoxical sequence starting at $n > 0$. Let $m > 0$ be a lower bound for all odd iterates $T^i(n)$ with $i < j$. Let q be the number of odd steps. Then:*

$$\frac{\log 2}{\log(3+1/m)} \leq \frac{q}{j} < \frac{\log 2}{\log 3}. \tag{4.28}$$

Proof. This follows from Lemma 148 and the algebraic properties derived in Lemma 149. □

Since the harmonic mean h is not less than the minimum m of the odd iterates of T, Lemma 150 is a weaker version of Corollary 4.2 in Rozier–Terracol.

Lemma 151. *For any $j \geq 2$ and any $k \in \mathbb{N}$, the value $j \cdot \log 2/\log 3$ is not an integer:*

$$\frac{j \log 2}{\log 3} \neq k. \tag{4.29}$$

Proof. If $j \log 2/\log 3 = k$, then $\xi = \log 2/\log 3 = k/j \in \mathbb{Q}$, contradicting the irrationality of ξ (Lemma 135). □

Lemma 152. *For any $j \geq 2$, the floor of $j \log 2/\log 3$ is strictly less than $j \log 2/\log 3$:*

$$\left\lfloor \frac{j \log 2}{\log 3} \right\rfloor < \frac{j \log 2}{\log 3}. \tag{4.30}$$

Proof. By definition, $\lfloor x \rfloor \leq x$. The floor must be strictly less because $j \log 2/\log 3$ is not an integer by Lemma 151. □

Lemma 153. *Let $j, q, m \in \mathbb{N}$ with $j, q, m > 0$. Suppose the ratio bounds*

$$\frac{\log 2}{\log(3+1/m)} \leq \frac{q}{j} < \frac{\log 2}{\log 3} \tag{4.31}$$

hold. Then:

1. $2^{j/q} - 3 > 0$*, i.e.,* $2^{j/q} > 3$*;*
2. $m \leq \dfrac{1}{2^{j/q} - 3}$.

Proof. The upper bound $q/j < \log 2/\log 3$ implies $q \log 3 < j \log 2$, hence $\log 3 < (j/q) \log 2 = \log(2^{j/q})$, so $2^{j/q} > 3$.

The lower bound $\log 2/\log(3+1/m) \leq q/j$ implies $j \log 2 \leq q \log(3+1/m)$, i.e., $(j/q) \log 2 \leq \log(3+1/m)$, so $2^{j/q} \leq 3 + 1/m$. This gives $2^{j/q} - 3 \leq 1/m$, and since $2^{j/q} - 3 > 0$, we can take reciprocals (reversing the inequality) to get $m \leq 1/(2^{j/q} - 3)$. □

Lemma 154. *Let $j, q, m \in \mathbb{N}$ with $j \geq 2$ and $m > 0$. Suppose the ratio bounds of Lemma 149 hold for q/j. Define:*

$$r = \frac{j}{\lfloor j \log 2/\log 3 \rfloor}. \tag{4.32}$$

Then $2^r - 3 > 0$ and $m \leq 1/(2^r - 3)$.

Proof. From $q \leq \lfloor j \log 2/\log 3 \rfloor$ (by the upper bound $q < j \log 2/\log 3$), we have $r = j/\lfloor j \log 2/\log 3 \rfloor \leq j/q$. This means $2^r \leq 2^{j/q}$, so $1/(2^r - 3) \geq 1/(2^{j/q} - 3)$. By Lemma 153, $m \leq 1/(2^{j/q} - 3) \leq 1/(2^r - 3)$.

To show $2^r > 3$: since $\lfloor j \log 2/\log 3 \rfloor < j \log 2/\log 3$ (strictly, by Lemma 152), $r = j/\lfloor j \log 2/\log 3 \rfloor > j/(j \log 2/\log 3) = \log 3/\log 2$. Therefore $r \log 2 > \log 3$, i.e., $\log(2^r) > \log 3$, so $2^r > 3$. □

Corollary 155. *Let (j, n) be a paradoxical pair with $j \geq 2$ and $n > 0$. Let $m > 0$ be a lower bound for all odd iterates $T^k(n)$ with $k < j$. Define $r = j/\lfloor j \log 2/\log 3 \rfloor$. Then:*

$$2^r - 3 > 0 \quad \textit{and} \quad m \leq \frac{1}{2^r - 3}. \tag{4.33}$$

Proof. Combine the ratio bounds from Lemma 150 with Lemma 154. □

Since the harmonic mean h is not less than the minimum m of the odd iterates of T, Lemma 155 is weaker than Corollary 4.3 in Rozier–Terracol.

Chapter 5

Eliahou & Verger-Gaugry (2025)

5.1 Uniform Distribution Correspondence

Using the compact map T, Eliahou & Verger-Gaugry (2025)[2] established that their Conjecture 2.4 (every positive integer eventually reaches 1 under the map T) is equivalent to Conjecture 2.5 (for every positive integer n, the proportion of even iterates among the first k iterates of T tends to 1/2 as $k \to \infty$).

See Definition 81 for $Q(k, n)$ as the number of odd steps in the first k iterations of T starting from n.

Lemma 156. *For all $a, b, n \in \mathbb{N}$,*

$$Q(a + b,\, n) = Q(a,\, n) + Q(b,\, T^a(n)).$$

Proof. The count of odd steps over the first $a + b$ iterates splits into the count over the first a iterates starting from n, plus the count over the next b iterates starting from $T^a(n)$. This follows directly from the additivity of finite sums and the identity $T^{a+j}(n) = T^j(T^a(n))$. □

Lemma 157. *For all $j \in \mathbb{N}$,*

$$T^{2j}(1) = 1.$$

Proof. By induction on j. The base case $j = 0$ is trivial. For the inductive step, one uses $T(1) = 2$ (since 1 is odd: $T(1) = (3 \cdot 1 + 1)/2 = 2$) and $T(2) = 1$ (since 2 is even: $T(2) = 2/2 = 1$), so $T^{2(j+1)}(1) = T^2(T^{2j}(1)) = T^2(1) = T(2) = 1$. □

Lemma 158. *For all $j \in \mathbb{N}$,*

$$T^{2j+1}(1) = 2.$$

Proof. We have $T^{2j+1}(1) = T(T^{2j}(1)) = T(1) = 2$ by Lemma 157 and the fact that $T(1) = 2$. □

Lemma 159. *For all $i \in \mathbb{N}$, the parity indicator of $T^i(1)$ satisfies*

$$X(T^i(1)) = \begin{cases} 1 & \text{if } i \equiv 0 \pmod 2, \\ 0 & \text{if } i \equiv 1 \pmod 2. \end{cases}$$

Proof. Write $i = 2j$ or $i = 2j + 1$. In the even case, $T^{2j}(1) = 1$ by Lemma 157, and 1 is odd so $X(1) = 1$. In the odd case, $T^{2j+1}(1) = 2$ by Lemma 158, and 2 is even so $X(2) = 0$. □

Lemma 160. *For all $j \in \mathbb{N}$,*

$$Q(2j,\, 1) = j.$$

Proof. By induction on j. The base case is trivial. For the inductive step, use Lemma 156 to split the $2(j+1)$ steps as $2j$ steps followed by 2 steps starting from $T^{2j}(1) = 1$ (by Lemma 157). The count for the last 2 steps starting from 1 is 1 (since $T^0(1) = 1$ is odd and $T^1(1) = 2$ is even), so the total is $j + 1$. □

Lemma 161. *For all $j \in \mathbb{N}$,*

$$Q(2j + 1,\, 1) = j + 1.$$

Proof. Apply Lemma 156 to split the $2j + 1$ steps as $2j$ steps followed by 1 step starting from $T^{2j}(1) = 1$ (by Lemma 157). The count for the first $2j$ steps is j by Lemma 160, and the single additional step starts at 1 which is odd, contributing 1 more. Total: $j + 1$. □

Lemma 162. *For all $m \in \mathbb{N}$, letting $s = Q(m,\, 1)$,*

$$m \le 2s + 1 \quad \textit{and} \quad 2s \le m + 1.$$

In other words, $2s \in \{m, m + 1\}$.

Proof. Consider the two cases according to the parity of m.

- If $m = 2j$, then $s = j$ by Lemma 160, so $2s = m$ and both bounds hold with equality (up to ± 1).
- If $m = 2j + 1$, then $s = j + 1$ by Lemma 161, so $2s = 2j + 2 = m + 1$, and again both bounds hold.

□

Lemma 163. *Let $n \ge 1$ and suppose there exists $k_0 \ge 1$ such that $T^{k_0}(n) = 1$. Then the proportion of even iterates satisfies*

$$\lim_{k \to \infty} \frac{k - Q(k,\, n)}{k} = \frac{1}{2}.$$

Proof. Let $s_k = Q(k, n)$. For $k \ge k_0$, write $k = k_0 + (k - k_0)$ and apply Lemma 156:

$$s_k = Q(k_0, n) + Q(k - k_0, 1),$$

using that $T^{k_0}(n) = 1$. Let $c = Q(k_0, n)$, which is a fixed constant with $c \le k_0$. By Lemma 162 applied to $m = k - k_0$, we have $2Q(k - k_0, 1) \in \{k - k_0,\, k - k_0 + 1\}$, so:

$$2s_k \le k + (k_0 + 1) \quad \text{and} \quad k \le 2s_k + (k_0 + 1).$$

Using the order-topology characterisation of convergence, we show the limit equals $1/2$:

- *Lower bound:* For $a < 1/2$, set $\varepsilon = 1 - 2a > 0$. Choose N large enough that $(k_0 + 1)/\varepsilon < N$ and $k \ge \max(N, k_0 + 1)$. Then from $k \le 2s_k + (k_0 + 1)$ we get $(k - s_k)/k > a$.
- *Upper bound:* For $b > 1/2$, set $\varepsilon = 2b - 1 > 0$. Choose N large enough that $(k_0 + 1)/\varepsilon < N$ and $k \ge \max(N, k_0 + 1)$. Then from $2s_k \le k + (k_0 + 1)$ we get $(k - s_k)/k < b$.

In both cases, the bound follows by a simple linear arithmetic argument once k is large enough relative to $(k_0 + 1)/\varepsilon$. □

Lemma 164. *For every n with $1 \le n \le 9$, there exists $k \ge 1$ such that $T^k(n) = 1$.*

Proof. Verified by explicit computation for each value $n \in \{1, 2, \dots, 9\}$: $T^2(1) = 1$, $T^1(2) = 1$, $T^5(3) = 1$, $T^2(4) = 1$, $T^4(5) = 1$, $T^6(6) = 1$, $T^{11}(7) = 1$, $T^3(8) = 1$, $T^{13}(9) = 1$. ☐

Lemma 165. *For every odd $n \ge 10$,*

$$10 \cdot (3n + 1) \le 31n.$$

Proof. This is equivalent to $30n + 10 \le 31n$, i.e. $10 \le n$, which holds by assumption. ☐

Lemma 166. *Let $k, n \in \mathbb{N}$ and suppose that $T^j(n) \ge 10$ for all $j \le k$. Let $s = Q(k, n)$. Then*

$$10^s \cdot 2^k \cdot T^k(n) \le 31^s \cdot n.$$

Proof. By induction on k.

Base case ($k = 0$): both sides equal n.

Inductive step: Assume the bound holds for k, with $s_k = Q(k, n)$, and that $T^j(n) \ge 10$ for all $j \le k + 1$. Let $n_k = T^k(n)$ and $s_{k+1} = s_k + X(n_k)$, where $X(n_k) \in \{0, 1\}$ is the parity indicator.

Case n_k even ($X(n_k) = 0$): Then $T(n_k) = n_k/2$, so $2T(n_k) = n_k$, and $s_{k+1} = s_k$.

$$\begin{aligned} 10^{s_k} \cdot 2^{k+1} \cdot T(n_k) &= 10^{s_k} \cdot 2^k \cdot (2T(n_k)) \\ &= 10^{s_k} \cdot 2^k \cdot n_k \\ &\le 31^{s_k} \cdot n, \end{aligned}$$

where the last step uses the induction hypothesis.

Case n_k odd ($X(n_k) = 1$): Then $T(n_k) = (3n_k + 1)/2$, so $2T(n_k) = 3n_k + 1$, and $s_{k+1} = s_k + 1$. Since $n_k \ge 10$ and n_k is odd, Lemma 165 gives $10(3n_k + 1) \le 31n_k$.

$$\begin{aligned} 10^{s_k+1} \cdot 2^{k+1} \cdot T(n_k) &= 10^{s_k} \cdot 2^k \cdot (10 \cdot (3n_k + 1)) \\ &\le 10^{s_k} \cdot 2^k \cdot (31\, n_k) \\ &= 31 \cdot (10^{s_k} \cdot 2^k \cdot n_k) \\ &\le 31 \cdot 31^{s_k} \cdot n \\ &= 31^{s_k+1} \cdot n, \end{aligned}$$

where the second-to-last step uses the induction hypothesis. ☐

Lemma 167. *For all $s, k \in \mathbb{N}$ with $5s + 1 \le 3k$,*

$$31^s < 2^k \cdot 10^s.$$

Proof. By strong induction on s, with step size 3.

Base cases:

- $s = 0$: Need $1 < 2^k$. From the hypothesis $5 \cdot 0 + 1 \le 3k$ we get $k \ge 1$, so $2^k \ge 2 > 1$.
- $s = 1$: Need $31 < 2^k \cdot 10$. From $5 + 1 \le 3k$ we get $k \ge 2$, so $2^k \cdot 10 \ge 40 > 31$.
- $s = 2$: Need $961 < 2^k \cdot 100$. From $11 \le 3k$ we get $k \ge 4$, so $2^k \cdot 100 \ge 1600 > 961$.

Inductive step ($s \geq 3$, write $s = s' + 3$): Assume $31^{s'} < 2^{k'} \cdot 10^{s'}$ for all valid k', where $5s' + 1 \leq 3k'$. Apply the induction hypothesis with $k' = k - 5$ (valid since $5(s' + 3) + 1 \leq 3k$ implies $5s' + 1 \leq 3(k - 5)$, and $k \geq 6$):

$$31^{s'+3} = 31^3 \cdot 31^{s'} < 31^3 \cdot 2^{k-5} \cdot 10^{s'}.$$

Using the key arithmetic fact $31^3 = 29791 < 32000 = 2^5 \cdot 10^3$:

$$31^3 \cdot 2^{k-5} \cdot 10^{s'} \leq 2^5 \cdot 10^3 \cdot 2^{k-5} \cdot 10^{s'} = 2^k \cdot 10^{s'+3}.$$ □

Lemma 168. *Let $m \geq 1$ be a minimal counterexample to Conjecture 2.4, meaning:*

- *$T^k(m) \neq 1$ for all $k \geq 1$;*
- *for every $n < m$ with $n \geq 1$, there exists $k \geq 1$ with $T^k(n) = 1$.*

Then $T^j(m) \geq m$ for all $j \in \mathbb{N}$.

Proof. Suppose for contradiction that $T^j(m) < m$ for some j. The case $j = 0$ gives $m < m$, a contradiction. For $j = j' + 1$: let $n' = T^{j'+1}(m)$. Since $n' < m$ and $n' \geq 1$ (because iterates of positive integers stay positive), the minimality of m gives some $k \geq 1$ with $T^k(n') = 1$. But then $T^{k+(j'+1)}(m) = T^k(T^{j'+1}(m)) = T^k(n') = 1$, contradicting the assumption that m never reaches 1. □

Lemma 169. *Let $n \in \mathbb{N}$ and suppose that*

$$\lim_{k\to\infty} \frac{k - Q(k, n)}{k} = \frac{1}{2}.$$

Then there exists $k \geq 1$ such that

$$5 \cdot Q(k, n) + 1 \leq 3k.$$

Proof. Since the limit is $1/2$, the sequence eventually lies in the open interval $(2/5, 3/5)$. Choose N such that for all $k \geq N$ the ratio $(k - s_k)/k > 2/5$, where $s_k = Q(k, n)$. Set $k_0 = \max(N, 1)$. Then $(k_0 - s_{k_0})/k_0 > 2/5$, i.e.

$$k_0 - s_{k_0} > \tfrac{2}{5} k_0,$$

so $\frac{3}{5} k_0 > s_{k_0}$, which gives $5s_{k_0} < 3k_0$, i.e. $5s_{k_0} + 1 \leq 3k_0$. □

Lemma 170. *Conjectures 2.4 and 2.5 are equivalent:*

- Conjecture 2.4: *For every $n \geq 1$, there exists $k \geq 1$ such that $T^k(n) = 1$.*
- Conjecture 2.5: *For every $n \geq 1$,*

$$\lim_{k\to\infty} \frac{k - Q(k, n)}{k} = \frac{1}{2}.$$

Proof. **(2.4 ⇒ 2.5).** Let $n \geq 1$. By Conjecture 2.4, there exists $k_0 \geq 1$ with $T^{k_0}(n) = 1$. Lemma 163 then directly gives the desired Tendsto conclusion.

(2.5 ⇒ 2.4). Suppose for contradiction that Conjecture 2.4 fails. Then there exists some $n_0 \geq 1$ with $T^k(n_0) \neq 1$ for all $k \geq 1$. Among all such counterexamples, let m be the minimal one (using Nat.find); formally, let

$$P(n) \;\equiv\; (n \geq 1) \wedge (\forall k \geq 1,\ T^k(n) \neq 1),$$

and set $m = \min\{n : P(n)\}$. The properties of m are:

1. $m \geq 1$ and $T^k(m) \neq 1$ for all $k \geq 1$.

2. For every $n < m$ with $n \geq 1$, there exists $k \geq 1$ with $T^k(n) = 1$.

Step 1: $m \geq 10$. By Lemma 164, every n with $1 \leq n \leq 9$ reaches 1. So if $m \leq 9$, property (1) would be contradicted. Hence $m \geq 10$.

Step 2: $T^j(m) \geq m$ for all j. This is Lemma 168 applied to m with its minimality property. In particular, all iterates satisfy $T^j(m) \geq m \geq 10$.

Step 3: Extract k with a good odd-step ratio. Since Conjecture 2.5 holds (by assumption), the ratio $(k - s_k)/k \to 1/2$ for $n = m$. By Lemma 169, there exists $k \geq 1$ with

$$5 \cdot Q(k, m) + 1 \leq 3k.$$

Step 4: Derive a contradiction. Let $s = Q(k, m)$.

- From Step 2, all iterates $T^j(m) \geq 10$ for $j \leq k$, so Lemma 166 gives:

$$10^s \cdot 2^k \cdot T^k(m) \leq 31^s \cdot m.$$

- Lemma 167 with $5s + 1 \leq 3k$ gives:

$$31^s < 2^k \cdot 10^s.$$

Combining these two inequalities with $T^k(m) \geq m$ (Step 2):

$$31^s \cdot m \geq 10^s \cdot 2^k \cdot T^k(m) \geq 10^s \cdot 2^k \cdot m.$$

Dividing by $10^s \cdot m > 0$:

$$(31/10)^s \geq 2^k,$$

i.e. $31^s \geq 2^k \cdot 10^s$, contradicting Lemma 167. □

5.2 Base 3/2 Number System

5.2.1 The Number System in Rational Base 3/2

The following definitions and results formalise portions of the paper by Eliahou and Verger-Gaugry [2], which studies the representation of natural numbers in the rational base 3/2 and its connections to the Collatz problem. Every natural number n admits a unique representation

$$n = \frac{1}{2} \sum_{i=0}^{k} a_i \left(\frac{3}{2}\right)^i, \qquad a_i \in \{0, 1, 2\},$$

and the digits a_i can be extracted by repeated division of $2n$ by 3.

Definition 171. The *least-significant digit* of n in rational base 3/2 is

$$\text{lsd}(n) = (2n) \bmod 3.$$

This is the digit a_0 of Proposition 2.2 of [2].

Definition 172. The *parent* of n in the base-3/2 digit tree is

$$\mathrm{parent}(n) = \lfloor 2n/3 \rfloor.$$

Removing the least-significant digit from $\langle n \rangle$ gives $\langle \mathrm{parent}(n) \rangle$.

Lemma 173. *For all* $n \in \mathbb{N}$,

$$2n = 3 \cdot \mathrm{parent}(n) + \mathrm{lsd}(n).$$

Proof. This is the division algorithm applied to $2n$ divided by 3. ☐

Lemma 174. *For all* $n \in \mathbb{N}$, $\mathrm{lsd}(n) < 3$.

Proof. By the definition of the remainder modulo 3. ☐

Lemma 175. *For all* $n \geq 1$, $\mathrm{parent}(n) < n$.

Proof. We have $\mathrm{parent}(n) = \lfloor 2n/3 \rfloor \leq 2n/3 < n$ whenever $n \geq 1$. ☐

Lemma 176. *For all* $n \in \mathbb{N}$,

$$\mathrm{lsd}(n) \equiv \mathrm{parent}(n) \pmod 2.$$

This is part of Proposition 2.2 of [2].

Proof. From $2n = 3 \cdot \mathrm{parent}(n) + \mathrm{lsd}(n)$, reducing modulo 2 gives $0 \equiv \mathrm{parent}(n) + \mathrm{lsd}(n) \pmod 2$. ☐

Definition 177. The *base-3/2 representation* of n, written $\langle n \rangle$, is the list of digits (least-significant first) defined recursively:

$$\mathrm{digits}_{3/2}(n) = \begin{cases} [] & \text{if } n = 0, \\ \mathrm{lsd}(n) :: \mathrm{digits}_{3/2}(\mathrm{parent}(n)) & \text{if } n \geq 1. \end{cases}$$

Termination is guaranteed by Lemma 175.

Lemma 178. *Every digit in* $\mathrm{digits}_{3/2}(n)$ *is less than* 3.

Proof. By strong induction on n. If $n = 0$ the digit list is empty. Otherwise the first digit is $\mathrm{lsd}(n) < 3$ (Lemma 174), and the remaining digits belong to $\mathrm{digits}_{3/2}(\mathrm{parent}(n))$, which satisfy the bound by the induction hypothesis (since $\mathrm{parent}(n) < n$). ☐

Lemma 179. $|\,\mathrm{digits}_{3/2}(n)| = 0$ *if and only if* $n = 0$.

Proof. If $n = 0$ the list is empty. Conversely, if $n \geq 1$ the list has the form $\mathrm{lsd}(n) :: (\cdots)$, whose length is at least 1. ☐

Definition 180. The *evaluation* of a list of base-3/2 digits (least-significant first) is

$$\mathrm{eval}_{3/2}([]) = 0, \qquad \mathrm{eval}_{3/2}(a :: \mathit{rest}) = \frac{a + 3 \cdot \mathrm{eval}_{3/2}(\mathit{rest})}{2}.$$

For admissible digit lists (those produced by $\mathrm{digits}_{3/2}$), the division is always exact.

Lemma 181. *For all* $n \in \mathbb{N}$,

$$\mathrm{eval}_{3/2}(\mathrm{digits}_{3/2}(n)) = n.$$

Proof. By strong induction on n. For $n = 0$ both sides are 0. For $n \geq 1$:

$$\begin{aligned} \mathrm{eval}_{3/2}(\mathrm{lsd}(n) :: \mathrm{digits}_{3/2}(\mathrm{parent}(n))) &= \frac{\mathrm{lsd}(n) + 3 \cdot \mathrm{parent}(n)}{2} && \text{(by the induction hypothesis)} \\ &= \frac{2n}{2} = n, && \text{(by Lemma 173)} \end{aligned}$$

where the division is exact since $\mathrm{lsd}(n) + 3 \cdot \mathrm{parent}(n) = 2n$. □

Definition 182. The function $U : \mathbb{N} \to \mathbb{N}$ (Notation 3.7 of [2]) is defined by

$$U(n) = \begin{cases} (3n+2)/2 & \text{if } n \text{ is even,} \\ (3n+1)/2 & \text{if } n \text{ is odd.} \end{cases}$$

Equivalently, $U(n) = (3n + 2 - (n \bmod 2))/2$.

Lemma 183. *When n is odd, $U(n) = T(n)$.*

Proof. Both evaluate to $(3n + 1)/2$. □

Lemma 184. *For all $n \in \mathbb{N}$,*

$$U(n) \equiv T(n) + n + 1 \pmod 2.$$

This is Remark 3.8 of [2].

Proof. Split into cases according to the parity of n and verify by direct computation:

- If $n = 2k$, then $U(n) = 3k + 1$ and $T(n) = k$, so $U(n) \bmod 2 = (k + 1) \bmod 2 = (k + 2k + 1) \bmod 2 = (T(n) + n + 1) \bmod 2$.
- If $n = 2k+1$, then $U(n) = 3k+2$ and $T(n) = 3k+2$, and again both sides match modulo 2.

□

Lemma 185. *For all $n \in \mathbb{N}$,*

$$\mathrm{parent}(U(n)) = n.$$

This is Proposition 3.9 of [2].

Proof. Split by parity of n and verify the integer arithmetic:

- If n is even: $U(n) = (3n+2)/2$, so $\mathrm{parent}(U(n)) = \lfloor 2 \cdot (3n+2)/2 \,/\, 3 \rfloor = \lfloor (3n+2)/3 \rfloor = n$.
- If n is odd: $U(n) = (3n + 1)/2$, so $\mathrm{parent}(U(n)) = \lfloor (3n + 1)/3 \rfloor = n$.

□

Lemma 186. *When n is odd,* $\mathrm{lsd}(U(n)) = 1$.

Proof. $U(n) = (3n + 1)/2$, so $\mathrm{lsd}(U(n)) = (2 \cdot (3n + 1)/2) \bmod 3 = (3n + 1) \bmod 3 = 1$. □

Lemma 187. *When n is even,* $\mathrm{lsd}(U(n)) = 2$.

Proof. $U(n) = (3n + 2)/2$, so $\mathrm{lsd}(U(n)) = (3n + 2) \bmod 3 = 2$. □

Lemma 188. *For all $n \in \mathbb{N}$, $U(n) \geq 1$.*

Proof. In both cases $(3n+2)/2 \geq 1$ and $(3n+1)/2 \geq 1$. $\square$

Lemma 189. *For all $n \in \mathbb{N}$, the base-3/2 digits of $U(n)$ are obtained by prepending 1 (if n is odd) or 2 (if n is even) to $\langle n \rangle$:*

$$\langle U(n) \rangle = \begin{cases} 1 :: \langle n \rangle & \textit{if } n \textit{ is odd,} \\ 2 :: \langle n \rangle & \textit{if } n \textit{ is even.} \end{cases}$$

This is Proposition 3.9 of [2] (combined form).

Proof. Since $U(n) \geq 1$ (Lemma 188), we can unfold the recursive definition to get $\langle U(n) \rangle = \text{lsd}(U(n)) :: \langle \text{parent}(U(n)) \rangle$. By Lemma 185, $\text{parent}(U(n)) = n$. The digit is 1 when n is odd (Lemma 186) and 2 when n is even (Lemma 187). $\square$

Lemma 190. *For odd n (Corollary 3.2 of [2]),*

$$\langle T(n) \rangle = 1 :: \langle n \rangle.$$

That is, the base-3/2 representation of $T(n) = (3n+1)/2$ is obtained by prepending digit 1 to $\langle n \rangle$.

Proof. Since n is odd, $T(n) = (3n+1)/2 \geq 1$. We verify that $\text{lsd}(T(n)) = 1$ and $\text{parent}(T(n)) = n$ by direct arithmetic. $\square$

Lemma 191. *For even $n \geq 1$ (Proposition 3.1, digit 0),*

$$\langle 3n/2 \rangle = 0 :: \langle n \rangle.$$

Proof. We verify $\text{lsd}(3n/2) = 0$ and $\text{parent}(3n/2) = n$ by the arithmetic of even n: since n is even, $3n/2$ is an integer and $2 \cdot (3n/2) = 3n$, whose remainder modulo 3 is 0 and whose quotient by 3 is n. $\square$

Lemma 192. *For even n (Proposition 3.1, digit 2),*

$$\langle (3n+2)/2 \rangle = 2 :: \langle n \rangle.$$

Proof. Since n is even, $(3n+2)/2$ is a positive integer. We verify $\text{lsd}((3n+2)/2) = 2$ and $\text{parent}((3n+2)/2) = n$ by computing $2 \cdot (3n+2)/2 = 3n+2$, which gives remainder 2 and quotient n upon division by 3. $\square$

Lemma 193. *If $\text{lsd}(n) = 1$, then $\text{parent}(n)$ is odd.*

Proof. Since $\text{lsd}(n) \equiv \text{parent}(n) \pmod 2$ (Lemma 176) and $\text{lsd}(n) = 1$, we have $\text{parent}(n) \bmod 2 = 1$. $\square$

Lemma 194. *If $\text{lsd}(n) = 0$, then $\text{parent}(n)$ is even.*

Proof. Same reasoning: $\text{lsd}(n) = 0$ and the parity congruence give $\text{parent}(n) \bmod 2 = 0$. $\square$

Lemma 195. *If $\text{lsd}(n) = 2$, then $\text{parent}(n)$ is even.*

Proof. Since $2 \bmod 2 = 0$, the parity congruence gives $\text{parent}(n) \bmod 2 = 0$. $\square$

Definition 196. The k-fold iteration of U is defined by $U^0(n) = n$ and $U^{k+1}(n) = U(U^k(n))$.

Definition 197. The *saturated number* $\mathrm{sat}(k) = U^k(0)$ is the largest natural number whose base-3/2 representation has length k (Proposition 3.10 of [2]).

The first few values are $\mathrm{sat}(0) = 0$, $\mathrm{sat}(1) = 1$, $\mathrm{sat}(2) = 2$, $\mathrm{sat}(3) = 4$, $\mathrm{sat}(4) = 7$, $\mathrm{sat}(5) = 11$, $\mathrm{sat}(6) = 17$, $\mathrm{sat}(7) = 26$.

Lemma 198. *For $k \geq 1$, every digit in $\langle \mathrm{sat}(k) \rangle$ is either 1 or 2 (no digit 0 appears). This is Proposition 3.5 of [2].*

Proof. By induction on k. For $k = 0$ the claim is vacuously true (the list is empty). For $k+1$, we have $\mathrm{sat}(k+1) = U(\mathrm{sat}(k))$, so by Lemma 189, $\langle \mathrm{sat}(k+1) \rangle$ is $\langle \mathrm{sat}(k) \rangle$ with either 1 or 2 prepended. The prepended digit is in $\{1, 2\}$ by construction, and the remaining digits are in $\{1, 2\}$ by the induction hypothesis (or are empty when $k = 0$). □

Lemma 199. *For all $k \in \mathbb{N}$ (Proposition 3.6 of [2]),*

$$\langle \mathrm{sat}(k+1) \rangle = \begin{cases} 1 :: \langle \mathrm{sat}(k) \rangle & \text{if } \mathrm{sat}(k) \text{ is odd,} \\ 2 :: \langle \mathrm{sat}(k) \rangle & \text{if } \mathrm{sat}(k) \text{ is even.} \end{cases}$$

Proof. This is a direct application of Lemma 189 to $n = \mathrm{sat}(k)$. □

Definition 200. The *cyclic permutation* $\sigma = (2\ 1\ 0)$ on the digit set $\{0, 1, 2\}$:

$$\sigma(0) = 2, \quad \sigma(1) = 0, \quad \sigma(2) = 1.$$

Formally, $\sigma(d) = (d + 2) \bmod 3$.

Definition 201. Given a digit list (least-significant first), splitAtFirstZero returns the pair (*prefix*, *suffix*) where *prefix* consists of all nonzero digits before the first 0, and *suffix* consists of the digits after the first 0. If no 0 is present, the suffix is empty.

Lemma 202. *For all $n \geq 1$,*

$$\mathrm{lsd}(n+1) = \sigma(\mathrm{lsd}(n)).$$

Proof. Direct arithmetic: $\mathrm{lsd}(n+1) = (2(n+1)) \bmod 3 = (2n+2) \bmod 3 = ((2n) \bmod 3 + 2) \bmod 3 = \sigma(\mathrm{lsd}(n))$. □

Lemma 203. *If $n \geq 1$ and $\mathrm{lsd}(n) = 0$, then $\mathrm{parent}(n+1) = \mathrm{parent}(n)$ (no carry).*

Proof. When $\mathrm{lsd}(n) = 0$, we have $2n \equiv 0 \pmod 3$, so $\lfloor 2(n+1)/3 \rfloor = \lfloor (2n+2)/3 \rfloor = \lfloor 2n/3 \rfloor$ since the remainder goes from 0 to 2 without crossing 3. □

Lemma 204. *If $n \geq 1$ and $\mathrm{lsd}(n) \neq 0$, then $\mathrm{parent}(n+1) = \mathrm{parent}(n) + 1$ (carry propagates).*

Proof. When $\mathrm{lsd}(n) \in \{1, 2\}$, we have $2n \bmod 3 \geq 1$, so adding 2 causes the quotient $\lfloor 2n/3 \rfloor$ to increase by 1. □

Lemma 205. ***(Proposition 2.3, Odometer.)*** *Let $\langle n \rangle = u\,0\,v$ where $v \in \{1, 2\}^*$ (i.e. v is the block of nonzero digits before the first 0, reading from the least-significant end, and u is the part after the 0). Then*

$$\langle n+1 \rangle = u\ 2\ v^\sigma,$$

where v^σ denotes the list obtained by applying σ to each digit of v. In other words, incrementing by one applies σ to all leading nonzero digits (carrying), replaces the first 0 with 2, and leaves the remaining digits unchanged.

Proof. By strong induction on n.

Base case $(n = 0)$*:* $\langle 0 \rangle = []$ splits trivially, and $\langle 1 \rangle = [2]$ (since $\mathrm{lsd}(1) = 2$ and $\mathrm{parent}(1) = 0$). The formula gives $[] \cdot 2 \cdot [] = [2]$.

Inductive step $(n \geq 1)$*:* Write $\langle n \rangle = \mathrm{lsd}(n) :: \langle \mathrm{parent}(n) \rangle$.

Case $\mathrm{lsd}(n) = 0$ *(no carry):* The split gives prefix $= []$ and suffix $= \langle \mathrm{parent}(n) \rangle$. By Lemma 202, $\mathrm{lsd}(n+1) = \sigma(0) = 2$. By Lemma 203, $\mathrm{parent}(n+1) = \mathrm{parent}(n)$. So $\langle n+1 \rangle = 2 :: \langle \mathrm{parent}(n) \rangle$, which matches the formula.

Case $\mathrm{lsd}(n) \neq 0$ *(carry):* The split prefix is $\mathrm{lsd}(n) :: (\text{prefix from rest})$. By Lemma 202, the first digit becomes $\sigma(\mathrm{lsd}(n))$. By Lemma 204, $\mathrm{parent}(n+1) = \mathrm{parent}(n) + 1$. The induction hypothesis (applied to $\mathrm{parent}(n) < n$) handles the remaining digits. □

Definition 206. The *scaled evaluation* of a digit list $[a_0, \dots, a_k]$ is

$$\mathrm{scaledEval}([a_0, \dots, a_k]) = \sum_{i=0}^{k} 2^{k-i} \cdot 3^i \cdot a_i.$$

Recursively: $\mathrm{scaledEval}([]) = 0$ and $\mathrm{scaledEval}(a :: \mathit{rest}) = 2^{|\mathit{rest}|} \cdot a + 3 \cdot \mathrm{scaledEval}(\mathit{rest})$.

Lemma 207. *For all* $n \in \mathbb{N}$*,*

$$\mathrm{scaledEval}(\mathrm{digits}_{3/2}(n)) = 2^{|\,\mathrm{digits}_{3/2}(n)|} \cdot n.$$

Proof. By strong induction on n. For $n = 0$, both sides are 0. For $n \geq 1$, let $L = |\,\mathrm{digits}_{3/2}(\mathrm{parent}(n))|$. Then $|\,\mathrm{digits}_{3/2}(n)| = L + 1$ and:

$$\begin{aligned}
\mathrm{scaledEval}(\langle n \rangle) &= 2^L \cdot \mathrm{lsd}(n) + 3 \cdot \mathrm{scaledEval}(\langle \mathrm{parent}(n) \rangle) \\
&= 2^L \cdot \mathrm{lsd}(n) + 3 \cdot 2^L \cdot \mathrm{parent}(n) && \text{(by IH)} \\
&= 2^L \cdot (\mathrm{lsd}(n) + 3 \cdot \mathrm{parent}(n)) \\
&= 2^L \cdot 2n && \text{(by Lemma 173)} \\
&= 2^{L+1} \cdot n. \square
\end{aligned}$$

Lemma 208. ***(Proposition 4.5 of [2].)*** *Let* $n \geq 1$ *with* $\langle n \rangle = a_k \cdots a_0$ *having* $k+1$ *digits. Then* n *is even if and only if*

$$\sum_{i=0}^{k} 2^{k-i} \cdot 3^i \cdot a_i \equiv 0 \pmod{2^{k+2}}.$$

Equivalently, using the scaled evaluation:

$$n \bmod 2 = 0 \quad \iff \quad \mathrm{scaledEval}(\langle n \rangle) \bmod 2^{L+1} = 0,$$

where $L = |\langle n \rangle|$.

Proof. By Lemma 207, $\mathrm{scaledEval}(\langle n \rangle) = 2^L \cdot n$, so modulo $2^{L+1} = 2^L \cdot 2$:

$$2^L \cdot n \bmod (2^L \cdot 2) = 2^L \cdot (n \bmod 2).$$

The left-hand side is 0 if and only if $n \bmod 2 = 0$ (since $2^L \geq 1$, so $2^L \cdot (n \bmod 2) = 0$ implies $n \bmod 2 = 0$). □

Definition 209. The number of odd values among the first k terms of the U-orbit starting from n_0 is

$$\mathrm{numOddU}(k, n_0) = \#\{j \in \{0, \dots, k-1\} : U^j(n_0) \text{ is odd}\}.$$

Definition 210. The *U-orbit density conjecture* for a fixed starting point n_0 asserts that

$$\lim_{k\to\infty} \frac{\mathrm{numOddU}(k, n_0)}{k} = \frac{1}{2}.$$

Definition 211. The *universal U-orbit density conjecture* asserts that U-orbit density(n_0) holds for every $n_0 \in \mathbb{N}$.

Lemma 212. *Conjecture 4.2 (Definition* **??***) is equivalent to the U-orbit density conjecture for $n_0 = 0$.*

Proof. Since $\mathrm{sat}(j) = U^j(0)$, the count of odd values $\mathrm{numOnesSat}(k)$ is exactly $\mathrm{numOddU}(k, 0)$. The two statements are definitionally equal. □

Chapter 6

Literature

Bibliography

[1] CRANDALL, RICHARD E., *"On the "3x+1" problem."*, Mathematics of computation 32.144 (1978): 1281-1292.

[2] ELIAHOU, SHALOM, AND JEAN-LOUIS VERGER-GAUGRY, *"The number system in rational base* $3/2$ *and the* $3x+1$ *problem."*, Comptes Rendus. Mathématique 363.G4 (2025): 329-336.

[3] EVERETT, C. J., *"Iteration of the number-theoretic function* $f(2n) = n, f(2n+1) = 3n+2$*."*, Advances in Mathematics 25.1 (1977): 42-45.

[4] LAGARIAS, JEFFREY C., "The 3x+ 1 problem: An overview.", in: *The ultimate challenge: The 3x+1 Problem* (2010): 3-29.

[5] ROZIER, OLIVIER AND TERRACOL, CLAUDE, *Paradoxical behavior in Collatz sequences*, arXiv preprint arXiv:2502.00948, 2025

[6] TERRAS, RIHO, *"A stopping time problem on the positive integers."*, Acta Arithmetica 30 (1976): 241-252.