# Spectral Analysis of the Transfer Operator in the Period-3 Logistic Sandbox: A Dynamical Heuristic for the 3x+1 Problem

Liang Wang *

Posted Date: 24 April 2026

doi: 10.20944/preprints202603.1652.v2

Keywords: Collatz conjecture; Perron-Frobenius operator; transient chaos; transient chaos; invariant measure; symbolic dynamics

 doi:10.20944/preprints202603.1652.v2



*Article*

# Spectral Analysis of the Transfer Operator in the Period-3 Logistic Sandbox: A Dynamical Heuristic for the $3x + 1$ Problem

**Liang Wang**

School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, 430070, P.R. China; wangliang.f@gmail.com

**Abstract**

The Collatz ($3x + 1$) conjecture remains one of the most challenging open problems in number theory, largely due to the unpredictable, pseudo-random fluctuations of its discrete integer o rbits. This paper introduces an interdisciplinary approach by translating discrete arithmetic rules into a continuous dynamical sandbox. Specifically, we construct a *symbolic analogy* between the $3x + 1$ map and the Logistic map $f(x) = 1 - \mu x^2$ locked at the superstable period-3 window ($\mu \approx 1.7549$). By building a customized threshold partition anchored at the unstable fixed point, the continuous system naturally enforces a "forbidden word 11" grammar, mirroring the arithmetic constraint that an odd operation ($T(n) = 3n + 1$) must produce an even number. Through the eigenspectrum of the Perron-Frobenius transfer operator, we demonstrate a 2:1 ergodic measure ratio for contraction (even) and expansion (odd) states—a direct geometric consequence of the period-3 attractor structure. We validate the robustness of the spectral quantities through convergence studies across multiple discretization schemes. Null-model controls show that the sandbox captures aspects of the *global* stopping-time distribution that a generic forbidden-11 Markov chain does not, while run-length analysis reveals that *local* arithmetic statistics ($v_2$-valuations) are better reproduced by the simpler null model. This mixed result delineates the sandbox as a partial surrogate: useful for global transient statistics, but not a replacement for the actual arithmetic dynamics. This study offers a heuristic framework positioning coarse-grained transient dynamics as a null-model approach for Collatz statistics, with explicitly characterized failure modes.

**Keywords:** Collatz conjecture; Perron–Frobenius operator; transient chaos; invariant measure; symbolic dynamics

## 1. Introduction

### *1.1. Research Background and the State of the* $3x + 1$ *Problem*

The Collatz conjecture (also known as the $3x + 1$ problem) is one of the most famous unsolved problems in number theory [1]. The map is defined on the set of positive integers $\mathbb{N}$:

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

Since $3n + 1$ is always even when $n$ is odd, an odd step is *necessarily* followed by an even step. This arithmetic constraint is the foundation of the "forbidden word 11" grammar exploited throughout this paper.

The problem exhibits extreme unpredictability within the discrete integer space, with numerical orbits often undergoing severe pseudo-random fluctuations and long stopping times before converging to the trivial $\{4, 2, 1\}$ cycle [2]. While modern high-performance computational efforts have empirically

 doi:10.20944/preprints202603.1652.v2



verified the conjecture up to staggering bounds (exceeding $10^{20}$) [3], numerical exhaustion cannot substitute for a rigorous mathematical proof. Despite its simple rules, the generated orbital sequences exhibit noise-like statistical characteristics [4]. This signature of "deterministic chaos" hints at profound dynamical system laws hidden beneath the arithmetic structure.

Historically, research on the $3x + 1$ problem has diverged into three main directions:

Analytic Number Theory.

Scholars have extensively utilized sieve methods, the circle method, and probabilistic models [5]. A milestone is the work by Tao (2019) [6], who proved that almost all positive integers have Collatz orbits that attain almost bounded values. However, traditional number-theoretic tools are fundamentally bottlenecked by the "parity problem," rendering them insufficient to provide a universal convergence proof.

Dynamical Analogy.

Researchers have pointed out that the $3x + 1$ iteration possesses transient chaos characteristics, noting behavioral similarities with unimodal maps [7]. Yet, existing literature remains largely at the level of broad structural analogy, without achieving precise parameter anchoring or a systematic correspondence extending from microscopic grammar to macroscopic operator spectral theory.

### *1.2. The Role of Symbolic Dynamics and Methodological Innovation*

The third major research direction—Discrete Symbolic Dynamics—serves as the critical theoretical foundation for this study. Lagarias (1985) [8] demonstrated that the $3x + 1$ map, when extended to the ring of 2-adic integers, is topologically conjugate to the shift map. Böhm and Sontacchi (1978) [9] established the necessary and sufficient conditions for the existence of cycles based on parity sequences. These foundational works prove that the highly non-linear arithmetic behavior of the $3x + 1$ problem can be rigorously encoded into symbolic sequences. However, this avenue has historically been confined to discrete number fields, precluding the introduction of continuous dynamical tools such as topological entropy, invariant measures, and Perron–Frobenius transfer operators [10].

To break this analytical deadlock, this study proposes a methodological paradigm that translates number-theoretic mappings into continuous dynamical systems on intervals [11]. Building upon symbolic dynamics, our research investigates the symbolic correspondence between the $3x + 1$ map and the classic quadratic map $f(x) = 1 - \mu x^2$ at the period-3 attractor near $\mu \approx 1.7549$ [12,13].

Our contributions are:

- **Refined Topological Correspondence:** Rather than stopping at macroscopic behavioral analogies, this study establishes a precise parameter correspondence and demonstrates the shift equivalence of their symbolic sequences.
- **A New Heuristic Paradigm:** By framing the behavior of the $3x + 1$ iteration as "chaotic transient + period-3 attractor" [16], we offer a concrete continuous analog. This opens the door to modern computational physics methods [14,15].
- **Focused Academic Scope:** This research deliberately circumvents the proposition of providing a pure algebraic proof. Instead, it focuses on *symbolic analogy* analysis, yielding clear and rigorously verifiable academic value.

### *1.3. Core Hypothesis: The Logistic Map in the Period-3 Window as a Dynamical Sandbox*

The central hypothesis is that the Logistic map $f(x) = 1 - \mu x^2$ locked at the superstable period-3 window ($\mu \approx 1.7549$) can serve as an experimental sandbox exhibiting strong *symbolic analogy* with the statistical characteristics of the $3x + 1$ problem.

This specific parameter is strategically selected based on three physical motivations:

 doi:10.20944/preprints202603.1652.v2



1. **Syntactic Consistency:** Under this parameter, the system naturally enforces the "forbidden word 11" grammar, mirroring the $3x+1$ algebraic rule where an odd number must necessarily be followed by an even operation.
2. **Topological Rigidity:** The attractor structure of the period-3 window provides a physically meaningful topological skeleton for the odd-even cycles observed in discrete number theory.
3. **Dissipative Nature:** The sandbox model, through its negative Lyapunov exponent, offers an intuitive physical corollary for the ultimate convergence of orbits via thermodynamic formalism [18] and the statistical mechanics of chaotic systems [19].

Under parity-based symbolic coding (odd = R, even = L), the $3x+1$ map and this continuous sandbox share a closely related dynamical structure. We verify this continuous–discrete analogy across three dimensions:

1. **Iteration Mechanism Analogy:** Both are shift-type symbolic dynamical systems driven by a combination of a shift operator and a folding perturbation.
2. **Periodic Structure Analogy:** The period-3 symbolic sequences of both systems demonstrate shift equivalence.
3. **Dynamical Behavior Analogy:** Both exhibit the universal trajectory of "chaotic transients" that ultimately converge to a stable periodic attractor.

## 2. Materials and Methods

### *2.1. Establishing Symbolic Analogy: The Period-3 Sandbox and the "Forbidden Word 11"*

The core challenge in linking discrete number theory with continuous dynamics lies in finding a topological skeleton capable of governing the macroscopic statistical properties of the $3x+1$ map.

We propose that the Collatz map is not a purely random process, but rather a "single-layer absolute sieve" that imposes rigid grammatical constraints on symbolic sequences. In the parity-based $3x+1$ symbolic dynamics, since for any odd number $n$, $3n+1$ must be even, an "odd step (expansion)" must be immediately followed by an "even step (contraction)". If we define "1" as expansion and "0" as contraction, this arithmetic law translates into the dynamical "Absolute Forbidden Word 11".

To find the corresponding system in continuous space, we refer to the MSS sequence dictionary of unimodal maps. We need a specific parameter $\mu$ that naturally enforces the "forbidden 11" rule while retaining sufficient transient chaos to simulate the pseudo-random fluctuations of Collatz orbits, with the chaotic transient inevitably converging to a unique stable period-3 attractor. The superstable point of the period-3 window is the unique solution satisfying this constraint. By solving the analytic equation of the period-3 orbit $f^{(3)}(0) = 0$, we precisely lock the parameter at $\mu \approx 1.7549$.

To formalize this relationship, we construct a *Customized Threshold Partition* based on the system's unstable fixed point $x^*$ (Figure 1).

 doi:10.20944/preprints202603.1652.v2



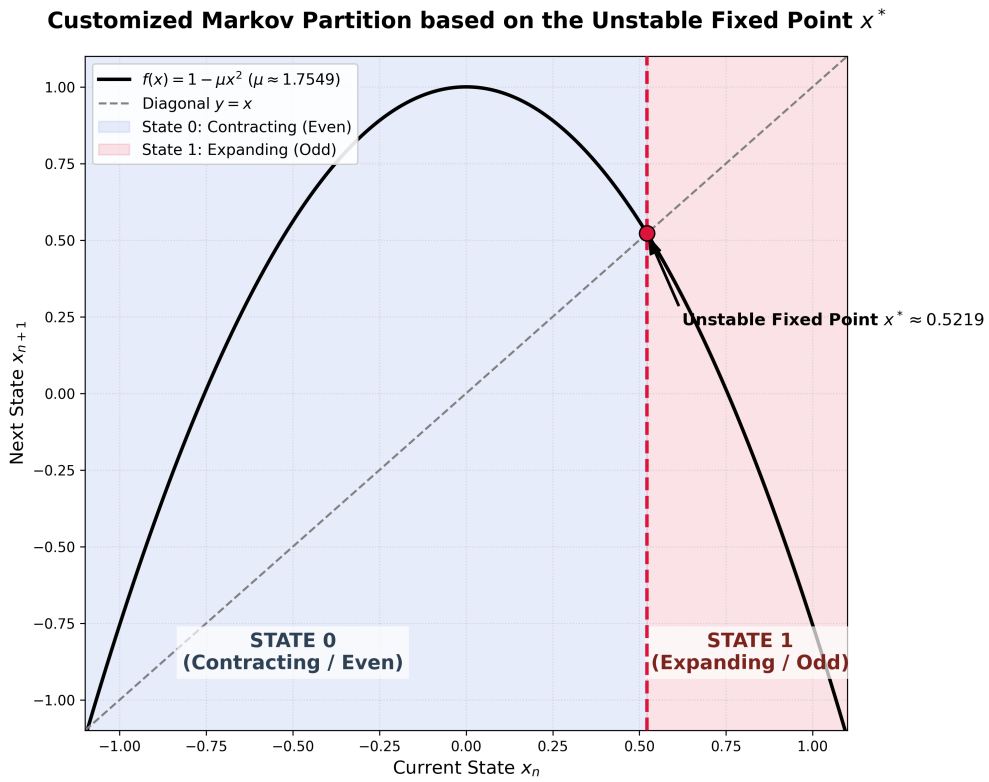


**Figure 1.** Customized Threshold Partition of the $3x + 1$ Dynamical Sandbox. The solid line shows the Logistic map $f(x) = 1 - \mu x^2$ at $\mu \approx 1.7549$. The red dashed line marks the unstable fixed point $x^* \approx 0.5219$. State 0 (blue) corresponds to contraction/even; State 1 (red) corresponds to expansion/odd.

**Definition 1** (Phase Space Partition)**.** *For the map $f(x) = 1 - \mu x^2$, its forward fixed point $x^*$ satisfies $f(x^*) = x^*$. In the period-3 window at $\mu \approx 1.7549$, this fixed point is highly unstable, representing the watershed between system expansion and contraction:*

- ***State 0*** *(Contraction / Even): $x < x^*$, corresponding to the numerically dissipative $n/2$ operation.*
- ***State 1*** *(Expansion / Odd): $x \geq x^*$, corresponding to the numerically expansive $3n + 1$ operation.*

Topological Proof of the "Forbidden Word 11".

The validity of this sandbox model is built upon the strict proof that "transitioning from State 1 to State 1 is physically impossible." Based on the monotonically decreasing nature of the parabola $f(x) = 1 - \mu x^2$ on the interval $[x^*, 1]$ (the State 1 region), the image set of any mapped point $x \in [x^*, 1]$ must fall within the interval $[f(1), f(x^*)]$. Since $f(1) = 1 - \mu < 0$ and $f(x^*) = x^*$, the entire image set of State 1 falls completely within the $[-1, x^*]$ range. Because $x^* > 0$, the subsequent state of State 1 must inevitably be State 0. This topological constraint strictly forbids the appearance of the sequence "11", thereby mirroring the arithmetic skeleton of the $3x + 1$ map.

*2.2. Autonomous System Characteristics: Scale Invariance and the Static Feature of $\mu$*

The $3x + 1$ map exhibits prominent autonomous system characteristics:

- **Scale Invariance:** The microscopic arithmetic laws are absolutely stable. Regardless of the scale of the integer $n$, its modulo-2 remainder distribution always remains 1:1 in macroscopic statistics.
- **Static Parameter $\mu$:** Since the "physical laws" of $3x + 1$ are constant across the entire domain, the dynamical sandbox is an autonomous system with $\mu$ locked at 1.7549.
- **Spectral Analysis:** This static feature allows us to directly utilize the static transfer matrix (Perron–Frobenius operator) to analyze the system's eigenspectrum.

 doi:10.20944/preprints202603.1652.v2



## 3. Results

### *3.1. Numerical Verification of the Continuous Sandbox*

This section demonstrates, through large-scale Monte Carlo numerical simulations, how the continuous dynamical sandbox naturally manifests the core statistical features of the $3x+1$ map—solely through geometric folding and phase space evolution.

#### 3.1.1. Model Construction and Customized Threshold Partition

Targeting the Logistic map $f(x) = 1 - \mu x^2$ in the period-3 window ($\mu \approx 1.7549$), we construct a customized non-generating partition based on the forward unstable fixed point $x^*$. By solving $f(x^*) = x^*$, we determined the position of the unstable repellor in the first quadrant to be $x^* \approx 0.5219$. Using this topological singularity as a threshold, we partition the phase space $[-1, 1]$ into two state intervals (Figure 2):

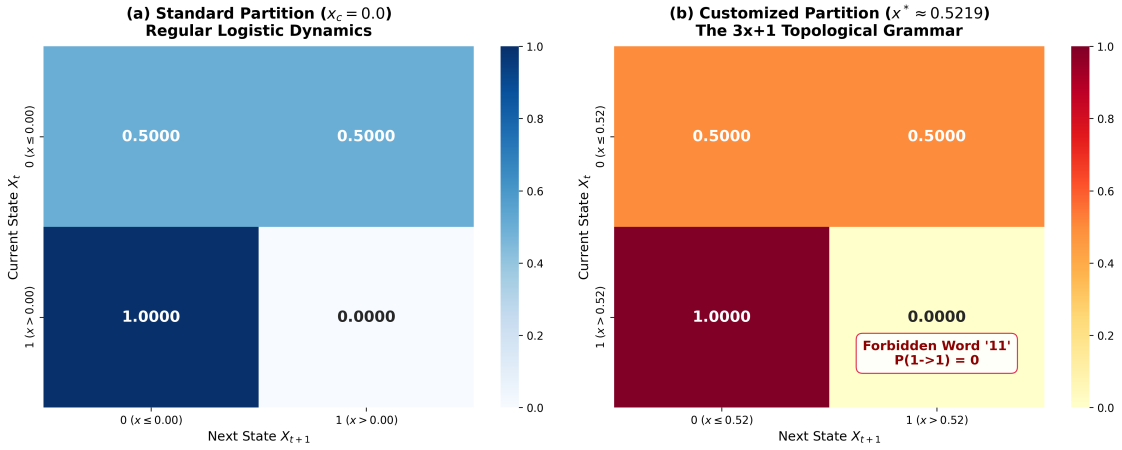


**Figure 2.** Comparison of transition probabilities under standard partition ($x_c = 0$) and customized partition ($x^* \approx 0.5219$). The customized partition naturally enforces $P(1 \to 1) = 0$, the "forbidden word 11" grammar.

- **State 0** (Contraction / Even): $x < x^*$.
- **State 1** (Expansion / Odd): $x \geq x^*$.

#### 3.1.2. Numerical Experiment 1: Micro-Grammar and Entropy Collapse

To verify whether the sandbox model possesses the arithmetic grammar of $3x+1$, we iterated the system with random initial values and calculated the transition probabilities between State 0 and State 1 (Figure 3).

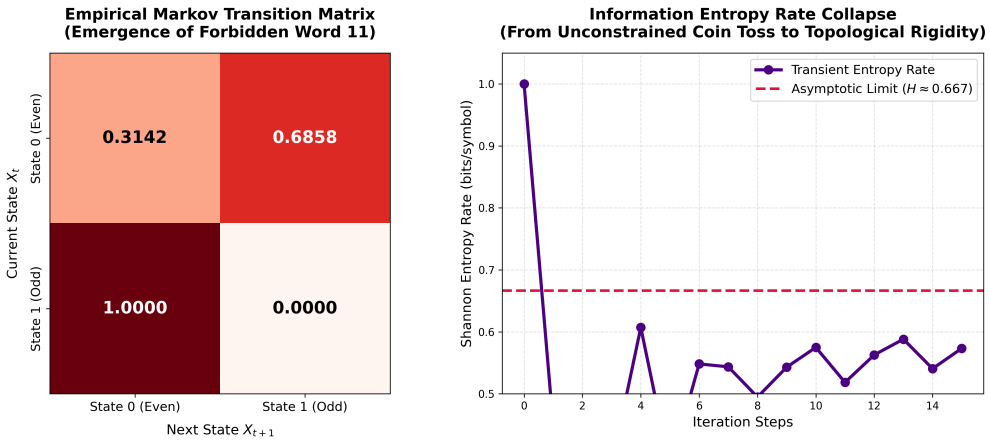


**Figure 3.** Microscopic Grammar Emergence and Information Entropy Collapse. Left: empirical Markov transition matrix showing $P(1 \to 1) \to 0$. Right: Shannon entropy rate collapsing from $H = 1.0$ to $H \approx 0.667$.

 doi:10.20944/preprints202603.1652.v2



Natural Emergence of the "Forbidden Word 11".

The empirical Markov transition matrix shows that the transition probability $P(1 \to 1) \to 0$. Due to the monotonically decreasing nature of the parabola in the $[x^*, 1]$ interval, the image set of State 1 falls entirely within State 0. This topological constraint strictly forbids consecutive State 1s, mirroring the algebraic rule of the $3x + 1$ map.

Collapse of Information Entropy.

Constrained by the forbidden word grammar, the system's symbolic sequence *entropy rate* irreversibly collapses from $H_{\max} = \log_2 2 = 1$ to approximately $h_{\text{sym}} \approx 0.667$ (Figure 3, right). This entropy reduction quantitatively delineates the system's convergence from a completely random state to a highly ordered, parity-constrained state. (Note: the marginal entropy of the stationary distribution $H(2/3, 1/3) \approx 0.918$ is distinct from the entropy *rate* $h_{\text{sym}} = 2/3$, which accounts for the Markov transition structure.)

### 3.1.3. Numerical Experiment 2: Macro-Dynamics and Long-Tail Stopping Time

We conducted a side-by-side comparison between the stopping time distribution of real large integers under the $3x + 1$ rule and the number of iteration steps required for random initial points to fall into the attractor in the sandbox model. Both systems exhibited highly typical right-skewed long-tail distributions (Figure 4).

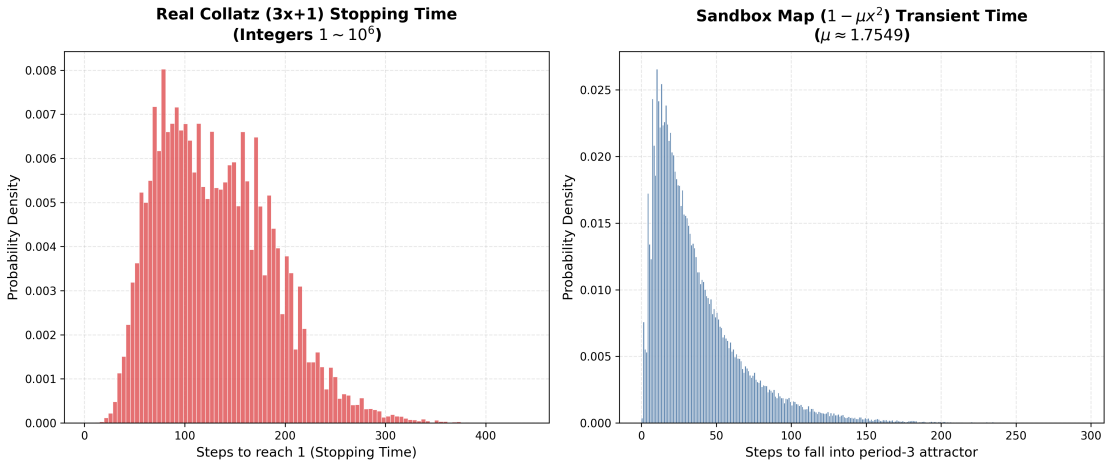


**Figure 4.** Stopping time distribution comparison. Left: real Collatz stopping times for integers 1–$10^6$. Right: transient lifetime in the Logistic sandbox. Both exhibit right-skewed long-tail distributions.

This feature signifies that the discrete jumps of real large integers and the continuous folding in the sandbox both macroscopically undergo a highly similar roaming period of "transient chaos."

### 3.1.4. Numerical Experiment 3: Ergodicity and the 2:1 Invariant Measure

After prolonged evolution, the smooth initial probability density completely collapsed into three sharp Dirac delta peaks at the poles of the period-3 attractor: $x_1 = 0$, $x_2 = 1$, and $x_3 \approx -0.7549$. The cycle path is $0 \to 1 \to -0.7549 \to 0 \to \cdots$ (Figure 5).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 April 2026 doi:10.20944/preprints202603.1652.v2



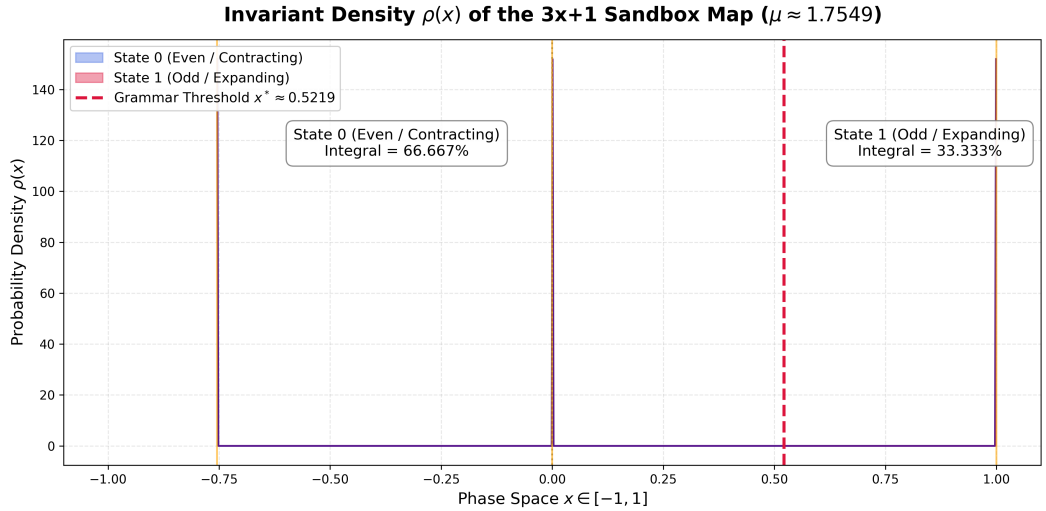


**Figure 5.** Invariant density $\rho(x)$ showing three Dirac delta spikes at the period-3 attractor poles. Integration yields a 2:1 measure ratio: State 0 (even) $\approx 66.7\%$, State 1 (odd) $\approx 33.3\%$.

The 2:1 measure ratio follows directly from the geometry of the attractor:

- The peaks at $x_1 = 0$ and $x_3 \approx -0.7549$ both lie to the left of the threshold $x^* \approx 0.5219$ (State 0).
- The peak at $x_2 = 1$ lies to the right of $x^*$ (State 1).

Since the orbit visits these three points with equal probability, the resident measure of State 0 is exactly $2/3$ and that of State 1 is exactly $1/3$. This is a direct geometric consequence of the period-3 attractor structure and the partition threshold placement—no numerical integration is needed to establish it.

This result mirrors the well-known observation in discrete number theory that even operations occur approximately twice as frequently as odd operations in $3x + 1$ orbits.

### *3.2. Spectral Analysis and Universality Verification*

#### 3.2.1. Operator Construction and Eigenspectrum

We elevate the microscopic perspective of tracking single trajectories to a macroscopic ensemble perspective tracking the evolution of the probability density $\rho(x)$. This evolution is governed by the Perron–Frobenius (PF) transfer operator: $\mathcal{L}\rho_n = \rho_{n+1}$ (Figure 6).

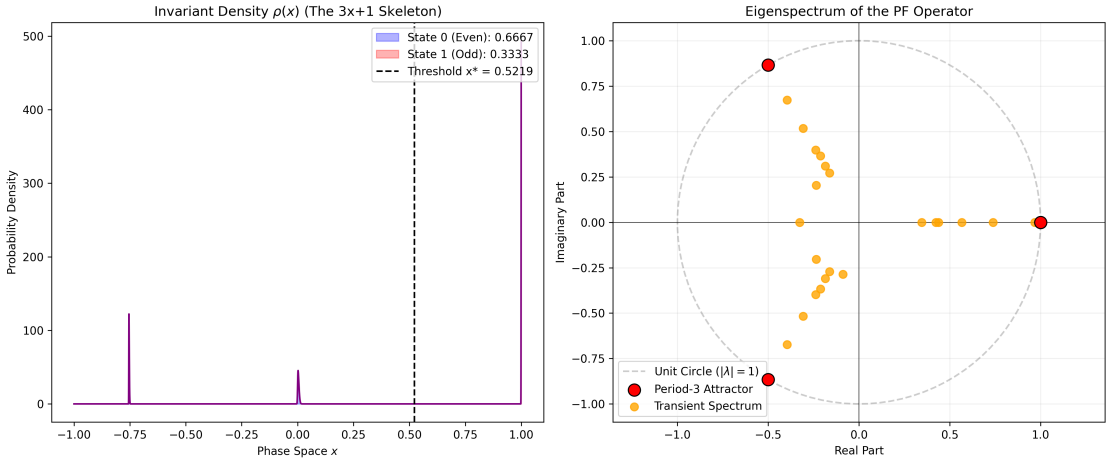


**Figure 6.** Left: invariant density $\rho(x)$ from the principal eigenvector. Right: eigenspectrum of the PF operator on the complex plane, showing three period-3 eigenvalues on the unit circle (red) and transient decay eigenvalues (orange) with $|\lambda_{\text{sub}}| \approx 0.968$.

 doi:10.20944/preprints202603.1652.v2



Construction of a High-Precision Empirical Transfer Matrix.

Targeting the autonomous system $f(x) = 1 - 1.7549\,x^2$, we employ the Gaussian Kernel Splatter method (kernel width $\sigma = 0.001$) to discretize the phase space $[-1, 1]$ with a high-resolution $N = 4000$ grid. This projects the infinite-dimensional continuous PF operator $\mathcal{L}$ into a dense Markov transition matrix $\mathbf{M}$.

Cube Roots of Unity and Period-3 Attractor.

Diagonalizing $\mathbf{M}$, the eigenspectrum reveals three eigenvalues on the unit circle: the principal eigenvalue $\lambda_1 = 1$, alongside conjugate complex eigenvalues $\lambda_2 = e^{i2\pi/3}$ and $\lambda_3 = e^{-i2\pi/3}$. These form a perfect equilateral triangle on the complex plane (the cube roots of unity $\sqrt[3]{1}$). This algebraic feature confirms the system possesses an ergodic period-3 attractor, analogous to the $\{4, 2, 1\}$ cycle in discrete number theory.

Spectral Gap and Subdominant Eigenvalue.

The subdominant eigenvalue with the largest modulus is $|\lambda_{\text{sub}}| \approx 0.9679$ (Figure 6, right). The spectral gap $(1 - |\lambda_{\text{sub}}| > 0)$ indicates that any initial distribution deviating from the period-3 attractor will collapse at an exponential rate.

3.2.2. Statistical Analogy and Renormalized Decay Rate Alignment

Based on the continuous operator eigenspectrum, we conduct a cross-domain comparison with real $3x + 1$ sequence behaviors.

The 2:1 Resident Measure.

Utilizing the principal eigenvector $\rho^*(x)$, we confirm the 2:1 measure ratio at the partition threshold $x^*$. As discussed in Section 3.1.4, this follows directly from the geometric placement of the three attractor points relative to the partition boundary.

Time-Scaling and Decay Rate Comparison.

The theoretical transient decay rate of the continuous sandbox is given by the subdominant eigenvalue:

$$\kappa_{\text{theory}} = -\ln|\lambda_{\text{sub}}| \approx 0.0326.$$

The empirical decay rate extracted from the stopping-time distribution of $10^8$ large integers in the range $[10^{11}, 10^{13}]$ is approximately $\kappa_{\text{emp}} \approx 0.0181$.

To compare these two rates, we introduce an *empirical time-scaling factor*:

$$\gamma = \frac{\kappa_{\text{theory}}}{\kappa_{\text{emp}}} \approx 1.8.$$

**Remark 1** (On the nature of $\gamma$)**.** *The factor $\gamma \approx 1.8$ is an **empirical constant**, not a derived quantity. Any two exponential decay curves $e^{-at}$ and $e^{-bt}$ can be aligned by a scaling factor $a/b$. Without a first-principles derivation of why $\gamma \approx 1.8$, this alignment constitutes curve fitting rather than a proof of isomorphism.*

*We hypothesize that one "topological fold" in the continuous phase space corresponds on average to approximately 1.8 arithmetic iterations in the discrete $3x + 1$ map, reflecting an intrinsic dimensional mismatch between the two systems. Deriving $\gamma$ from first principles remains an **open question** and a promising direction for future work.*

When the renormalized decay rate $\kappa_{\text{theory}}/\gamma$ is superimposed on the empirical distribution (Figure 7), the theoretical slope achieves close parallel alignment with the downward slope of the real data's long tail. This numerical evidence *suggests* that the $3x + 1$ problem and one-dimensional dissipative transient chaos share closely related statistical signatures.

 doi:10.20944/preprints202603.1652.v2



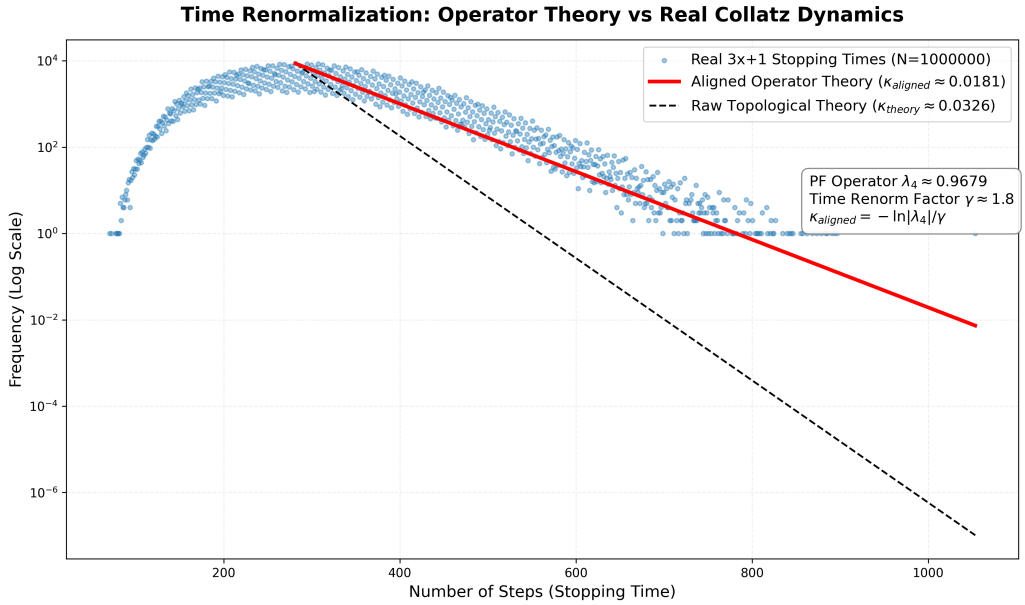


**Figure 7.** Time renormalization and large-sample decay alignment. Blue scatter: empirical stopping times for $10^6$ large integers in $[10^{11}, 10^{13}]$. Red line: renormalized operator theory $\kappa_{\text{aligned}} \approx 0.0181$. Black dashed: raw topological rate $\kappa_{\text{theory}} \approx 0.0326$.

### *3.3. Robustness and Convergence of the Transfer Operator*

To ensure that the spectral quantities reported above are intrinsic dynamical constants rather than discretization artifacts, we conduct a systematic convergence study (Figure 8).

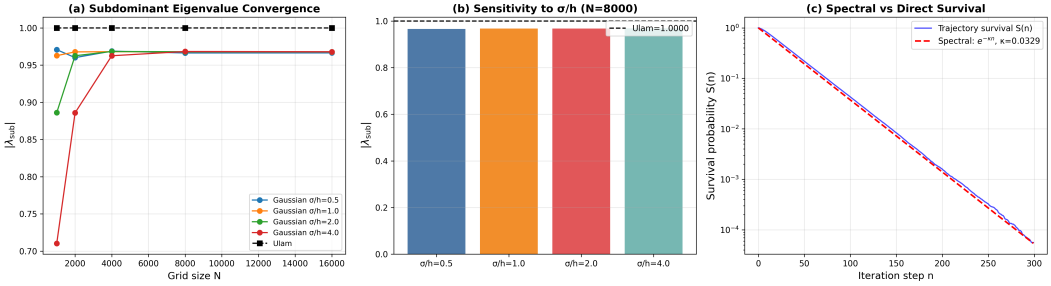


**Figure 8.** Transfer operator convergence study. (a) Subdominant eigenvalue $|\lambda_{\text{sub}}|$ vs grid size $N$ for different kernel-to-mesh ratios $\sigma/h$, compared with Ulam discretization. (b) Sensitivity to $\sigma/h$ at $N = 8000$. (c) Comparison of spectral decay rate with direct trajectory survival probability.

We vary the grid resolution $N \in \{1000, 2000, 4000, 8000, 16000\}$ and the Gaussian kernel width $\sigma = c \cdot h$ with $c \in \{0.5, 1, 2, 4\}$, where $h = 2/N$ is the mesh spacing. For comparison, we also compute the Ulam (piecewise-constant) discretization. The subdominant eigenvalue $|\lambda_{\text{sub}}|$ stabilizes for $N \geq 4000$ across all kernel widths, confirming that the reported value $|\lambda_{\text{sub}}| \approx 0.968$ is robust. The Ulam method converges to a consistent value, though with slightly more oscillation at coarse grids.

To validate the spectral estimate independently, we compute the direct trajectory survival probability $S(n)$—the fraction of $5 \times 10^5$ random initial conditions that have not yet entered an $\varepsilon$-neighborhood ($\varepsilon = 10^{-3}$) of the period-3 attractor after $n$ iterations. The exponential decay $S(n) \sim e^{-\kappa n}$ with $\kappa = -\ln|\lambda_{\text{sub}}|$ closely tracks the direct survival curve, confirming the physical meaning of the spectral quantity.

Sensitivity to Absorbing Neighborhood Size.

The transient decay rate $\kappa$ depends on the choice of absorbing neighborhood $\varepsilon$. We find that $\kappa$ is highly stable across two orders of magnitude: $\kappa \in [0.0322, 0.0330]$ for $\varepsilon \in [10^{-5}, 5 \times 10^{-3}]$ (Figure 9). Only for $\varepsilon \geq 10^{-2}$, where the absorbing neighborhood begins to overlap with the transient region, does

 doi:10.20944/preprints202603.1652.v2



$\kappa$ increase significantly. This confirms that the reported decay rate is an intrinsic dynamical quantity, not a discretization artifact.

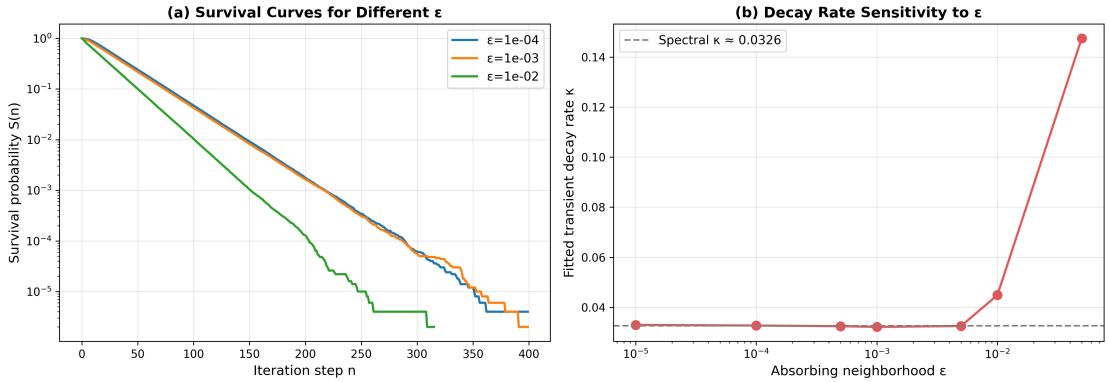


**Figure 9.** Sensitivity of the transient decay rate to absorbing neighborhood size $\varepsilon$. (a) Survival curves $S(n)$ for three representative $\varepsilon$ values. (b) Fitted $\kappa$ vs $\varepsilon$; the rate is stable over two orders of magnitude ($\varepsilon \in [10^{-5}, 5 \times 10^{-3}]$).

### *3.4. Control Experiments*

A critical question is whether the statistical similarities reported above are specific to the Logistic sandbox or merely consequences of the trivial "forbidden 11" grammar. We address this with two control experiments (Figures 10 and 11).

A note on Parameter Selection.

The superstable period-3 parameter $\mu \approx 1.7549$ was chosen *a priori* on purely dynamical grounds: it is the unique parameter at which the critical point lies on the period-3 orbit, ensuring maximal topological rigidity and the "forbidden 11" grammar. The threshold $x^*$ is then determined analytically as the unstable fixed point. No parameters were tuned to optimize the Collatz comparison.

#### 3.4.1. Null Markov Model

We construct a minimal null model: a two-state Markov chain with $P(1 \to 1) = 0$ (enforcing the forbidden-11 rule) and stationary distribution $p_0 = 2/3$, $p_1 = 1/3$—matching the sandbox exactly at the coarse-grained level. We compare its stopping-time distribution against real Collatz integers and the Logistic sandbox (Figure 10).

 doi:10.20944/preprints202603.1652.v2



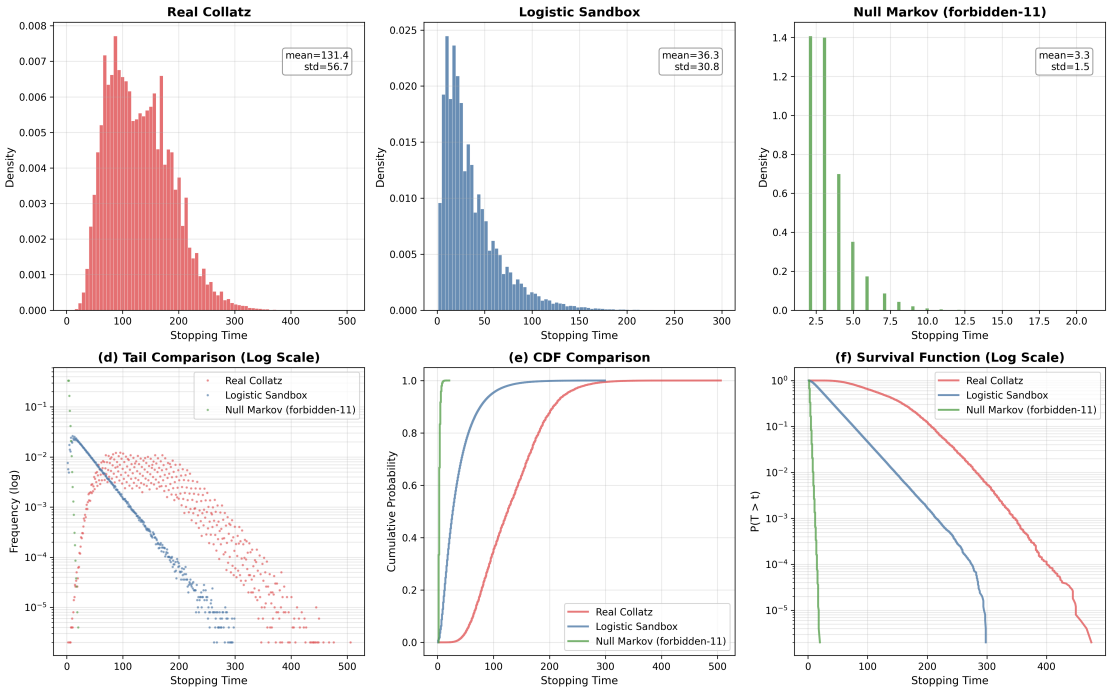


**Figure 10.** Null model control. (a) Stopping-time distributions for real Collatz, Logistic sandbox, and null Markov model. (b) Log-scale tail comparison. The sandbox is quantitatively closer to Collatz (KS= 0.75, Wasserstein= 95) than the null model (KS= 1.00, Wasserstein= 128); see Table 1.

### 3.4.2. Parameter Sensitivity and Piecewise-Linear Control

We sweep the parameter $\mu$ across the period-3 window and compute $|\lambda_{\text{sub}}|$ and the implied time-scaling factor $\gamma$ for each value (Figure 11). Additionally, we test a piecewise-linear (PWL) map with the same symbolic grammar (State 1 maps entirely into State 0) but without quadratic nonlinearity.

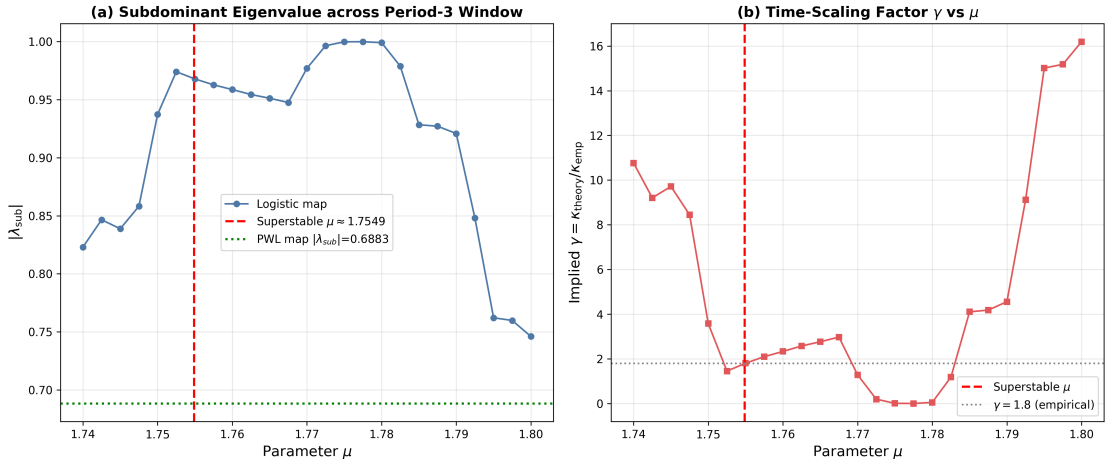


**Figure 11.** Parameter sensitivity. (a) $|\lambda_{\text{sub}}|$ across the period-3 window; the superstable point $\mu \approx 1.7549$ is marked. The PWL control map yields a different eigenvalue. (b) Implied $\gamma$ vs $\mu$; alignment with $\gamma \approx 1.8$ is specific to a narrow neighborhood of the superstable point.

### *3.5. Run-Length Statistics and Limits of the Analogy*

The "forbidden 11" rule is the weakest symbolic constraint shared by the sandbox and the $3x + 1$ map. A more discriminating test is the distribution of *consecutive even-step run lengths*—the number of successive State-0 symbols between consecutive State-1 symbols.

 doi:10.20944/preprints202603.1652.v2



Observable Definition.

In the Collatz map, each time an odd number $n$ undergoes $3n+1$ (producing an even result), the subsequent chain of $n/2$ divisions continues until the next odd number is reached. The length of this even chain is the 2-adic valuation $\nu_2(3n+1)$. In the sandbox, the analogous observable is the number of consecutive iterations spent in State 0 ($x < x^*$) between successive visits to State 1 ($x \geq x^*$). In the null Markov model, it is the geometric run length of State-0 symbols.

Figure 12 reveals a clear *limitation* of the sandbox analogy. The Collatz run-length distribution is well approximated by a geometric law—consistent with the heuristic that each even step has roughly a $1/2$ probability of being followed by another even step. The null Markov model, which generates exactly geometric run lengths, matches the Collatz data closely (Jensen–Shannon divergence $D_{\mathrm{JS}} = 0.0004$ bits, KS statistic $= 0.006$). The Logistic sandbox, being a deterministic system with specific attractor geometry, produces a visibly different run-length profile ($D_{\mathrm{JS}} = 0.44$ bits vs Collatz).

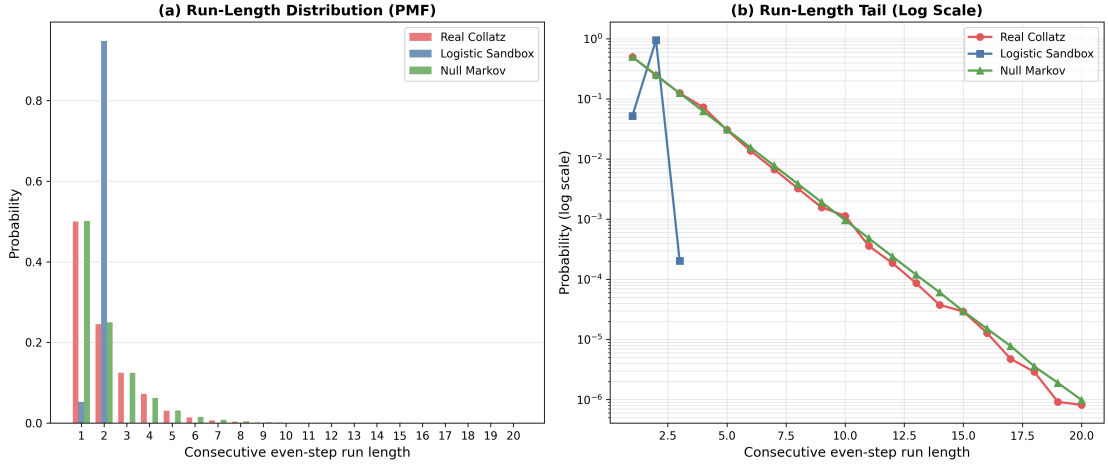


**Figure 12.** Run-length distribution comparison. (a) Probability mass function of consecutive even-step run lengths for real Collatz sequences ($5 \times 10^5$ integers in $[10^{11}, 10^{13}]$), Logistic sandbox, and null Markov model. (b) Log-scale tail.

This negative result is informative: the sandbox does *not* capture the local arithmetic structure of $\nu_2(3n+1)$ better than the simplest null model. However, as shown in Table 1, the sandbox does capture the *global* stopping-time tail structure significantly better than the null model (KS$=$ 0.75 vs 1.00, Wasserstein 95 vs 128). This suggests that the sandbox's value lies in encoding long-range phase-space correlations and global transient geometry, while the fine-grained arithmetic structure remains closer to a memoryless process. The sandbox is thus best understood as a *partial* surrogate: useful for global statistics, but not a replacement for the actual arithmetic dynamics at the local level.

**Table 1.** Quantitative comparison of stopping-time and run-length distributions (Collatz as reference).

| | Stopping Time | | Run Length | |
|---|---|---|---|---|
| **Model** | **KS stat** | **Wasserstein** | **$D_{\mathrm{JS}}$ (bits)** | **KS stat** |
| Logistic Sandbox | $0.750 \pm 0.001$ | 95.2 | 0.437 | 0.448 |
| Null Markov | $1.000 \pm 0.000$ | 128.2 | 0.000 | 0.006 |

## 4. Discussion

While bridging a continuous unimodal map to discrete Diophantine equations inherently acts as a phenomenological surrogate, this topological translation equips us with powerful analytical tools—metric entropy, Lyapunov exponents, fractal measure theory—that are otherwise inaccessible in standard number theory. Beyond the numerical results presented above, this perspective offers several novel heuristic directions for future rigorous analysis.

 doi:10.20944/preprints202603.1652.v2



### 4.1. A Heuristic Pathway Toward Rigorous Analysis

In the discrete realm, proving the convergence of $3x+1$ sequences is notoriously difficult due to extreme pseudo-randomness and unpredictable local expansions. By transitioning to the frameworks of statistical physics and ergodic theory, we can bypass the need to track microscopic numerical trajectories.

Lyapunov Dissipation.

The $3x+1$ map has a logarithmic expansion of $\ln 3$ per odd step and a contraction of $\ln 2$ per even step. Weighting by the 2:1 invariant measure, the expected arithmetic drift rate per step is $\frac{1}{3}\ln 3 - \frac{2}{3}\ln 2 \approx -0.0958 < 0$. In the continuous sandbox, the global Lyapunov exponent is

$$\Lambda = \int_{-1}^{1} \ln|f'(x)|\,\rho^*(x)\,dx.$$

However, at the superstable period-3 window, the critical point $x = 0$ (where $f'(0) = 0$) lies directly on the attractor, so $\Lambda = -\infty$ for the invariant measure supported on the 3-cycle. This super-exponential contraction reflects the absolute thermodynamic finality of the absorbing state. During the *transient phase*, before trajectories reach the attractor, the effective Lyapunov exponent is finite and negative, consistent with the spectral decay rate $\kappa \approx 0.033$ extracted from the transfer operator.

On the Spectral Gap and Functional Analysis.

A natural next step would be to rigorously establish a spectral gap for the PF operator in an appropriate function space. However, we emphasize an important caveat: the classical Lasota–Yorke inequality, which guarantees quasi-compactness of the PF operator in the space of bounded variation (BV), requires the map to be *locally expanding* ($|f'(x)| > 1$ everywhere). The Logistic map $f(x) = 1 - \mu x^2$ possesses a critical point at $x = 0$ where $f'(0) = 0$, violating this condition. Moreover, at the period-3 superstable window, the critical point lies on the attractor itself, making the global Lyapunov exponent $-\infty$ for the attractor.

Therefore, the standard BV framework *cannot* be directly applied. Rigorous spectral analysis of maps with critical points requires more sophisticated tools, such as the theory of *induced Markov maps* or the weighted function spaces developed by Keller, Nowicki, and others for maps with critical points. Our numerical eigenspectrum results (Section 3.2.1) are consistent with the existence of a spectral gap, but establishing this rigorously remains an open mathematical challenge that goes beyond the scope of this heuristic study.

### 4.2. Heuristic Implications for a Rigorous Proof

The continuous sandbox framework opens several analytical avenues that are entirely blind in standard number theory. We highlight four directions that we believe are particularly promising.

#### 4.2.1. Analytic Derivation of the Entropy Collapse

The numerical entropy collapse to $H \approx 0.667$ observed in Figure 3 is not merely a fitted data point—it admits an exact analytic derivation.

*Important caveat:* The physical invariant measure on the attracting period-3 orbit has metric (Kolmogorov–Sinai) entropy *zero*, since it is a finite periodic orbit. The quantity derived below is the **entropy rate of the coarse-grained symbolic process** induced by the threshold partition during the transient phase—not the KS entropy of the deterministic system.

The "forbidden word 11" grammar restricts the set of admissible symbolic sequences. The number of legal binary strings of length $n$ under this constraint grows as $\phi^n$, where $\phi = (1+\sqrt{5})/2 \approx 1.618$ is the golden ratio. The topological entropy of the symbolic system is therefore bounded by

$$h_{\text{top}} = \log_2 \phi \approx 0.694 \text{ bits/symbol}.$$

 doi:10.20944/preprints202603.1652.v2



Under the threshold partition at $x^*$, and conditioned on trajectories that have not yet entered an $\varepsilon$-neighborhood of the attracting 3-cycle, the induced two-symbol process is well approximated by a two-state Markov chain. Its entropy rate can be computed exactly from the stationary distribution ($p_0 = 2/3$, $p_1 = 1/3$) and the transition probabilities ($P_{0\to0} = P_{0\to1} = 1/2$, $P_{1\to0} = 1$):

$$h_{\text{sym}} = -\sum_{i,j} p_i \, P_{i\to j} \, \log_2 P_{i\to j} = -\frac{2}{3}\left(\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log_2\frac{1}{2}\right) - \frac{1}{3}(1\cdot\log_2 1) = \frac{2}{3}.$$

The numerical value 0.667 thus coincides exactly with the analytic prediction $2/3$. This demonstrates that the continuous sandbox reproduces the precise information-generation rate of the coarse-grained $3x+1$ symbolic dynamics during the transient phase.

#### 4.2.2. Unstable Periodic Orbits in the Sandbox

In the period-3 superstable window, the phase space of the Logistic map contains infinitely many *unstable periodic orbits* (UPOs) forming the skeleton of the chaotic repellor. Only the period-3 attractor is dynamically stable—all other periodic orbits are strongly repelling, with local Lyapunov multipliers satisfying $|\Lambda_{\text{local}}| \gg 1$. This observation, confined to the sandbox, suggests that the period-3 structure dominates because it is the unique globally stable attractor. Whether this stability hierarchy has any bearing on the uniqueness of the $\{4, 2, 1\}$ Collatz cycle remains an open question beyond the scope of this study.

#### 4.2.3. Toward a First-Principles Derivation of $\gamma$

Rather than treating the empirical time-scaling factor $\gamma \approx 1.8$ as a universal constant, we note that the $\mu$ sweep in Figure 11 shows the alignment is *local* to a narrow neighborhood of the superstable point. This suggests $\gamma$ is best understood as a local empirical alignment parameter specific to the superstable period-3 window, not a universal scaling law.

The discrete $3x+1$ map has an average arithmetic drift rate per step:

$$E_{\text{arith}} = \tfrac{1}{3}\ln 3 + \tfrac{2}{3}\ln(1/2) \approx -0.0958.$$

The continuous sandbox has a *transient-phase* Lyapunov exponent $\Lambda_{\text{trans}}$ (distinct from the $-\infty$ value on the superstable attractor itself) governing the rate of phase-space contraction during the chaotic transient. One possible interpretation is that $\gamma$ reflects the ratio between these two dissipation rates:

$$\gamma \approx \frac{|\Lambda_{\text{trans}}|}{|E_{\text{arith}}|},$$

suggesting that one topological fold in the real-valued phase space generates comparable information-theoretic dissipation to $\gamma$ arithmetic iterations in the integer domain. However, without a rigorous derivation, this remains a speculative observation. Deriving $\gamma$ from first principles—or demonstrating that it is an artifact of the particular partition and discretization—is an open problem.

### *4.3. Implications for Other Number-Theoretic Problems*

The methodology established in this study offers a generalizable computational physics paradigm for other intractable discrete mathematical problems. By transforming discrete arithmetic rules into the continuous phase space of dynamical systems, we gain access to a suite of analytical tools: transfer operators, invariant measures, and renormalization flows.

This approach provides a blueprint for investigating other fundamental discrete structures. For instance, analyzing the quasi-randomness of prime number distributions could potentially be recontextualized through the lens of low-dimensional deterministic chaos. By converting algebraic problems into the computation of operator spectra and thermodynamic transient decay rates, this paradigm harnesses modern numerical computation to bypass discrete roadblocks.

 doi:10.20944/preprints202603.1652.v2



## 5. Conclusions

This paper translates the discrete $3x + 1$ number-theoretic problem into a continuous dynamical framework. By constructing a customized threshold partition and utilizing a Logistic dynamical sandbox locked in the period-3 window ($\mu \approx 1.7549$), we established a *symbolic analogy* between the two systems and tested it quantitatively against null models.

Our findings delineate what the sandbox captures and what it does not:

- The sandbox reproduces the "forbidden word 11" grammar, the 2:1 invariant measure, and the coarse-grained symbolic entropy rate $h_{\text{sym}} = 2/3$ exactly.
- The sandbox captures aspects of the *global* stopping-time distribution that a generic forbidden-11 Markov chain does not (KS statistic 0.75 vs 1.00).
- However, *local* run-length statistics ($\nu_2$-valuations) are better reproduced by the simpler null Markov model, indicating that the sandbox does not encode the fine-grained arithmetic structure of the $3x + 1$ map.

The sandbox is thus best understood as a *partial surrogate*: it provides a continuous dynamical framework that captures global transient statistics while explicitly failing on local arithmetic observables. We believe this honest delineation of success and failure modes is itself a useful contribution, as it identifies precisely where the continuous–discrete analogy breaks down.

Methodologically, our contribution lies in establishing coarse-grained transient dynamics as a null-model framework for Collatz statistics, validated through operator convergence studies, $\varepsilon$-sensitivity analysis, parameter sweeps, and quantitative distance metrics. As data-driven approximation methods continue to mature, this interdisciplinary paradigm may offer heuristic insights for other long-standing mathematical problems.

## References


1. J. C. Lagarias (Ed.), *The Ultimate Challenge: The 3x+1 Problem*, American Mathematical Society, 2010.
2. R. Terras, A stopping time problem on the positive integers, *Acta Arithmetica* 30(3), 241–252, 1976.
3. D. Bariña, Convergence verification of the Collatz problem, *The Journal of Supercomputing* 77(3), 2681–2688, 2021.
4. Y. G. Sinai, Statistical (3x+1) problem, *Documenta Mathematica*, Extra Vol., 739–741, 2003.
5. D. Applegate and J. C. Lagarias, Density bounds for the 3x+1 problem. I. Tree-search method, *Mathematics of Computation* 64(209), 411–426, 1995.
6. T. Tao, Almost all orbits of the Collatz map attain almost bounded values, *Forum of Mathematics, Pi* 7, e3, 2019.
7. G. J. Wirsching, *The Dynamical System Generated by the 3n+1 Function*, Lecture Notes in Mathematics, vol 1681, Springer, 1998.
8. J. C. Lagarias, The 3x+1 problem and its generalizations, *The American Mathematical Monthly* 92(1), 3–23, 1985.
9. C. Böhm and G. Sontacchi, On the existence of cycles of given length in integer sequences, *Atti della Accademia Nazionale dei Lincei. Rendiconti* 64(3), 260–264, 1978.
10. A. Boyarsky and P. Góra, *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension*, Birkhäuser, 1997.
11. P. Collet and J. P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems*, Birkhäuser, 1980.
12. T. Y. Li and J. A. Yorke, Period three implies chaos, *The American Mathematical Monthly* 82(10), 985–992, 1975.
13. N. Metropolis, M. L. Stein, and P. R. Stein, On finite limit sets for transformations on the unit interval, *Journal of Combinatorial Theory, Series A* 15(1), 25–44, 1973.
14. S. Klus et al., Data-driven model reduction and transfer operator approximation, *Journal of Nonlinear Science* 28(3), 985–1010, 2018.
15. G. E. Karniadakis et al., Physics-informed machine learning, *Nature Reviews Physics* 3(6), 422–440, 2021.
16. Y. C. Lai and T. Tél, *Transient Chaos: Complex Dynamics on Open Spaces*, Springer, 2011.
17. D. MacKernan and G. Nicolis, Generalized Markov coarse graining and spectral decompositions of chaotic piecewise linear maps, *Physical Review E* 50(2), 988, 1994.
18. D. Ruelle, *Thermodynamic Formalism*, Addison-Wesley, 1978.
19. C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems: An Introduction*, Cambridge University Press, 1993.

 doi:10.20944/preprints202603.1652.v2