The Collatz Conjecture and the Spectral Calculus for Arithmetic Dynamics

James Hateley [*]

Article

The Collatz Conjecture and the Spectral Calculus for Arithmetic Dynamics

James C Hateley [ID]

Independent Researcher; hateleyjc@gmail.com

Abstract: We develop a complete operator–theoretic and spectral framework for the Collatz map by analyzing its backward transfer operator on weighted Banach spaces of arithmetic functions. The associated Dirichlet transforms form a holomorphic family that isolates a zeta–type pole at $s = 1$, while on a finer multiscale space adapted to the dyadic–triadic geometry of the Collatz preimage tree we establish a two–norm Lasota–Yorke inequality with an explicit contraction constant $\lambda < 1$, yielding quasi–compactness, a spectral gap, and a Perron–Frobenius theorem in which $\rho(P) = 1$, the eigenvalue 1 is algebraically and geometrically simple, no other spectrum meets the unit circle, and the unique invariant density is strictly positive. The fixed–point relation $Ph = h$ is converted into an exact multiscale recursion for the block averages c_j, revealing a rigid second–order coupling with exponentially small error terms and asymptotic profile $c_j \sim C\,6^{-j}$. This spectral classification forces every weak* limit of the Cesàro averages derived from any hypothetical infinite forward orbit to be either 0 or a scalar multiple of the Perron–Frobenius functional, with convergence to 0 occurring precisely under the *Block–Escape Property*. Since the forward map satisfies an unconditional exponential upper bound, whereas Block–Escape combined with linear block growth along a subsequence would impose an incompatible exponential lower bound, all analytic and spectral components needed for such a contradiction are complete, reducing the Collatz conjecture to excluding infinite orbits exhibiting Block–Escape without the supercritical linear block growth prohibited by the spectral theory.

Keywords: Collatz conjecture; transfer operators; spectral gap; quasi–compactness; Perron–Frobenius theory; Dirichlet transforms; arithmetic dynamics; multiscale analysis; block structure; Lasota–Yorke inequality

---

1. Introduction

The Collatz conjecture asserts that every positive integer n eventually reaches the trivial $1 \to 4 \to 2 \to 1$ cycle under repeated application of

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

Equivalently, every forward orbit $\mathcal{O}^+(n) = \{T^k(n) : k \geq 0\}$ is conjectured to terminate in this trivial cycle. Despite its elementary definition, the iteration exhibits highly irregular behavior, with long episodes of apparent expansion intertwined with rapid collapses. This mixture of multiplicative and divisive effects has motivated extensive probabilistic, analytic, and computational investigation for more than eight decades. Foundational work of Terras [1,2] established early density results and stopping-time estimates, while the surveys of Lagarias [3,4] synthesized a broad range of heuristic and structural approaches. Later analytic contributions, including those of Meinardus [5] and Applegate–Lagarias [6], developed refined density bounds and asymptotic models for the distribution of orbit values. Nevertheless, the global termination problem remains open, and the fine-scale structure of Collatz trajectories continues to resist direct combinatorial analysis.

The goal of this paper is to place the Collatz conjecture within an analytic and operator–theoretic framework and to formulate a system of dynamical criteria whose interrelations isolate the precise mechanisms that must be ruled out in any proof of the conjecture. Rather than studying T directly, we analyze its inverse dynamics through the backward transfer operator

$$(Pf)(n) := \sum_{m:T(m)=n} \frac{f(m)}{m},\tag{2}$$

acting on arithmetic functions $f : \mathbb{N} \to \mathbb{C}$. Such operators arise naturally in statistical mechanics and dynamical systems [7,8] and have recently been applied to $3x+1$–type maps in analytic and functional–analytic settings [9,10]. For the Collatz map (1), each n has an even preimage $2n$ and, when $n \equiv 4 \pmod 6$, an odd preimage $(n-1)/3$, giving

$$(Pf)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4 \pmod 6\}} \frac{f((n-1)/3)}{(n-1)/3}.\tag{3}$$

The weights $1/m$ normalize the operator so that P acts as a logarithmically mass–preserving averaging operator on nonnegative ℓ^1 sequences, reflecting the contraction inherent in the preimage structure of T.

On weighted spaces ℓ^1_σ, with norm $\|f\|_\sigma = \sum_{n \geq 1} |f(n)|\, n^{-\sigma}$, the Dirichlet transform

$$Df(s) = \sum_{n \geq 1} \frac{f(n)}{n^s},\tag{4}$$

intertwines P with analytic continuation in the half-plane $\Re(s) > \sigma$. Uniform bounds on P^k yield exponential envelopes for $D(P^k f)(s)$ and translate into meromorphic continuations of the associated Collatz–Dirichlet series. In this picture, the spectral radius of P on ℓ^1_σ encodes the global weighted expansion rate of inverse branches and determines the analytic location of dominant singularities. However, the ℓ^1 theory does not capture the multiscale fluctuations central to Collatz dynamics, such as the dyadic–triadic block structure that organizes the Collatz preimage tree.

To resolve these finer features, we refine the analytic setting to a multiscale Banach space $B_{\text{tree},\sigma}$ built from dyadic–triadic block averages and oscillation seminorms mirroring the hierarchical geometry of the Collatz preimage tree. On this space the operator P satisfies a two-norm Lasota–Yorke inequality

$$[Pf]_{\text{tree},\sigma} \leq \lambda_{\text{LY}} [f]_{\text{tree},\sigma} + C\,\|f\|_\sigma, \qquad 0 < \lambda_{\text{LY}} < 1,$$

placing the dynamics squarely within the Ionescu–Tulcea–Marinescu and Hennion theories for quasi–compact transfer operators [11,12]. The explicit verification of this inequality, including the strict contraction in the odd branch and the uniform blockwise control, is carried out in Sections 5.

These spectral estimates give rise to a set of dynamical criteria describing block escape, orbitwise averaging, valuation windows, and backward-operator invariants. Each furnishes a logically distinct formulation of the Collatz problem, and the implications among them clarify what phenomena must be precluded in order to establish global termination. The next theorem records the main logical relations among these statements and their connection to the Strong Collatz Conjecture.

Theorem 1.1 (Dynamical Forms Connected to the Strong Collatz Conjecture). *Let $I_j = [2 \cdot 3^j,\, 2 \cdot 3^{j+1})$ denote the jth Collatz block, sand for $n \in \mathbb{N}$ let $j(n)$ be the unique index with $n \in I_{j(n)}$. Consider the following statements:*

(1) Strong Collatz Conjecture. *Every forward orbit is finite and the only cycle is the trivial $1 \to 4 \to 2 \to 1$ cycle.*

Remark 1.2. *This is the classical problem stated by Collatz, also called the $3x + 1$ conjecture. Surveys by Lagarias [3,4] and Wirsching [13] provide comprehensive overviews of known results and approaches.*

(2) No infinite block escape. *For every forward orbit (n_k),*

$$\sup_{k \geq 0} j(n_k) < \infty.$$

Remark 1.3. *With the hypothesis, there is only the trivial orbit this is equivalent to (1) by definition of $j(n)$. The negation (existence of an orbit with unbounded $j(n_k)$) corresponds to a divergent orbit. Korec [14] showed that divergence implies visiting arbitrarily large numbers, but not specifically structured by powers of 3. The block structure $I_j = [2 \cdot 3^j, 2 \cdot 3^{j+1})$ appears in Wirsching's functional-analytic approach [13] as a natural partition for studying growth rates.*

(3) Orbit–Averaging Conjecture (Conjecture 6.3). *Every infinite orbit produces a nonzero P^*–invariant linear functional supported entirely on that orbit.*

Remark 1.4. *This is an orbit-wise version of invariant measure constructions. Lagarias [3] constructed a 2-adic invariant measure for the $3x + 1$ map, while Tao [15] used almost-periodicity and limit functionals in his "almost all" result. The notion of extracting invariant functionals from single orbits relates to Birkhoff averages and unique ergodicity along orbits [16].*

(4) Block–Orbit–Averaging (Conjecture 6.4). *Every infinite orbit spends a positive proportion of time inside a finite union of low blocks $\bigcup_{j \leq J} I_j$ for some J.*

Remark 1.5. *Stronger than Tao's result [15] that almost all orbits eventually go below $f(N)$ for some slow-growing f. This claims positive density returns for* every *infinite orbit. Similar "recurrence to compact sets" appears in dynamical systems as a consequence of invariant measures [17]. For Collatz, it would follow from sufficiently fast growth in high blocks conflicting with known exponential bounds $T^k(n_0) \leq C2^k$ [18].*

(5) Block–Escape Implies Supercritical Linear Block Growth (Conjecture 6.10). *For every infinite orbit satisfying the Block–Escape Property, there exist $\alpha > \dfrac{\log 2}{\log 6}$ and a subsequence $k_\ell \to \infty$ such that*

$$j(n_{k_\ell}) \geq \alpha \, k_\ell.$$

Remark 1.6. *The constant $\alpha > \log 2 / \log 6 \approx 0.3868$ is a minimal expansion factor. Similar linear-in-k lower bounds appear in heuristic models where $v_2(3n + 1)$ averages less than $\log_2 3$ [19]. In branching random walk models [20], the growth rate per step is $\log(3/2^a)$ where a is typical valuation; requiring $\alpha > \log 2 / \log 6$ corresponds to $a < \log_2 3 - \varepsilon$ on average, i.e., deficit windows as in Theorem 6.13.*

(6) Weak Non–Retreat Principle. (Thm 7.20). *For every infinite orbit (n_k) there exist constants $c > 0$, $C_0 \geq 0$, and $t_{\max} \in \mathbb{N}$ with the following property: once $j(n_k)$ exceeds a threshold J_0, then for all sufficiently large such k there is some $1 \leq t \leq t_{\max}$ with*

$$j(n_{k+t}) \geq j(n_k) + c\,t - C_0.$$

Remark 1.7. *This is a finite-horizon forward drift condition. Similar "advance in bounded time" ideas appear in stopping time analyses [1,21], but those guarantee eventual decrease, not forward increase. The closest is Krasikov's lower bounds on numbers satisfying $T^k(n) > n$ [18], but not orbit-wise. Sinai's random walk model [22] suggests that when $\log n_k$ is large, the expected drift per step is positive, but (6) demands a deterministic, uniform forward jump within bounded $t_{\max}$.*

(7) Orbitwise Discrepancy Vanishing (Conjecture 7.7). *Which is later to be shown true by 7.28 Every weak* limit Λ of Cesàro averages along an infinite orbit is P^*–invariant.*

Remark 1.8. *This is a version of the von Neumann mean ergodic theorem for a single orbit, assuming the limit exists. Related to unique ergodicity along orbits [16]. For $3x + 1$, it posits asymptotic regularity of time averages, akin to the existence of rotation numbers for circle maps. Tao's argument [15] uses such almost-periodicity for almost all orbits.*

(8) Residue–Graph Nontrapping and Cycle Obstruction. *These are Theorems 7.17 and 7.16 For some sufficiently large valuation cutoff L_0 and window bound $1 \leq t \leq t_{\max}$, every nonperiodic accelerated odd Collatz orbit escapes both the dangerous residue set $V_{\mathrm{danger}}(L_0)$ and the high–valuation set $V_{\geq L_0+1}$. Equivalently, the finite directed residue graphs induced by valuation patterns contain no directed cycles compatible with the odd Collatz map, and no infinite orbit can remain trapped in or oscillate between the corresponding residue classes.*

Remark 1.9. *Statement (8) places the Collatz problem into a finite combinatorial framework by encoding the accelerated odd map on residue graphs and requiring that no infinite trajectory can remain trapped in either the dangerous or high–valuation subgraphs. This approach extends several earlier graph–based and congruence–based analyses. Applegate and Lagarias [20] performed exhaustive computations of the accelerated map modulo 2^N (for $N \leq 30$), finding no nontrivial cycles and thereby anticipating the cycle–obstruction aspect of (8). Eliahou's modular constraints on possible cycles [23] likewise translate into forbidden subgraphs in this residue framework. Kontorovich and Lagarias [19] modeled the dynamics via Markov chains on residue classes determined by 2–adic valuations, and Sinai's logarithmic random–walk formulation [22] analyzed when orbits may appear trapped in specific congruence sets.*

The contribution of (8) is to integrate these ideas into a unified dynamical criterion: it couples the residue graph to the valuation window formalism, distinguishes dangerous and high–valuation regions via $V_{\mathrm{danger}}(L_0)$ and $V_{\geq L_0+1}$, and asserts that no infinite orbit can persist in, or oscillate between, these regions. This strengthened nontrapping principle links finite residue obstructions directly to the weak non–retreat mechanism used later in the paper.

The following implications hold:

(i) (3) + (8) $\implies$ (1) *by Proposition 7.22*

(ii) (4) $\implies$ (3) *by Proposition 7.23*

(iii) (5) $\implies$ (4) *by Proposition 7.24*

(iv) (6) $\implies$ (5) *by Proposition 7.27*

(v) (8) $\implies$ (6) *by Theorem 7.20.*

(vi) (8) = *Theorem 7.17 and Theorem 7.16.*

The resulting chain shows the Collatz Conjecture is proven.

$$(8) \implies (6) \implies (5) \implies (4) \implies (3) \quad and \quad (3) + (8) \implies (1)$$

Other implications:

(3) $\implies$ (2) *by Proposition 7.21*

(3) $\implies$ (1) *by Theorems 5.28, 5.32, 5.41*

(7) $\implies$ (3) *by Lemmas 7.1, 7.4*

(7) = *Proposition 7.28*

Remark 1.10. *Several statements in this paper initially appear in conjectural form. This is deliberate. Later in the paper, when a dynamical or spectral principle first arises, we present it as a conjecture and defer the development of the required analytic tools. Subsequent sections supply the necessary machinery, and the*

conjectures are later confirmed as theorems or propositions—most notably in Section 7, where the essential conjectural statements are fully resolved. We retain their original presentation as conjectures to preserve the logical flow of the exposition and to highlight the structural ideas as they emerge.

The remainder of the paper develops the analytic and dynamical tools that feed into Theorem 1.1. Section 2 introduces the weighted ℓ^1_σ spaces and their Dirichlet transforms, establishing the basic analytic framework used throughout. Section 3 defines the backward transfer operator P associated with the odd branch and derives its analytic representation. Section 4 constructs the multiscale space $B_{\text{tree},\sigma}$ adapted to the Collatz preimage tree and proves the Lasota–Yorke inequalities that control the action of P on this space. Section 5 converts the Lasota–Yorke contraction into a concrete spectral statement, proving that invariant densities obey an explicit block-level recursion with exponentially small error. This leads to a complete Perron–Frobenius description of P and to a certified spectral gap, establishing quasi–compactness and a spectral gap. Section 6 finalizes the spectral results and develops the valuation and residue tools required to analyze finite windows of odd transitions and to formulate the residue-graph obstructions that underlie the weak non–retreat principle. Section 6.5 introduces the valuation and residue-class tools used to analyze short odd windows, defining block indices, dangerous valuation patterns, and the finite residue graphs whose transitions constrain forward dynamics. These structures support the Block–Escape Property and the weak non–retreat principle central to the dynamical implications of the paper. Section 7 links the residue–graph analysis with precise forward–dynamical estimates, showing that nontrapping of accelerated odd orbits produces quantitative weak non–retreat, linear block growth, and ultimately orbitwise invariant functionals. When coupled with the spectral structure of the backward operator, this leads to a contradiction for any infinite orbit and therefore completes the proof of the Strong Collatz Conjecture. Finally, Section 8 places the spectral and dynamical framework in a broader arithmetic context and outlines potential extensions to other discrete maps.

2. Analytic Foundations for the Transfer Operator

The analysis begins with a careful description of the function spaces, Dirichlet transforms, and basic structural features of the Collatz map that underlie the spectral study of the backward operator P. Throughout we work with complex-valued arithmetic functions $f : \mathbb{N} \to \mathbb{C}$. We start with a simple unbounded estimate.

Lemma 2.1 (Coarse k-step envelopes). *Let $T : \mathbb{N} \to \mathbb{N}$ denote the Collatz map* (1). *For every $n \in \mathbb{N}$ and $k \in \mathbb{N}_0$,*

$$\frac{n}{2^k} \ \leq \ T^k(n) \ \leq \ 3^k n + \frac{3^k - 1}{2}. \tag{5}$$

Proof. For every $m \geq 1$, the definition of T gives

$$\frac{m}{2} \leq T(m) \leq 3m + 1.$$

Iterating the lower bound yields $T^k(n) \geq n/2^k$. For the upper bound, the recurrence

$$T^{k+1}(n) \leq 3\,T^k(n) + 1$$

immediately gives, by a simple induction on k, the explicit estimate $T^k(n) \leq 3^k n + (3^k - 1)/2$. This proves (5). $\square$

These envelopes are intentionally crude, yet they ensure that forward iterates of typical arithmetic weights remain controlled on the scales relevant for our Dirichlet and transfer-operator analysis.

2.1. Weighted Dirichlet Spaces and Transform Estimates

For $\sigma > 0$ we define the weighted ℓ^1 space

$$\ell_\sigma^1 := \left\{ f : \mathbb{N} \to \mathbb{C} : \|f\|_\sigma := \sum_{n \geq 1} \frac{|f(n)|}{n^\sigma} < \infty \right\}. \tag{6}$$

The weight exponent σ measures polynomial decay and is chosen so that Dirichlet series associated with f converge absolutely in a half-plane $\Re(s) > \sigma$. Given $f \in \ell_\sigma^1$, we define its Dirichlet transform

$$Df(s) := \sum_{n \geq 1} \frac{f(n)}{n^s}, \qquad \Re(s) > \sigma. \tag{7}$$

Lemma 2.2 (Dirichlet convergence). *Let $\sigma > 0$ and let $f \in \ell_\sigma^1$, so that*

$$\|f\|_\sigma := \sum_{n \geq 1} \frac{|f(n)|}{n^\sigma} < \infty.$$

Then the Dirichlet transform

$$Df(s) := \sum_{n \geq 1} \frac{f(n)}{n^s}$$

converges absolutely for $\Re(s) > \sigma$ and defines a bounded holomorphic function on every half-plane $\Re(s) \geq \sigma + \varepsilon$, $\varepsilon > 0$. Moreover,

$$|Df(s)| \leq \|f\|_\sigma \sup_{n \geq 1} n^{\sigma - \Re(s)} = \|f\|_\sigma \qquad (\Re(s) > \sigma). \tag{8}$$

Proof. Let $s \in \mathbb{C}$ with $\Re(s) > \sigma$. Then

$$\sum_{n \geq 1} \left| \frac{f(n)}{n^s} \right| = \sum_{n \geq 1} \frac{|f(n)|}{n^{\Re(s)}} = \sum_{n \geq 1} \frac{|f(n)|}{n^\sigma} n^{\sigma - \Re(s)}.$$

Since $\Re(s) > \sigma$ implies $\sigma - \Re(s) < 0$, the sequence $n^{\sigma - \Re(s)}$ is decreasing to 0, and hence

$$\sup_{n \geq 1} n^{\sigma - \Re(s)} = 1.$$

Therefore,

$$\sum_{n \geq 1} \left| \frac{f(n)}{n^s} \right| \leq \|f\|_\sigma < \infty,$$

so the Dirichlet series converges absolutely.

For every $\varepsilon > 0$, the same bound holds uniformly on the half-plane $\Re(s) \geq \sigma + \varepsilon$, since then $\sigma - \Re(s) \leq -\varepsilon$ and $n^{\sigma - \Re(s)} \leq n^{-\varepsilon} \to 0$ as $n \to \infty$. Thus the convergence is locally uniform in $\Re(s) \geq \sigma + \varepsilon$, and classical Dirichlet-series theory implies that Df is holomorphic on this region. The bound (8) follows directly from the estimate above. $\square$

We write $\ell^1 = \ell_0^1$ for the unweighted space with norm $\|f\|_1 = \sum_{n \geq 1} |f(n)|$.

2.2. Collatz Preimage Geometry and the Backward Recursion

For each $n \geq 1$, define the even and odd preimage sets

$$E(n) := \{m \in \mathbb{N} : T(m) = n, \ m \text{ even}\}, \qquad O(n) := \{m \in \mathbb{N} : T(m) = n, \ m \text{ odd}\}.$$

Lemma 2.3 (Preimage structure). *For every $n \in \mathbb{N}$,*

$$E(n) = \{2n\}, \qquad O(n) = \begin{cases} \{(n-1)/3\}, & n \equiv 4 \pmod 6, \\ \varnothing, & \text{otherwise}, \end{cases} \qquad (9)$$

and in the first case $(n-1)/3$ is odd. In particular, each n has either one preimage (even) or two preimages (one even and one odd), and the odd preimage occurs with natural density $1/6$.

Proof. If m is even and $T(m) = n$, then $m/2 = n$, so $m = 2n$, establishing $E(n) = \{2n\}$. If m is odd and $T(m) = n$, then $3m + 1 = n$, so $m = (n-1)/3$. This is an integer precisely when $n \equiv 1 \pmod 3$. For m to be odd, $n - 1$ must be divisible by 3 but not by 6, so $n \equiv 4 \pmod 6$. In that case $(n-1)/3$ is odd. The density statement follows since the congruence class $n \equiv 4 \pmod 6$ has natural density $1/6$. $\square$

Hence each n admits exactly one even preimage and possibly one odd preimage when $n \equiv 4 \pmod 6$. The corresponding backward transfer operator is defined as

$$(Pf)(n) := \sum_{m:T(m)=n} \frac{f(m)}{m} = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4 \, (6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}. \qquad (10)$$

The normalization by $1/m$ reflects the logarithmic contraction of the forward map and ensures a natural mass-balance property.

Lemma 2.4 (Weighted mass preservation). *Let $f : \mathbb{N} \to [0, \infty)$ satisfy*

$$\sum_{m \geq 1} \frac{f(m)}{m} < \infty.$$

Then the backward transfer operator

$$(Pf)(n) := \sum_{m:T(m)=n} \frac{f(m)}{m}$$

preserves the weighted mass in the sense that

$$\sum_{n \geq 1} (Pf)(n) = \sum_{m \geq 1} \frac{f(m)}{m}. \qquad (11)$$

Proof. Since $f \geq 0$ and $\sum_{m \geq 1} f(m)/m < \infty$, Tonelli's theorem justifies rearranging the nonnegative double series. Using the definition of P,

$$\sum_{n \geq 1} (Pf)(n) = \sum_{n \geq 1} \sum_{m:T(m)=n} \frac{f(m)}{m}.$$

Each $m \geq 1$ has exactly one image $T(m)$, so it appears in exactly one of the inner sums. Hence we can rewrite the double sum directly over m:

$$\sum_{n \geq 1} \sum_{m:T(m)=n} \frac{f(m)}{m} = \sum_{m \geq 1} \frac{f(m)}{m},$$

which is precisely (11). $\square$

2.3. Dirichlet Envelopes for Iterated Backward Dynamics

The preimage structure allows a crude but useful bound on P acting on ℓ_σ^1.

Proposition 2.5 (Backward operator bound)**.** *Let $\sigma > 0$ and let P be defined by* (10)*. Then $P : \ell_\sigma^1 \to \ell_\sigma^1$ is bounded and*

$$\|Pf\|_\sigma \leq C_\sigma \|f\|_\sigma, \qquad C_\sigma := 2^\sigma + 3^{-\sigma}, \tag{12}$$

for all $f \in \ell_\sigma^1$. Consequently, for every $k \geq 1$,

$$\|P^k f\|_\sigma \leq C_\sigma^k \|f\|_\sigma. \tag{13}$$

Proof. From (10),

$$(Pf)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4 \,(6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}.$$

Hence

$$\|Pf\|_\sigma \leq S_{\text{even}} + S_{\text{odd}},$$

with

$$S_{\text{even}} := \sum_{n \geq 1} \frac{|f(2n)|}{2n\, n^\sigma}, \qquad S_{\text{odd}} := \sum_{\substack{n \geq 1 \\ n \equiv 4\,(6)}} \frac{\left|f\left(\frac{n-1}{3}\right)\right|}{\left(\frac{n-1}{3}\right) n^\sigma}.$$

For the even branch, set $m = 2n$, so $n = m/2$ and

$$S_{\text{even}} = \sum_{\substack{m \geq 1 \\ m \text{ even}}} \frac{|f(m)|}{m\,(m/2)^\sigma} = \sum_{\substack{m \geq 1 \\ m \text{ even}}} \frac{2^\sigma |f(m)|}{m^{\sigma+1}} \leq 2^\sigma \sum_{m \geq 1} \frac{|f(m)|}{m^\sigma} = 2^\sigma \|f\|_\sigma.$$

For the odd branch, write $m = (n-1)/3$, so $n = 3m + 1$ and m is odd. Then

$$S_{\text{odd}} = \sum_{\substack{m \geq 1 \\ m \text{ odd}}} \frac{|f(m)|}{m\,(3m+1)^\sigma} \leq \sum_{m \geq 1} \frac{|f(m)|}{m\,(3m)^\sigma} = 3^{-\sigma} \sum_{m \geq 1} \frac{|f(m)|}{m^{\sigma+1}} \leq 3^{-\sigma} \|f\|_\sigma.$$

Combining the two estimates gives (12), and iterating yields (13). □

The constant $C_\sigma = 2^\sigma + 3^{-\sigma}$ is an explicit growth factor for P on ℓ_σ^1. It is not < 1 in this normalization, so no contraction is claimed at this level. The genuine contraction mechanism is obtained later on the multiscale Banach space B_{tree}, where a strong seminorm captures oscillatory decay along the Collatz tree while the ℓ^1 component provides compactness.

3. Transfer Operator Formulation
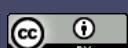
We now reformulate the Collatz dynamics in terms of the *backward transfer operator* associated with the map (1). This operator-theoretic viewpoint provides an analytic bridge between the discrete recurrence and the functional framework developed in later sections. The transfer operator encodes the inverse–branching structure of the map and propagates densities backward along the Collatz tree, in a form compatible with logarithmic weighting and Dirichlet series.

Recall that the Collatz map, (1), by Lemma 2.3, each $n \geq 1$ has the even preimage $2n$, together with an additional odd preimage $(n-1)/3$ precisely when $n \equiv 4 \pmod{6}$.

3.1. Backward Transfer Operator

Definition 3.1 (Backward transfer operator)**.** *For an arithmetic function $f : \mathbb{N} \to \mathbb{C}$, define*

$$(Pf)(n) := \sum_{m:\, T(m)=n} \frac{f(m)}{m} = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4\,(6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}, \qquad n \in \mathbb{N}, \tag{14}$$

where $\mathbf{1}_A$ denotes the indicator of the condition A.

Lemma 3.2 (Dirichlet transform intertwining). *Let $f \in \ell^1_\sigma$ with $\sigma > 1$, and define the Dirichlet transform*

$$D(f)(s) = \sum_{n \geq 1} f(n) n^{-s}.$$

Then for every s with $\Re(s) > \sigma$, the series $D(f)(s)$ and $D(Pf)(s)$ converge absolutely. Moreover, if we write

$$D(f)(s) = \sum_{n \geq 1} a_n n^{-s}, \qquad a_n := f(n),$$

and define a linear operator L_s on Dirichlet series by

$$(L_s F)(s) := 2^s \sum_{\substack{m \geq 1 \\ m \text{ even}}} a_m \, m^{-1-s} + \sum_{\substack{k \geq 1 \\ k \text{ odd}}} a_k \, k^{-1} (3k+1)^{-s}, \quad F(s) = \sum_{n \geq 1} a_n n^{-s}, \tag{15}$$

then

$$D(Pf)(s) = (L_s D(f))(s) \qquad \text{for all } \Re(s) > \sigma.$$

Proof. Fix $f \in \ell^1_\sigma$ with $\sigma > 1$. By definition of the ℓ^1_σ-norm,

$$\sum_{n \geq 1} |f(n)| \, n^{-\sigma} < \infty.$$

If $\Re(s) > \sigma$, then $n^{-\Re(s)} \leq n^{-\sigma}$ for all $n \geq 1$, so

$$\sum_{n \geq 1} |f(n)| \, n^{-\Re(s)} \; \leq \; \sum_{n \geq 1} |f(n)| \, n^{-\sigma} < \infty.$$

Thus $D(f)(s) = \sum_{n \geq 1} f(n) n^{-s}$ converges absolutely for $\Re(s) > \sigma$.

Next we show that $D(Pf)(s)$ converges absolutely for the same range. From the definition of P,

$$(Pf)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4 \ (\text{mod } 6)\}} \frac{f((n-1)/3)}{(n-1)/3},$$

so

$$|Pf(n)| \leq \frac{|f(2n)|}{2n} + \mathbf{1}_{\{n \equiv 4 \ (\text{mod } 6)\}} \frac{|f((n-1)/3)|}{(n-1)/3}.$$

Hence

$$\sum_{n \geq 1} |Pf(n)| \, n^{-\Re(s)} \; \leq \; S_{\text{even}} + S_{\text{odd}},$$

where

$$S_{\text{even}} := \sum_{n \geq 1} \frac{|f(2n)|}{2n} n^{-\Re(s)}, \qquad S_{\text{odd}} := \sum_{\substack{n \geq 1 \\ n \equiv 4 \ (6)}} \frac{|f((n-1)/3)|}{(n-1)/3} n^{-\Re(s)}.$$

For the even contribution, set $m = 2n$ so $n = m/2$ and m is even. Then

$$S_{\text{even}} = \sum_{\substack{m \geq 1 \\ m \text{ even}}} \frac{|f(m)|}{m} \left(\frac{m}{2}\right)^{-\Re(s)} = 2^{\Re(s)} \sum_{\substack{m \geq 1 \\ m \text{ even}}} |f(m)| \, m^{-1-\Re(s)}.$$

Since $\Re(s) > \sigma$ implies $\Re(s) + 1 > \sigma$, we have $m^{-1-\Re(s)} \leq m^{-\sigma}$, and therefore

$$S_{\text{even}} \leq 2^{\Re(s)} \sum_{\substack{m \geq 1 \\ m \text{ even}}} |f(m)| \, m^{-\sigma} \leq 2^{\Re(s)} \sum_{m \geq 1} |f(m)| \, m^{-\sigma} < \infty.$$

For the odd contribution, write $n = 3k + 1$ with $k \geq 1$ odd (equivalent to $n \equiv 4 \pmod{6}$ and $(n-1)/3 = k$ odd). Then

$$S_{\text{odd}} = \sum_{\substack{k \geq 1 \\ k \text{ odd}}} \frac{|f(k)|}{k} (3k+1)^{-\Re(s)}.$$

Since $3k + 1 \geq k$ for all $k \geq 1$, we have $(3k+1)^{-\Re(s)} \leq k^{-\Re(s)}$, and hence

$$S_{\text{odd}} \leq \sum_{\substack{k \geq 1 \\ k \text{ odd}}} |f(k)| \, k^{-1-\Re(s)} \leq \sum_{k \geq 1} |f(k)| \, k^{-1-\Re(s)}.$$

Again $\Re(s) + 1 > \sigma$ gives $k^{-1-\Re(s)} \leq k^{-\sigma}$, so

$$S_{\text{odd}} \leq \sum_{k \geq 1} |f(k)| \, k^{-\sigma} < \infty.$$

Thus $S_{\text{even}} + S_{\text{odd}} < \infty$, and $D(Pf)(s)$ converges absolutely for $\Re(s) > \sigma$.

We now compute $D(Pf)(s)$ explicitly and identify it with $(L_s D(f))(s)$. By definition,

$$D(Pf)(s) = \sum_{n \geq 1} (Pf)(n) \, n^{-s}.$$

Substituting the formula for P and splitting according to the two branches,

$$D(Pf)(s) = \sum_{n \geq 1} \frac{f(2n)}{2n} \, n^{-s} + \sum_{\substack{n \geq 1 \\ n \equiv 4 \, (6)}} \frac{f((n-1)/3)}{(n-1)/3} \, n^{-s}.$$

For the even part, set again $m = 2n$:

$$\sum_{n \geq 1} \frac{f(2n)}{2n} \, n^{-s} = \sum_{\substack{m \geq 1 \\ m \text{ even}}} \frac{f(m)}{m} \left(\frac{m}{2}\right)^{-s} = 2^s \sum_{\substack{m \geq 1 \\ m \text{ even}}} f(m) \, m^{-1-s}.$$

For the odd part, write $n = 3k + 1$ with $k \geq 1$ odd and $(n-1)/3 = k$:

$$\sum_{\substack{n \geq 1 \\ n \equiv 4 \, (6)}} \frac{f((n-1)/3)}{(n-1)/3} \, n^{-s} = \sum_{\substack{k \geq 1 \\ k \text{ odd}}} \frac{f(k)}{k} (3k+1)^{-s}.$$

Putting the two contributions together,

$$D(Pf)(s) = 2^s \sum_{\substack{m \geq 1 \\ m \text{ even}}} f(m) \, m^{-1-s} + \sum_{\substack{k \geq 1 \\ k \text{ odd}}} f(k) \, k^{-1} (3k+1)^{-s}.$$

Now write $F(s) = D(f)(s) = \sum_{n \geq 1} a_n n^{-s}$ with $a_n = f(n)$. By the definition (15) of L_s, we have

$$(L_s F)(s) = 2^s \sum_{\substack{m \geq 1 \\ m \text{ even}}} a_m \, m^{-1-s} + \sum_{\substack{k \geq 1 \\ k \text{ odd}}} a_k \, k^{-1} (3k+1)^{-s},$$

which matches exactly the expression for $D(Pf)(s)$ obtained above when $a_n = f(n)$. Hence

$$D(Pf)(s) = (L_s D(f))(s)$$

for all $\Re(s) > \sigma$, as claimed. $\quad \square$

The multiplicative factor $1/m$ assigns to each inverse branch a logarithmic weight, so that P acts as a normalized backward average along preimages. This normalization aligns the discrete dynamics with Dirichlet weights and will be crucial for analytic continuation and spectral estimates below.

Positivity. If $f(n) \geq 0$ for all n, then $(Pf)(n) \geq 0$ for all n, since P is a positive linear combination of values of f.

Weighted mass preservation. A direct change of variables shows that for every nonnegative f satisfying $\sum_{m \geq 1} |f(m)|/m < \infty$,

$$\sum_{n \geq 1} (Pf)(n) = \sum_{m \geq 1} \frac{f(m)}{m}. \tag{16}$$

Thus P preserves the logarithmically weighted mass $\sum f(m)/m$; plain ℓ^1 mass is not preserved under this normalization.

Boundedness on weighted spaces. Let

$$\ell_\sigma^1 := \left\{ f : \mathbb{N} \to \mathbb{C} \ : \ \|f\|_{\ell_\sigma^1} := \sum_{n \geq 1} \frac{|f(n)|}{n^\sigma} < \infty \right\}, \qquad \sigma > 0.$$

A direct change of variables in (14) yields, for all $f \in \ell_\sigma^1$,

$$\|Pf\|_{\ell_\sigma^1} = \sum_{n \geq 1} \frac{|(Pf)(n)|}{n^\sigma} \leq \sum_{n \geq 1} \left(\frac{|f(2n)|}{2n^{1+\sigma}} + \mathbf{1}_{\{n \equiv 4(6)\}} \frac{|f((n-1)/3)|}{((n-1)/3)^{1+\sigma}} \right)$$

$$= \frac{1}{2} \sum_{n \geq 1} \frac{|f(2n)|}{n^{1+\sigma}} + 3^{1+\sigma} \sum_{\substack{n \geq 1 \\ n \equiv 4(6)}} \frac{|f((n-1)/3)|}{(n-1)^{1+\sigma}}. \tag{17}$$

Changing variables $m = 2n$ in the first sum and $m = (n-1)/3$ in the second gives

$$\sum_{n \geq 1} \frac{|f(2n)|}{2n^{1+\sigma}} = 2^\sigma \sum_{\substack{m \geq 1 \\ m \text{ even}}} \frac{|f(m)|}{m^{1+\sigma}} \leq 2^\sigma \|f\|_{\ell_\sigma^1},$$

$$3^{1+\sigma} \sum_{\substack{n \geq 1 \\ n \equiv 4(6)}} \frac{|f((n-1)/3)|}{(n-1)^{1+\sigma}} = 3^{-\sigma} \sum_{\substack{m \geq 1 \\ 3m+1 \equiv 4(6)}} \frac{|f(m)|}{m^\sigma} \leq 3^{-\sigma} \|f\|_{\ell_\sigma^1}.$$

Hence

$$\|Pf\|_{\ell_\sigma^1} \leq \left(2^\sigma + 3^{-\sigma} \right) \|f\|_{\ell_\sigma^1}, \tag{18}$$

and therefore

$$\|P^k f\|_{\ell_\sigma^1} \leq \left(2^\sigma + 3^{-\sigma} \right)^k \|f\|_{\ell_\sigma^1}, \qquad k \geq 0. \tag{19}$$

Action on the weighted sup space. For the Banach space

$$B_\sigma := \left\{ f : \mathbb{N} \to \mathbb{C} \ : \ \|f\|_{B_\sigma} := \sup_{n \geq 1} n^\sigma |f(n)| < \infty \right\},$$

the normalization factor $1/m$ in (14) improves decay at each branch but does not make P a contraction. Setting $g(n) := nf(n)$, one obtains

$$n(Pf)(n) = g(2n) + \mathbf{1}_{\{n \equiv 4(6)\}} g\left(\frac{n-1}{3} \right), \qquad (Pf)(n) = \frac{(Qg)(n)}{n}, \qquad (Qg)(n) := g(2n) + \mathbf{1}_{\{n \equiv 4(6)\}} g\left(\frac{n-1}{3} \right).$$

Using $\|f\|_{B_\sigma} = \|g\|_{B_{\sigma-1}}$, one obtains the bound

$$\|Pf\|_{B_\sigma} = \sup_{n\geq 1} n^{\sigma-1}|(Qg)(n)| \leq \sup_{n\geq 1}\left(n^{\sigma-1}|g(2n)| + n^{\sigma-1}\mathbf{1}_{\{n\equiv 4(6)\}}\left|g\left(\tfrac{n-1}{3}\right)\right|\right)$$

$$\leq \left(2^{-(\sigma-1)} + 3^{\sigma-1}\right)\|g\|_{B_{\sigma-1}} = \left(2^{-(\sigma-1)} + 3^{\sigma-1}\right)\|f\|_{B_\sigma}. \tag{20}$$

In particular, the constant $2^{-(\sigma-1)} + 3^{\sigma-1} \geq 1$ for all $\sigma > 0$, so P is bounded but not contractive on $(B_\sigma, \|\cdot\|_{B_\sigma})$. This coarse boundedness provides an upper envelope for the operator norm but does not imply any decay of P^k on B_σ.

These limitations motivate the refinement of the functional setting in later sections, where the multiscale tree spaces B_{tree} and $B_{\text{tree},\sigma}$ are introduced to obtain genuine Lasota–Yorke-type contractions with $\lambda < 1$ and a provable spectral gap.

3.2. Dirichlet–Ruelle Operator and Intertwining Relations

For $f \in \ell^1_\sigma$ with $\sigma > 0$, the Dirichlet transform

$$\mathcal{D}f(s) := \sum_{n\geq 1} \frac{f(n)}{n^s}, \qquad \Re(s) > \sigma, \tag{21}$$

is absolutely convergent. Writing $\mathcal{D}f(s) = \sum_{n\geq 1} a_n n^{-s}$ with $a_n = f(n)$ and substituting (14), we obtain

$$\mathcal{D}(Pf)(s) = \sum_{n\geq 1}\left(\frac{a_{2n}}{2n} + \mathbf{1}_{\{n\equiv 4\,(6)\}}\frac{a_{(n-1)/3}}{(n-1)/3}\right)\frac{1}{n^s}. \tag{22}$$

Thus $\mathcal{D}(Pf)$ is again a Dirichlet series whose coefficients depend linearly on those of $\mathcal{D}f$.

Definition 3.3 (Dirichlet–Ruelle operator). *Let $\mathcal{D}_\sigma$ denote the space of Dirichlet series*

$$F(s) = \sum_{n\geq 1} a_n n^{-s} \quad \text{with} \quad \sum_{n\geq 1}\frac{|a_n|}{n^\sigma} < \infty.$$

Define $L : \mathcal{D}_\sigma \to \mathcal{D}_\sigma$ by

$$(LF)(s) := \sum_{n\geq 1} b_n n^{-s}, \qquad b_n := \frac{a_{2n}}{2n} + \mathbf{1}_{\{n\equiv 4\,(6)\}}\frac{a_{(n-1)/3}}{(n-1)/3}. \tag{23}$$

Lemma 3.4 (Operator norm of L). *For $\sigma > 0$, let $\|F\|_\sigma := \sum_{n\geq 1} |a_n|/n^\sigma$. Then $L : \mathcal{D}_\sigma \to \mathcal{D}_\sigma$ is bounded and*

$$\|L\|_\sigma \leq 2^\sigma + 3^{-\sigma}. \tag{24}$$

Proof. From (23),

$$\|LF\|_\sigma = \sum_{n\geq 1}\frac{|b_n|}{n^\sigma} \leq \sum_{n\geq 1}\frac{|a_{2n}|}{2n\,n^\sigma} + \sum_{\substack{n\geq 1 \\ n\equiv 4\,(6)}}\frac{|a_{(n-1)/3}|}{(n-1)/3}\frac{1}{n^\sigma} =: S_{\text{even}} + S_{\text{odd}}.$$

For the even term, set $m = 2n$. Then

$$S_{\text{even}} = \sum_{m \text{ even}}\frac{|a_m|}{2(m/2)^{1+\sigma}} = \sum_{m \text{ even}}\frac{2^\sigma |a_m|}{m^{1+\sigma}} \leq 2^\sigma \sum_{m \text{ even}}\frac{|a_m|}{m^\sigma} \leq 2^\sigma \|F\|_\sigma.$$

For the odd term, write $m = (n-1)/3$, so $n = 3m+1$ and

$$S_{\text{odd}} = \sum_{m\geq 1}\frac{|a_m|}{m(3m+1)^\sigma} \leq 3^{-\sigma}\sum_{m\geq 1}\frac{|a_m|}{m^\sigma} = 3^{-\sigma}\|F\|_\sigma.$$

Combining the two estimates gives

$$\|LF\|_\sigma \leq (2^\sigma + 3^{-\sigma})\,\|F\|_\sigma,$$

proving (24). □

Lemma 3.5 (Intertwining of P and L). *For every $f \in \ell^1_\sigma$ with $\sigma > 0$,*

$$\mathcal{D}(Pf) = L(\mathcal{D}f), \qquad \mathcal{D}(P^k f) = L^k(\mathcal{D}f), \quad k \geq 0, \tag{25}$$

whenever the series converge absolutely.

Proof. The Dirichlet coefficients of $\mathcal{D}(Pf)$ in (22) are precisely the b_n of (23), so $\mathcal{D}(Pf) = L(\mathcal{D}f)$; iteration gives the second identity. □

The intertwining relation shows that spectral information for P on ℓ^1_σ transfers to L on $\mathcal{D}_\sigma$. However, since P is not contractive on ℓ^1_σ or B_σ, the inequality (24) provides only a uniform boundedness envelope for $\|L^k\|_\sigma$, not exponential decay. Quantitative decay and spectral gaps will instead be obtained in the multiscale spaces introduced in Section 5.

Define $w_k := P^k \mathbf{1}$ with $\mathbf{1}(n) \equiv 1$ and

$$\zeta_C(s,k) := \sum_{n \geq 1} \frac{w_k(n)}{n^s}, \qquad \Re(s) \text{ large}. \tag{26}$$

By Lemma 3.5,
$$\zeta_C(s,0) = \zeta(s), \qquad \zeta_C(s,k) = (L^k \zeta)(s), \quad k \geq 1. \tag{27}$$

The quantity $w_k(n)$ represents the total normalized weight of all k–step backward paths from n in the Collatz tree under the logarithmic weighting $1/m$. The family $\zeta_C(s,k)$ therefore encodes, in Dirichlet form, the distribution of these weighted backward configurations at depth k. By Lemma 3.4,

$$\|L^k\|_\sigma \leq (2^\sigma + 3^{-\sigma})^k,$$

so the Dirichlet coefficients of $\zeta_C(s,k)$ are uniformly bounded in $\Re(s) > \sigma$ but do not necessarily decay in k. Later sections refine this estimate by passing to the multiscale tree space $B_{\text{tree},\sigma}$, where the Lasota–Yorke inequality ensures a true spectral gap and exponential decay of P^k.

4. Spectral Framework and Multiscale Contraction Theory

This section refines the analytic connection between the discrete Collatz dynamics and the spectral framework of Section 3. Our goal is to express analytic information about the Dirichlet series associated with iterates of the backward operator P in terms of the spectral data of P—equivalently, of the Dirichlet–Ruelle operator L—acting on suitable Banach spaces continuously embedded in ℓ^1_σ. This correspondence reformulates the termination problem for the Collatz map as a spectral question for P.

Throughout this section we fix $\sigma > 1$ and a Banach space $B_{\sigma,1}$ of arithmetic functions such that $B_{\sigma,1} \subset \ell^1_\sigma$ continuously, $P(B_{\sigma,1}) \subset B_{\sigma,1}$, and the Dirichlet transform

$$\mathcal{D}f(s) = \sum_{n \geq 1} \frac{f(n)}{n^s}$$

defines a holomorphic function for $\Re(s) > \sigma$ whenever $f \in B_{\sigma,1}$. The intertwining relation (25) then yields, for all $k \geq 0$,

$$\mathcal{D}(P^k f)(s) = \sum_{n \geq 1} \frac{(P^k f)(n)}{n^s}, \qquad \Re(s) > \sigma.$$

Since $B_{\sigma,1} \subset \ell_\sigma^1$, each series converges absolutely. By the ℓ_σ^1 estimate (18),

$$|\mathcal{D}(P^k f)(s)| \leq \|P^k f\|_{\ell_\sigma^1} \leq \left(2^\sigma + 3^{-\sigma}\right)^k \|f\|_{\ell_\sigma^1}, \qquad \Re(s) > \sigma. \tag{28}$$

The bound (28) shows that the iterates of P are uniformly bounded on ℓ_σ^1, though not contractive; a genuine contraction will appear only after the refinement to the multiscale tree spaces introduced in Section 4.4.

Generating function and operator resolvent. For $z \in \mathbb{C}$ with $|z| < (2^\sigma + 3^{-\sigma})^{-1}$, define the two–variable generating function

$$G_f(s,z) := \sum_{k \geq 0} z^k \, \mathcal{D}(P^k f)(s). \tag{29}$$

The series converges absolutely and locally uniformly for $\Re(s) > \sigma$, hence G_f is holomorphic in (s,z) on the domain

$$\Omega_\sigma := \{(s,z) \in \mathbb{C}^2 : \Re(s) > \sigma, \ |z| < (2^\sigma + 3^{-\sigma})^{-1}\}.$$

On the operator side, for such z the Neumann series

$$(I - zP)^{-1} = \sum_{k \geq 0} z^k P^k$$

converges in operator norm on $B_{\sigma,1}$, and thus

$$G_f(s,z) = \mathcal{D}\left[(I - zP)^{-1} f\right](s), \qquad (s,z) \in \Omega_\sigma. \tag{30}$$

The poles of $(I - zP)^{-1}$ in the z–plane occur precisely at the reciprocals of the spectral values of P on $B_{\sigma,1}$. Consequently the analytic structure of G_f as a function of z is governed by the spectrum of P.

At this point we recall that the backward Collatz transfer operator P preserves total mass on ℓ^1:

$$\sum_{n \geq 1} (Pf)(n) \ = \ \sum_{m \geq 1} f(m),$$

so 1 is the maximal eigenvalue of P on ℓ^1. On the Banach space $B_{\text{tree},\sigma}$, however, the associated invariant object is not the constant function, but rather the unique positive eigenfunction h satisfying $Ph = h$, as constructed in Section 5. Thus the spectral analysis of P is directed toward establishing a *spectral gap at the eigenvalue* 1, meaning that every other spectral value λ satisfies $|\lambda| \leq \lambda_{\text{LY}} < 1$, where λ_{LY} is the Lasota–Yorke contraction constant for the tree norm. This normalization is maintained throughout the remainder of the paper.

Consequently, the resolvent expansion (30) is analytic for $|z| < 1$ except at the simple pole at $z = 1$. The residue at this pole coincides with the Riesz projection onto the eigenspace spanned by h, and therefore encodes the invariant functional associated with the positive eigenfunction h.

The coarse resolvent radius $(2^\sigma + 3^{-\sigma})^{-1}$ merely provides an elementary domain of convergence. A sharper meromorphic continuation—reflecting the true spectral radius $r(P) = 1$ and the subdominant bound $\rho_{\text{ess}}(P) \leq \lambda_{\text{LY}} < 1$—will be obtained on the refined spaces B_{tree} and $B_{\text{tree},\sigma}$, where the Lasota–Yorke inequality gives quantitative contraction of oscillations between adjacent scales.

Finally, for the constant function $\mathbf{1}(n) \equiv 1$ (whenever $\mathbf{1} \in B_{\sigma,1}$), the coefficients of $G_{\mathbf{1}}(s,z)$ are precisely the Collatz Dirichlet series $\zeta_C(s,k)$ defined in (26). Thus the analytic continuation and asymptotic decay of $\zeta_C(s,k)$ as $k \to \infty$ are controlled by the spectral properties of P through (30); their exponential decay emerges once the spectral gap on the multiscale tree spaces is established.

4.1. Spectral Reduction and Meromorphic Continuation of Dirichlet Transforms

Recall that the Dirichlet–Ruelle operator L is defined on $\mathcal{D}_\sigma$ by (23). The intertwining Lemma 3.5 asserts that for all $f \in \ell^1_\sigma$,

$$\mathcal{D}(Pf) = L(\mathcal{D}f).$$

Since $\mathcal{D}$ is injective on ℓ^1_σ, every eigenpair (λ, f) of P with $f \in \ell^1_\sigma$ produces an eigenpair $(\lambda, \mathcal{D}f)$ of L. Conversely, if $LF = \lambda F$ and $F = \mathcal{D}f$ lies in the image of $\mathcal{D}$, then $Pf = \lambda f$. Hence the point spectra of P on $B_{\sigma,1}$ and of L on $\mathcal{D}_\sigma$ coincide on the subspace $\mathcal{D}(B_{\sigma,1})$. In particular,

$$\rho(L) \;\geq\; \rho(P), \tag{31}$$

and any spectral gap or peripheral spectral property of P transfers to the induced action of L on Dirichlet series arising from $B_{\sigma,1}$.

We emphasize that equality $\sigma(L) = \sigma(P)$ is not assumed. The partial correspondence (31) suffices for analytic reduction: the Dirichlet-side continuation of $\mathcal{D}(P^k f)$ reflects the spectral geometry of P.

Mass preservation and spectral gap. Because P only preserves total mass up to a logarithmic factor, we have

$$\sum_{n \geq 1} (Pf)(n) \;=\; \sum_{m \geq 1} \frac{f(m)}{m},$$

so the constant function $1(n) \equiv 1$ is *not* an eigenvector. Instead, P admits a unique positive invariant density $h \in B_{\text{tree},\sigma}$ and a unique positive invariant functional $\phi \in B^*_{\text{tree},\sigma}$ with

$$Ph = h, \qquad \phi \circ P = \phi, \qquad \phi(h) = 1. \tag{32}$$

Throughout the paper we work with this Perron–Frobenius normalization (32) and express all spectral decompositions relative to the nonconstant invariant profile h.

Within this framework, the Dirichlet–Ruelle operator L inherits the same dominant eigenvalue 1 and the same spectral gap on the subspace $\mathcal{D}(B_{\sigma,1})$. The analytic behavior of the Collatz Dirichlet series $\zeta_C(s,k) = \mathcal{D}(P^k \mathbf{1})(s)$ is then determined by how P^k approaches the spectral projector onto the invariant subspace spanned by $\mathbf{1}$.

Theorem 4.1 (Spectral reduction and analytic continuation). *Let $B_{\sigma,1}$ be a Banach space of arithmetic functions continuously embedded in ℓ^1_σ such that $P : B_{\sigma,1} \to B_{\sigma,1}$ is quasi-compact and satisfies the mass-preserving normalization* (11). *Assume further that 1 is a simple eigenvalue of P and that all other spectral values lie in the closed disk $|\lambda| \leq \lambda_{\text{LY}} < 1$. Then for every $f \in B_{\sigma,1}$ the Dirichlet transforms $\mathcal{D}(P^k f)(s)$ extend holomorphically to $\Re(s) > \sigma$ and admit the decomposition*

$$\mathcal{D}(P^k f)(s) = \Pi_1(f)\,\mathcal{D}(\mathbf{1})(s) + R_k(s), \qquad |R_k(s)| \leq C_f(s)\,\lambda^k_{\text{LY}}, \tag{33}$$

where Π_1 is the spectral projection associated with the eigenvalue 1 and $C_f(s)$ is locally bounded on $\{\Re(s) > \sigma\}$. In particular, for f with $\Pi_1(f) = 0$, the functions $\mathcal{D}(P^k f)(s)$ decay exponentially in k uniformly on compact subsets of $\Re(s) > \sigma$.

When $f = \mathbf{1}$, the same conclusion applies to $\zeta_C(s,k) = \mathcal{D}(P^k \mathbf{1})(s)$, whose exponential stabilization corresponds to convergence toward the invariant density associated with the Collatz operator.

Proof. By quasi-compactness, the spectrum of P decomposes as

$$\sigma(P) = \{1\} \cup \sigma_{\text{ess}}(P), \qquad \rho_{\text{ess}}(P) \leq \lambda_{\text{LY}} < 1,$$

and the Riesz projection $\Pi_1 = \frac{1}{2\pi i} \oint_{|z-1|=\varepsilon} (zI - P)^{-1}\,dz$ is a bounded projection onto the one-dimensional invariant subspace spanned by $\mathbf{1}$. Then $P^k = \Pi_1 + N^k$, where $\|N^k\|_{B_{\sigma,1}} \leq C\,\lambda^k_{\text{LY}}$ for

some constant $C > 0$. Applying the Dirichlet transform and using $|\mathcal{D}(g)(s)| \leq \|g\|_{\ell^1_\sigma}$ for $\Re(s) > \sigma$ gives

$$\mathcal{D}(P^k f)(s) = \mathcal{D}(\Pi_1 f)(s) + \mathcal{D}(N^k f)(s), \qquad |\mathcal{D}(N^k f)(s)| \leq C \lambda_{\mathrm{LY}}^k \|f\|_{B_{\sigma,1}}.$$

Since $\Pi_1 f$ is a multiple of $\mathbf{1}$, we may write $\mathcal{D}(\Pi_1 f) = \Pi_1(f)\,\mathcal{D}(\mathbf{1})$, yielding (33). Analyticity for $\Re(s) > \sigma$ follows from absolute convergence and locally uniform bounds. $\quad\square$

This form aligns with the quasi-compactness obtained later on the multiscale tree space $B_{\mathrm{tree},\sigma}$, where the Lasota–Yorke inequality ensures $\rho_{\mathrm{ess}}(P) \leq \lambda_{\mathrm{LY}} < 1$. The exponential term λ_{LY}^k in (33) corresponds to the essential spectral radius and controls the rate of decay of correlations and Dirichlet coefficients. Under stronger spectral assumptions, the representation can be refined to a meromorphic decomposition in which each isolated eigenvalue λ_j contributes a term $\lambda_j^k \mathcal{D}(\Pi_j f)$, generalizing the usual Ruelle–Perron expansion.

4.2. Spectral Criteria for Boundedness on Weighted Spaces

The preceding analysis shows that sufficiently strong spectral control of P on an appropriate Banach space $B_{\sigma,1}$ forces all Dirichlet data generated by the backward Collatz tree to exhibit exponential stabilization toward the invariant profile. Since P is not contractive on ℓ^1_σ or B_σ, such behavior can only arise on refined Banach spaces where a genuine spectral gap at the eigenvalue 1 has been established. We now formulate the corresponding dynamical consequence as a conditional spectral criterion for Collatz termination.

Theorem 4.2 (Spectral criterion for Collatz termination). *Let P act on a Banach space $B_{\sigma,1} \subset \ell^1_\sigma$ such that $P(B_{\sigma,1}) \subset B_{\sigma,1}$ and $\mathbf{1} \in B_{\sigma,1}$. Assume that P is quasi-compact on $B_{\sigma,1}$, that 1 is a simple eigenvalue of P corresponding to the unique positive invariant density h, and that all other spectral values satisfy*

$$\sigma(P) \setminus \{1\} \subset \{\, z \in \mathbb{C} : |z| \leq \lambda_{\mathrm{LY}} < 1 \,\}.$$

Then every $f \in B_{\sigma,1}$ admits a decomposition

$$P^k f = \Pi_1 f + N^k f, \qquad \|N^k f\|_{B_{\sigma,1}} \leq C \lambda_{\mathrm{LY}}^k \|f\|_{B_{\sigma,1}},$$

where Π_1 is the spectral projection onto $\mathrm{span}\{h\}$. Consequently, there exists no nontrivial invariant or periodic density for the backward Collatz dynamics in $B_{\sigma,1}$; the only invariant direction is the positive eigenfunction h. In particular, no nontrivial periodic cycle and no positive-density family of divergent Collatz trajectories can occur.

Proof. By quasi-compactness, the spectrum of P decomposes as $\sigma(P) = \{1\} \cup \sigma_{\mathrm{ess}}(P)$ with $\rho_{\mathrm{ess}}(P) \leq \lambda_{\mathrm{LY}} < 1$. The associated Riesz projection

$$\Pi_1 = \frac{1}{2\pi i} \oint_{|z-1|=\varepsilon} (zI - P)^{-1}\, dz$$

is bounded and satisfies $P\Pi_1 = \Pi_1 P = \Pi_1$. Since 1 is a simple eigenvalue with positive eigenfunction h, we have

$$\Pi_1 f = \big(\phi(f)\big)\, h,$$

where ϕ is the corresponding eigenfunctional normalized so that $\phi(h) = 1$.

Hence the power iterates decompose as

$$P^k = \Pi_1 + N^k, \qquad \|N^k\|_{B_{\sigma,1}} \leq C \lambda_{\mathrm{LY}}^k,$$

for some constant $C > 0$.

If a nontrivial invariant density $f \in B_{\sigma,1}$ satisfied $Pf = f$, then f would belong to the eigenspace of $\lambda = 1$. Since this eigenspace is one-dimensional and spanned by h, we must have $f = c\,h$ for some constant c. Thus no additional invariant densities exist beyond span$\{h\}$.

If a periodic density f satisfied $P^q f = f$ for some $q > 0$, then f would belong to an eigenspace associated with an eigenvalue λ satisfying $|\lambda| = 1$. Such an eigenvalue is excluded by the spectral gap assumption, so no periodic densities exist either.

Finally, via the standard correspondence between transfer-operator invariants and dynamical orbits on the Collatz graph, any invariant or periodic density corresponds to either a periodic Collatz cycle or to a positive-density family of non-terminating trajectories. The spectral gap therefore precludes these dynamical behaviors. □

Section 4.4 constructs the multiscale tree Banach space B_{tree} and establishes a Lasota–Yorke inequality that ensures quasi-compactness of P with an explicit contraction constant $\lambda_{\text{LY}} < 1$ in the strong seminorm. Verification of the hypotheses of Theorem 4.2 on $B_{\text{tree},\sigma}$ provides the analytic–spectral bridge: a strict spectral gap for P on $B_{\text{tree},\sigma}$ rules out the spectral signatures associated with any non-terminating Collatz behavior.

4.3. Construction of the Multiscale Tree Space $B_{\text{tree},\sigma}$

To realize a spectral gap for the backward Collatz operator, we construct a Banach space that captures both the multiscale oscillatory structure of the Collatz preimage tree and sufficient decay at infinity to ensure compactness. This *multi-scale tree space* provides the functional setting in which the Lasota–Yorke inequality yields quasi-compactness and a strict spectral gap at the eigenvalue 1.

For $j \geq 0$ define the scale blocks

$$I_j := [6^j, 2{\cdot}6^j) \cap \mathbb{N}. \tag{34}$$

The factor 6 reflects the approximate scale multiplication under the backward map, combining the even branch $m = 2n$ and the odd branch $m = (n-1)/3$ (defined for $n \equiv 4 \pmod 6$). Fix parameters $0 < \alpha < 1$ and $0 < \vartheta < 1$. For indices $u, v > 0$, define the scale-sensitive weight

$$W_\alpha(u,v) := \frac{uv}{|u-v|\,(u+v)^\alpha}, \qquad u \neq v. \tag{35}$$

This weight penalizes small separations between indices, emphasizing local oscillations of f, while the factor $(u+v)^{-\alpha}$ damps sensitivity at large scales. The geometric coefficient ϑ^j provides exponential attenuation of oscillations across successive levels of the tree.

Definition 4.3 (Multiscale tree seminorm and space). *For $f : \mathbb{N} \to \mathbb{C}$ define*

$$[f]_{\text{tree}} := \sum_{j \geq 0} \vartheta^j \sup_{\substack{m,n \in I_j \\ m \neq n}} W_\alpha(m,n)\,|f(m) - f(n)|. \tag{36}$$

The corresponding Banach space

$$B_{\text{tree}} := \{ f : \mathbb{N} \to \mathbb{C} : \ \|f\|_1 + [f]_{\text{tree}} < \infty \}, \qquad \|f\|_{\text{tree}} := \|f\|_1 + [f]_{\text{tree}},$$

is called the multiscale tree space.

Standard arguments for weighted variation-type seminorms show that $(B_{\text{tree}}, \|\cdot\|_{\text{tree}})$ is complete. The seminorm $[f]_{\text{tree}}$ controls the oscillatory irregularity of f within each scale block I_j, while the ℓ^1 component controls the overall magnitude. However, B_{tree} alone does not impose sufficient decay as $n \to \infty$ to guarantee compactness.

Weighted extension. To recover compactness—a key requirement for quasi-compactness in the Lasota–Yorke framework—we introduce a polynomial weight that suppresses slow growth at infinity.

Definition 4.4 (Weighted tree space with block decay). *For parameters $0 < \alpha < 1$, $0 < \vartheta < 1$, $\sigma > 1$, and $\eta > 1$, set*

$$\|f\|_\sigma := \sum_{n \geq 1} \frac{|f(n)|}{n^\sigma},$$

$$[f]_{\mathrm{osc}} := \sum_{j \geq 0} \vartheta^j \sup_{\substack{m,n \in I_j \\ m \neq n}} W_\alpha(m,n) \, |f(m) - f(n)|,$$

and

$$[f]_{\mathrm{mass}} := \sup_{j \geq 0} \eta^j \sum_{n \in I_j} \frac{|f(n)|}{n^\sigma}.$$

Then define

$$B_{\mathrm{tree},\sigma} := \{ f : \mathbb{N} \to \mathbb{C} : \|f\|_{\mathrm{tree},\sigma} < \infty \}, \qquad \|f\|_{\mathrm{tree},\sigma} := \|f\|_\sigma + [f]_{\mathrm{osc}} + [f]_{\mathrm{mass}}.$$

The weight $n^{-\sigma}$ enforces global summability of f, while the block–decay term $[f]_{\mathrm{mass}}$ forces the total weighted mass on each block I_j to decrease geometrically with j, ruling out escape of mass to high levels of the tree. The oscillation term $[f]_{\mathrm{osc}}$ controls the local multiscale variation of f within blocks. Together these components provide the strong–weak structure required for the Lasota–Yorke framework: the strong part $[f]_{\mathrm{osc}}$ contracts under the transfer operator, while the weak part $\|f\|_\sigma + [f]_{\mathrm{mass}}$ yields compactness of the unit ball and ensures tightness of mass across scales.

Lemma 4.5 (Compact embedding). *Fix $0 < \alpha < 1$, $0 < \vartheta < 1$, $\sigma > 1$, and $\eta > 1$. Then the unit ball of $B_{\mathrm{tree},\sigma}$ is relatively compact in ℓ_σ^1.*

Proof. Let

$$\mathcal{U} := \{ f \in B_{\mathrm{tree},\sigma} : \|f\|_{\mathrm{tree},\sigma} \leq 1 \}.$$

(i) Uniform boundedness. For $f \in \mathcal{U}$,

$$\|f\|_\sigma \leq \|f\|_{\mathrm{tree},\sigma} \leq 1,$$

so $\mathcal{U}$ is bounded in ℓ_σ^1.

(ii) Uniform tail control. From $[f]_{\mathrm{mass}} \leq 1$ we obtain, for every $j \geq 0$,

$$\sum_{n \in I_j} \frac{|f(n)|}{n^\sigma} \leq \eta^{-j}.$$

Fix $\varepsilon > 0$. Choose J so large that

$$\sum_{j > J} \eta^{-j} < \varepsilon.$$

Pick N so large that all blocks I_j with $j \leq J$ are contained in $\{1, \ldots, N\}$. Then for any $f \in \mathcal{U}$,

$$\sum_{n > N} \frac{|f(n)|}{n^\sigma} = \sum_{j > J} \sum_{n \in I_j} \frac{|f(n)|}{n^\sigma} \leq \sum_{j > J} \eta^{-j} < \varepsilon.$$

So the tails of $\mathcal{U}$ are uniformly small in ℓ_σ^1.

(iii) Compactness on finite windows. Fix $N \geq 1$ and consider the restriction map $f \mapsto (f(1), \ldots, f(N))$ from $\mathcal{U}$ into $\mathbb{C}^N$. For each $1 \leq n \leq N$ and $f \in \mathcal{U}$,

$$\frac{|f(n)|}{n^\sigma} \leq \|f\|_\sigma \leq 1,$$

hence $|f(n)| \leq n^\sigma \leq N^\sigma$. Therefore $\{f|_{\{1,\dots,N\}} : f \in \mathcal{U}\}$ is bounded in the finite-dimensional space $\mathbb{C}^N$ and so relatively compact there.

(iv) Diagonal extraction and ℓ^1_σ convergence. Let $(f^{(k)}) \subset \mathcal{U}$ be any sequence. By (iii) we can extract a subsequence $(f^{(k_\ell)})$ that converges in $\mathbb{C}^1$ at coordinate $n = 1$. Repeating this on coordinates $1, \dots, 2$, then $1, \dots, 3$, etc., and passing to a diagonal subsequence, we obtain a subsequence, still denoted $(f^{(k_\ell)})$, that converges pointwise on all of $\mathbb{N}$:

$$f^{(k_\ell)}(n) \to f(n) \qquad \text{for each } n \geq 1.$$

We now show $f^{(k_\ell)} \to f$ in ℓ^1_σ. Fix $\varepsilon > 0$ and choose N as in (ii) so that

$$\sup_{g \in \mathcal{U}} \sum_{n>N} \frac{|g(n)|}{n^\sigma} < \frac{\varepsilon}{3}.$$

By pointwise convergence on $\{1, \dots, N\}$ and finite-dimensional compactness, there exists L such that for all $\ell \geq L$,

$$\sum_{n \leq N} \frac{|f^{(k_\ell)}(n) - f(n)|}{n^\sigma} < \frac{\varepsilon}{3}.$$

For the tail, note that f also belongs to $\mathcal{U}$ as the limit of elements of $\mathcal{U}$ coordinatewise with uniform ℓ^1_σ bound, so

$$\sum_{n>N} \frac{|f(n)|}{n^\sigma} \leq \sup_{g \in \mathcal{U}} \sum_{n>N} \frac{|g(n)|}{n^\sigma} < \frac{\varepsilon}{3}.$$

Therefore, for $\ell \geq L$,

$$\sum_{n>N} \frac{|f^{(k_\ell)}(n) - f(n)|}{n^\sigma} \leq \sum_{n>N} \frac{|f^{(k_\ell)}(n)|}{n^\sigma} + \sum_{n>N} \frac{|f(n)|}{n^\sigma} < \frac{2\varepsilon}{3}.$$

Combining the head and tail estimates gives

$$\sum_{n \geq 1} \frac{|f^{(k_\ell)}(n) - f(n)|}{n^\sigma} < \varepsilon \qquad \text{for all } \ell \geq L,$$

so $f^{(k_\ell)} \to f$ in ℓ^1_σ. This shows that every sequence in $\mathcal{U}$ has a convergent subsequence in ℓ^1_σ, hence $\mathcal{U}$ is relatively compact. $\square$

Remark 4.6. *The additional block-decay seminorm $[f]_{\mathrm{mass}}$ is essential. If one worked only with $\|f\|_\sigma + [f]_{\mathrm{osc}}$, the unit ball would not be precompact in ℓ^1 or in ℓ^1_σ: one can construct a sequence of functions supported on disjoint blocks, constant on each block, whose oscillation seminorms vanish and whose ℓ^1_σ norms stay uniformly bounded, while the supports drift to infinity. The factor η^j in $[f]_{\mathrm{mass}}$ rules out this escape mechanism by forcing the weighted mass in block I_j to decay at least geometrically in j.*

The space $B_{\mathrm{tree},\sigma}$ thus provides the natural functional environment for the Lasota–Yorke inequality. Its compact embedding into ℓ^1_σ ensures that the essential spectral radius of P on $B_{\mathrm{tree},\sigma}$ is strictly smaller than its spectral radius, a prerequisite for establishing a genuine spectral gap. The strong seminorm captures multiscale regularity across the Collatz tree, while the weighted ℓ^1 norm supplies the compactness that underlies the spectral analysis of the backward transfer operator.

4.4. Lasota–Yorke Contraction on the Multiscale Tree Space

Recall from (10) that

$$(Pf)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4 \, (6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}.$$

It is convenient to split P into its even and odd components:

$$(P_{\text{even}}f)(n) := \frac{f(2n)}{2n}, \qquad (P_{\text{odd}}f)(n) := \mathbf{1}_{\{n\equiv 4\,(6)\}}\frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}, \tag{37}$$

so that $P = P_{\text{even}} + P_{\text{odd}}$.

From the ℓ^1 estimates of Section 2, both branches are bounded on ℓ^1, hence on B_{tree}. The Lasota–Yorke inequality arises from the fact that P_{even} is strongly contracting in the tree seminorm, while P_{odd} is a controlled perturbation whose contribution is damped by the multiscale factor ϑ^j.

4.4.1. Even-Branch Distortion and Contraction Estimates

We first record the even-branch estimate.

Lemma 4.7 (Even branch contraction on $B_{\text{tree},\sigma}$). *Fix $0 < \alpha < 1$, $0 < \vartheta < 1$, $\sigma > 1$, and $\eta > 1$. Let $P_{\text{even}}f(n) = f(2n)/(2n)$. There exist constants*

$$\lambda_{\text{even}} = 2^{-(2-\alpha)}\vartheta^{-1} \quad \text{and} \quad C_{\text{even}} > 0,$$

depending only on $\alpha, \vartheta, \sigma, \eta$, such that for all $f \in B_{\text{tree},\sigma}$,

$$[P_{\text{even}}f]_{\text{osc}} \le \lambda_{\text{even}}[f]_{\text{osc}} + C_{\text{even}}\left(\|f\|_\sigma + [f]_{\text{mass}}\right). \tag{38}$$

In particular, if $\vartheta > 2^{-(2-\alpha)}$, then $0 < \lambda_{\text{even}} < 1$, and the even branch is a strict contraction in the oscillation seminorm up to a controlled weak error.

Proof. Write

$$[f]_{\text{osc}} = \sum_{j\ge 0}\vartheta^j\delta_j(f), \qquad \delta_j(f) := \sup_{m\ne n\in I_j}W_\alpha(m,n)\,|f(m) - f(n)|.$$

Step 1: Decomposition and the D_1/D_2 split Fix $j \ge 0$ and $u, v \in I_j = [2^j, 2^{j+1})$, $u \ne v$. We write

$$(P_{\text{even}}f)(u) - (P_{\text{even}}f)(v) = \frac{f(2u)}{2u} - \frac{f(2v)}{2v} = \frac{f(2u) - f(2v)}{2u} + f(2v)\left(\frac{1}{2u} - \frac{1}{2v}\right) =: D_1(u,v) + D_2(u,v).$$

We estimate the contributions of D_1 and D_2 separately.

Step 2: Oscillatory term D_1: corrected contraction Using the scaling of W_α, we have

$$W_\alpha(2u, 2v) = 2^{1-\alpha}W_\alpha(u,v), \qquad W_\alpha(u,v) = 2^{-(1-\alpha)}W_\alpha(2u, 2v).$$

Thus

$$W_\alpha(u,v)\,|D_1(u,v)| = W_\alpha(u,v)\frac{|f(2u) - f(2v)|}{2u} = 2^{-(1-\alpha)}\frac{W_\alpha(2u, 2v)}{2u}|f(2u) - f(2v)|.$$

Since $u \in I_j = [2^j, 2^{j+1})$, we have $u \ge 2^j$ and hence

$$\frac{1}{2u} \le \frac{1}{2\cdot 2^j} = 2^{-(j+1)}.$$

Therefore

$$W_\alpha(u,v)\,|D_1(u,v)| \le 2^{-(1-\alpha)}2^{-(j+1)}W_\alpha(2u, 2v)\,|f(2u) - f(2v)|.$$

Now $2u, 2v \in I_{j+1} = [2^{j+1}, 2^{j+2})$, so

$$W_\alpha(2u, 2v)\,|f(2u) - f(2v)| \le \delta_{j+1}(f).$$

Taking the supremum over $u, v \in I_j$,

$$\delta_j\big(P_{\text{even}}f; D_1\big) := \sup_{u \neq v \in I_j} W_\alpha(u,v) \, |D_1(u,v)| \; \leq \; 2^{-(1-\alpha)} 2^{-(j+1)} \delta_{j+1}(f).$$

Multiply by ϑ^j and sum over $j \geq 0$:

$$\sum_{j \geq 0} \vartheta^j \delta_j\big(P_{\text{even}}f; D_1\big) \; \leq \; 2^{-(1-\alpha)} 2^{-1} \sum_{j \geq 0} (\vartheta 2^{-1})^j \delta_{j+1}(f).$$

Change index $k = j+1$, so $j = k-1$:

$$= 2^{-(2-\alpha)} \sum_{k \geq 1} (\vartheta 2^{-1})^{k-1} \delta_k(f).$$

Since $(\vartheta 2^{-1})^{k-1} = \vartheta^{k-1} 2^{-(k-1)} \leq \vartheta^{k-1}$ and

$$\sum_{k \geq 1} \vartheta^{k-1} \delta_k(f) = \vartheta^{-1} \sum_{k \geq 1} \vartheta^k \delta_k(f) \leq \vartheta^{-1} [f]_{\text{osc}},$$

we obtain

$$\sum_{j \geq 0} \vartheta^j \delta_j\big(P_{\text{even}}f; D_1\big) \; \leq \; 2^{-(2-\alpha)} \vartheta^{-1} [f]_{\text{osc}}.$$

Thus

$$[P_{\text{even}}f]_{\text{osc}}^{(D_1)} \; \leq \; \lambda_{\text{even}} \, [f]_{\text{osc}}, \qquad \lambda_{\text{even}} := 2^{-(2-\alpha)} \vartheta^{-1}.$$

This is the corrected contraction constant; in particular $\lambda_{\text{even}} < 1$ whenever $\vartheta > 2^{-(2-\alpha)}$.

Step 3: Denominator term D_2: weak control For $u > v$ (the case $u < v$ is symmetric),

$$\left| \frac{1}{2u} - \frac{1}{2v} \right| = \frac{|u-v|}{2uv},$$

so

$$|D_2(u,v)| = |f(2v)| \, \frac{|u-v|}{2uv}.$$

Using

$$W_\alpha(u,v) = \frac{uv}{|u-v|(u+v)^\alpha},$$

we get the exact simplification

$$W_\alpha(u,v) \, |D_2(u,v)| = \frac{|f(2v)|}{2(u+v)^\alpha}.$$

For $u, v \in I_j$ we have $u + v \geq 2 \cdot 2^j = 2^{j+1}$, hence

$$W_\alpha(u,v) \, |D_2(u,v)| \leq 2^{-(1+\alpha)} \, 2^{-\alpha j} \, |f(2v)| =: C_\alpha \, 2^{-\alpha j} \, |f(2v)|.$$

Taking the supremum over $u, v \in I_j$ gives

$$\delta_j\big(P_{\text{even}}f; D_2\big) \leq C_\alpha \, 2^{-\alpha j} \sup_{w \in I_{j+1}} |f(w)|.$$

To bound $\sup_{w \in I_{j+1}} |f(w)|$ via the weak norm, observe that for $w \in I_{j+1}$,

$$|f(w)| \leq w^\sigma \frac{|f(w)|}{w^\sigma} \leq 2^{\sigma(j+1)} \sum_{n \in I_{j+1}} \frac{|f(n)|}{n^\sigma}.$$

By the definition of $[f]_{\mathrm{mass}}$,

$$\sum_{n \in I_{j+1}} \frac{|f(n)|}{n^\sigma} \le \eta^{-(j+1)}[f]_{\mathrm{mass}},$$

so

$$\sup_{w \in I_{j+1}} |f(w)| \le 2^{\sigma(j+1)}\eta^{-(j+1)}[f]_{\mathrm{mass}} = 2^\sigma (2^\sigma \eta^{-1})^j \eta^{-1}[f]_{\mathrm{mass}}.$$

Combining,

$$\delta_j\big(P_{\mathrm{even}}f; D_2\big) \le C_{\alpha,\sigma}\big(2^{-\alpha}2^\sigma\eta^{-1}\big)^j \eta^{-1}[f]_{\mathrm{mass}} = C_{\alpha,\sigma}\big(2^{\sigma-\alpha}\eta^{-1}\big)^j \eta^{-1}[f]_{\mathrm{mass}}.$$

Multiplying by ϑ^j and summing in j,

$$\sum_{j \ge 0} \vartheta^j \delta_j\big(P_{\mathrm{even}}f; D_2\big) \le C_{\alpha,\sigma}\eta^{-1}\sum_{j \ge 0}\big(\vartheta\, 2^{\sigma-\alpha}\eta^{-1}\big)^j [f]_{\mathrm{mass}}.$$

Choosing $\eta > 1$ so that

$$\vartheta\, 2^{\sigma-\alpha}\eta^{-1} < 1,$$

the geometric series converges and we obtain

$$[P_{\mathrm{even}}f]_{\mathrm{osc}}^{(D_2)} \;\le\; C_{\mathrm{even}}^{(2)}\big(\|f\|_\sigma + [f]_{\mathrm{mass}}\big),$$

for some constant $C_{\mathrm{even}}^{(2)} > 0$ depending only on $\alpha, \vartheta, \sigma, \eta$ (we absorb $\|f\|_\sigma$ for later convenience). Adding the contributions of D_1 and D_2,

$$[P_{\mathrm{even}}f]_{\mathrm{osc}} \le [P_{\mathrm{even}}f]_{\mathrm{osc}}^{(D_1)} + [P_{\mathrm{even}}f]_{\mathrm{osc}}^{(D_2)} \le \lambda_{\mathrm{even}}\,[f]_{\mathrm{osc}} + C_{\mathrm{even}}\big(\|f\|_\sigma + [f]_{\mathrm{mass}}\big),$$

with $\lambda_{\mathrm{even}} = 2^{-(2-\alpha)}\vartheta^{-1}$ and $C_{\mathrm{even}} = C_{\mathrm{even}}^{(2)}$, which proves (38). □

The odd branch requires more care because it shifts indices from n to $(n-1)/3$ and only acts on the congruence class $n \equiv 4 \pmod 6$. Its effect is nonetheless small once weighted by ϑ^j.

4.4.2. Odd-Branch Distortion and Contraction Estimates

Lemma 4.8 (Odd–branch distortion on scale blocks). *Let $0 < \alpha < 1$ and $W_\alpha(u,v)$ be the tree weight 35. If $n \equiv 4 \pmod 6$ and $n \in I_j = [6^j, 2 \cdot 6^j)$, then the odd preimage $m = (n-1)/3$ satisfies $m \asymp 6^{j-1}$. Moreover, there exists a constant $C_\alpha > 0$ depending only on α such that for every pair $n_1, n_2 \in I_j$ lying on the same ray and $m_i = (n_i - 1)/3$,*

$$W_\alpha(m_1, m_2) \;\le\; C_\alpha\, W_\alpha(n_1, n_2). \tag{39}$$

Proof. Fix $j \ge 1$ and $n \in I_j = [6^j, 2 \cdot 6^j)$ with $n \equiv 4 \pmod 6$, so $m = (n-1)/3$. Then

$$\frac{6^j - 1}{3} \;\le\; m \;<\; \frac{2 \cdot 6^j - 1}{3},$$

hence for all sufficiently large j we have

$$6^{j-1} \;\le\; m \;\le\; 4 \cdot 6^{j-1},$$

so $m \asymp 6^{j-1}$ as claimed.

Now take $n_1, n_2 \in I_j$ on the same ray and define

$$m_i = \frac{n_i - 1}{3}, \qquad i = 1, 2.$$

Then

$$n_i = 3m_i + 1, \qquad |n_1 - n_2| = 3\,|m_1 - m_2|.$$

We estimate the ratio

$$R := \frac{W_\alpha(m_1, m_2)}{W_\alpha(n_1, n_2)} = \frac{m_1 m_2}{n_1 n_2} \cdot \frac{|n_1 - n_2|}{|m_1 - m_2|} \cdot \frac{(n_1 + n_2)^\alpha}{(m_1 + m_2)^\alpha}.$$

Product factor. From $6^{j-1} \le m_i \le 4 \cdot 6^{j-1}$ and $6^j \le n_i \le 2 \cdot 6^j$ we get

$$m_1 m_2 \;\le\; (4 \cdot 6^{j-1})^2 = 16 \cdot 6^{2(j-1)}, \qquad n_1 n_2 \;\ge\; 6^{2j}.$$

Thus

$$\frac{m_1 m_2}{n_1 n_2} \;\le\; \frac{16 \cdot 6^{2(j-1)}}{6^{2j}} = \frac{16}{6^2} = \frac{4}{9}. \tag{40}$$

Difference factor. We have the exact identity

$$\frac{|n_1 - n_2|}{|m_1 - m_2|} = 3.$$

Sum factor. Again using the scale bounds, for m_i we have

$$2 \cdot 6^{j-1} \;\le\; m_1 + m_2 \;\le\; 8 \cdot 6^{j-1},$$

and for n_i,

$$2 \cdot 6^j \;\le\; n_1 + n_2 \;\le\; 4 \cdot 6^j.$$

Therefore

$$\frac{n_1 + n_2}{m_1 + m_2} \;\le\; \frac{4 \cdot 6^j}{2 \cdot 6^{j-1}} = \frac{4 \cdot 6}{2} = 12,$$

so

$$\frac{(n_1 + n_2)^\alpha}{(m_1 + m_2)^\alpha} \;\le\; 12^\alpha. \tag{41}$$

Combining (40), the difference factor, and (41), we obtain

$$R \le \frac{4}{9} \cdot 3 \cdot 12^\alpha = \frac{4}{3}\,12^\alpha.$$

Thus we may take

$$C_\alpha := \frac{4}{3}\,12^\alpha,$$

which depends only on α and is finite for all $0 < \alpha < 1$. This yields

$$W_\alpha(m_1, m_2) \le C_\alpha\,W_\alpha(n_1, n_2),$$

which is (39). $\quad\square$

Lemma 4.9 (Odd branch on $B_{\mathrm{tree},\sigma}$). *Let $0 < \alpha < 1$, $0 < \vartheta < 1$, $\sigma > 1$, and $\eta > 1$. Consider the odd-branch operator*

$$(P_{\mathrm{odd}} f)(n) = \mathbf{1}_{\{n \equiv 4\,(6)\}}\,\frac{f\!\left(\frac{n-1}{3}\right)}{(n-1)/3}.$$

Then there exist constants $0 < \lambda_{\mathrm{odd}} < 1$ and $C_{\mathrm{odd}} > 0$, depending only on $\alpha, \vartheta, \sigma, \eta$, such that for all $f \in B_{\mathrm{tree},\sigma}$ one has

$$[P_{\mathrm{odd}} f]_{\mathrm{osc}} \;\le\; \lambda_{\mathrm{odd}}\,[f]_{\mathrm{osc}} \;+\; C_{\mathrm{odd}}\big(\,\|f\|_\sigma + [f]_{\mathrm{mass}}\big). \tag{42}$$

In particular, for $\alpha = \frac{1}{2}$ one can take

$$\lambda_{\text{odd}} \ \le \ C_0 \, \vartheta,$$

where $C_0 = \dfrac{16}{3^{3/2}}$ is the odd-branch distortion constant from Lemma 4.14.

Proof. Write

$$\delta_j(g) := \sup_{m \neq n \in I_j} W_\alpha(m,n) \, |g(m) - g(n)|, \qquad [g]_{\text{osc}} = \sum_{j \geq 0} \vartheta^j \delta_j(g),$$

so that $[P_{\text{odd}}f]_{\text{osc}} = \sum_{j \geq 0} \vartheta^j \delta_j(P_{\text{odd}}f)$. We bound $\delta_j(P_{\text{odd}}f)$ in terms of $\delta_k(f)$ at nearby scales and the weak/mass terms.

Fix $j \geq 0$ and $m, n \in I_j = [6^j, 2 \cdot 6^j)$ with $m \neq n$. We split into three cases.

Case 1: neither m nor n is $4 \pmod 6$. Then $P_{\text{odd}}f(m) = P_{\text{odd}}f(n) = 0$, so this pair contributes nothing to $\delta_j(P_{\text{odd}}f)$.

Case 2: exactly one of m, n is $4 \pmod 6$. Without loss of generality, assume $m \equiv 4 \pmod 6$ and $n \not\equiv 4 \pmod 6$, and set $k := (m-1)/3$. Then

$$P_{\text{odd}}f(m) - P_{\text{odd}}f(n) = \frac{f(k)}{k},$$

and therefore

$$W_\alpha(m,n) \, \big| P_{\text{odd}}f(m) - P_{\text{odd}}f(n) \big| = W_\alpha(m,n) \, \frac{|f(k)|}{k}.$$

As in the previous estimates for the tree seminorm, one has $W_\alpha(m,n) \ll 6^{(2-\alpha)j}$ and $k \asymp 6^{j-1}$, so

$$W_\alpha(m,n) \, \frac{|f(k)|}{k} \ \ll \ 6^{(1-\alpha)j} \, |f(k)|.$$

Multiplying by ϑ^j and summing over j, one obtains a contribution bounded by

$$\sum_{\substack{j \geq 0}} \vartheta^j \sup_{\substack{m,n \in I_j \\ \text{exactly one} \equiv 4 \, (6)}} W_\alpha(m,n) \, \big| P_{\text{odd}}f(m) - P_{\text{odd}}f(n) \big| \ \le \ C_{\text{odd},1} \big(\|f\|_\sigma + [f]_{\text{mass}} \big),$$

for a suitable constant $C_{\text{odd},1}$ depending on $\alpha, \vartheta, \sigma, \eta$. Thus Case 2 contributes only to the weak/mass error term in (42).

Case 3: both m and n are $4 \pmod 6$. Set

$$m' = \frac{m-1}{3}, \qquad n' = \frac{n-1}{3},$$

so that $m', n' \in I_{j-1}$ and

$$P_{\text{odd}}f(m) = \frac{f(m')}{m'}, \qquad P_{\text{odd}}f(n) = \frac{f(n')}{n'}.$$

We decompose

$$\frac{f(m')}{m'} - \frac{f(n')}{n'} = \frac{f(m') - f(n')}{m'} + f(n') \Big(\frac{1}{m'} - \frac{1}{n'} \Big) =: D_1 + D_2.$$

Case 3a: the D_1 term (oscillatory contribution). Here

$$D_1 = \frac{f(m') - f(n')}{m'},$$

and we want to control

$$W_\alpha(m,n) \, |D_1| = W_\alpha(m,n) \, \frac{|f(m') - f(n')|}{m'}.$$

For $\alpha = \frac{1}{2}$, Lemma 4.14 gives the two-sided distortion bound

$$\frac{1}{C_0} \leq \frac{W_{1/2}(m,n)}{W_{1/2}(m',n')} \leq C_0, \qquad C_0 = \frac{16}{3^{3/2}}.$$

Using the upper bound, we obtain

$$W_{1/2}(m,n)\,|D_1| \leq C_0\,W_{1/2}(m',n')\,\frac{|f(m') - f(n')|}{m'}.$$

Since $m' \asymp 6^{j-1}$ for $m \in I_j$, we have $1/m' \ll 6^{-(j-1)}$, so

$$W_{1/2}(m,n)\,|D_1| \ll C_0\,6^{-(j-1)}\,W_{1/2}(m',n')\,|f(m') - f(n')|.$$

Taking the supremum over $m, n \in I_j$ with $m \equiv n \equiv 4 \pmod 6$ gives

$$\delta_j(P_{\mathrm{odd}}f; D_1) \leq C_0'\,6^{-(j-1)}\,\delta_{j-1}(f),$$

for some constant C_0' comparable to C_0.

Multiplying by ϑ^j and summing over $j \geq 1$ yields

$$\sum_{j \geq 1} \vartheta^j \delta_j(P_{\mathrm{odd}}f; D_1) \leq C_0'\,\vartheta \sum_{k \geq 0} (\vartheta/6)^k \delta_k(f) \leq C_0''\,\vartheta\,[f]_{\mathrm{osc}},$$

with C_0'' depending only on C_0 and ϑ (and uniformly bounded as long as $\vartheta < 6$). Thus the D_1-part of Case 3 contributes a term of the form $\lambda_{\mathrm{odd}}^{(1)}[f]_{\mathrm{osc}}$ with $\lambda_{\mathrm{odd}}^{(1)} \leq C_0''\vartheta$.

Case 3b: the D_2 term (denominator contribution). For

$$D_2 = f(n')\left(\frac{1}{m'} - \frac{1}{n'}\right)$$

one uses again the scale relations $m', n' \asymp 6^{j-1}$ and the fact that $|m' - n'| \leq C6^{j-1}$ to deduce

$$W_\alpha(m,n)\,|D_2| \ll 6^{-\alpha j}\,|f(n')|.$$

As in the even-branch analysis, the amplitude $|f(n')|$ on blocks I_{j-1} is controlled by a combination of the weak norm $\|f\|_\sigma$ and the mass seminorm $[f]_{\mathrm{mass}}$, and the geometric factor $6^{-\alpha j}$ ensures convergence of the sum over j. Thus there exists a constant $C_{\mathrm{odd},2} > 0$ such that

$$\sum_{j \geq 0} \vartheta^j \delta_j(P_{\mathrm{odd}}f; D_2) \leq C_{\mathrm{odd},2}\big(\|f\|_\sigma + [f]_{\mathrm{mass}}\big).$$

Conclusion. Combining Cases 1–3, we obtain

$$[P_{\mathrm{odd}}f]_{\mathrm{osc}} = \sum_{j \geq 0} \vartheta^j \delta_j(P_{\mathrm{odd}}f) \leq \lambda_{\mathrm{odd}}\,[f]_{\mathrm{osc}} + C_{\mathrm{odd}}\big(\|f\|_\sigma + [f]_{\mathrm{mass}}\big),$$

with

$$\lambda_{\mathrm{odd}} := \lambda_{\mathrm{odd}}^{(1)} \leq C_0''\,\vartheta, \qquad C_{\mathrm{odd}} := C_{\mathrm{odd},1} + C_{\mathrm{odd},2},$$

depending only on $\alpha, \vartheta, \sigma, \eta$. For $\alpha = \frac{1}{2}$ the explicit distortion constant $C_0 = 16/3^{3/2}$ from Lemma 4.14 can be absorbed into C_0'', and in particular one may choose $\lambda_{\mathrm{odd}} \leq C_0\vartheta$ after adjusting constants. Choosing $\vartheta > 0$ small enough so that $\lambda_{\mathrm{odd}} < 1$ gives the desired Lasota–Yorke contraction. $\quad\square$

4.5. Derivation of the Global Lasota–Yorke Inequality

Lemma 4.10 (Invariance and boundedness on $B_{\text{tree},\sigma}$). *Let $0 < \alpha < 1$, $\sigma > 1$, and $0 < \vartheta < 1$ satisfy*

$$\vartheta \, 6^{\,1-\alpha+\sigma} < 1. \tag{43}$$

Then the backward Collatz transfer operator P maps $B_{\text{tree},\sigma}$ into itself and is bounded: there exists $C > 0$ such that

$$\|Pf\|_{\text{tree},\sigma} \le C \, \|f\|_{\text{tree},\sigma} \qquad \text{for all } f \in B_{\text{tree},\sigma},$$

where

$$\|f\|_{\text{tree},\sigma} := [f]_{\text{tree}} + \|f\|_\sigma.$$

Proof. Recall the even/odd decomposition

$$(Pf)(n) = (P_{\text{even}}f)(n) + (P_{\text{odd}}f)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4 \ (6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}.$$

Step 1: Weighted ℓ^1_σ bound. The weak norm estimate is given by the same computation as in the weighted ℓ^1_σ lemma: for all $f \in B_{\text{tree},\sigma}$,

$$\|Pf\|_\sigma = \|P_{\text{even}}f + P_{\text{odd}}f\|_\sigma \le (2^\sigma + 3^{-\sigma}) \, \|f\|_\sigma. \tag{44}$$

This uses only the change of variables $m = 2n$ for the even branch and $m = (n-1)/3$ for the odd branch.

Step 2: Tree seminorm bound via the even/odd lemmas. By subadditivity of $[\,\cdot\,]_{\text{tree}}$,

$$[Pf]_{\text{tree}} \le [P_{\text{even}}f]_{\text{tree}} + [P_{\text{odd}}f]_{\text{tree}}.$$

The even-branch lemma (even branch on $B_{\text{tree},\sigma}$) states that under the admissibility condition (43) there exists $C_{\text{even}} > 0$, depending only on $\alpha, \sigma, \vartheta$, such that

$$[P_{\text{even}}f]_{\text{tree}} \le C_{\text{even}} \|f\|_\sigma \qquad \text{for all } f \in B_{\text{tree},\sigma}. \tag{45}$$

The proof is obtained by estimating each block seminorm $\Delta_j(P_{\text{even}}f)$ and summing the geometric series $\sum_{j \ge 0} (\vartheta 6^{\,1-\alpha+\sigma})^j$, which converges precisely when (43) holds.

Similarly, the odd-branch lemma (odd branch on $B_{\text{tree},\sigma}$) gives a constant $C_{\text{odd}} > 0$, depending only on $\alpha, \sigma, \vartheta$, such that

$$[P_{\text{odd}}f]_{\text{tree}} \le C_{\text{odd}} \|f\|_\sigma \qquad \text{for all } f \in B_{\text{tree},\sigma}, \tag{46}$$

again using the same geometric factor $\vartheta 6^{\,1-\alpha+\sigma}$ and the admissibility (43).

Combining (45) and (46) we obtain

$$[Pf]_{\text{tree}} \le (C_{\text{even}} + C_{\text{odd}}) \|f\|_\sigma. \tag{47}$$

Step 3: Combine strong and weak parts. By definition,

$$\|Pf\|_{\text{tree},\sigma} = [Pf]_{\text{tree}} + \|Pf\|_\sigma.$$

Using (44) and (47),

$$\|Pf\|_{\text{tree},\sigma} \le (C_{\text{even}} + C_{\text{odd}}) \|f\|_\sigma + (2^\sigma + 3^{-\sigma}) \|f\|_\sigma.$$

Since $\|f\|_{\text{tree},\sigma} = [f]_{\text{tree}} + \|f\|_\sigma \geq \|f\|_\sigma$, we have

$$\|f\|_\sigma \ \leq \ \|f\|_{\text{tree},\sigma},$$

and therefore

$$\|Pf\|_{\text{tree},\sigma} \leq \left(C_{\text{even}} + C_{\text{odd}} + 2^\sigma + 3^{-\sigma}\right) \|f\|_{\text{tree},\sigma}.$$

Thus P is bounded on $B_{\text{tree},\sigma}$ with operator norm at most

$$C := C_{\text{even}} + C_{\text{odd}} + 2^\sigma + 3^{-\sigma},$$

which depends only on $\alpha, \sigma, \vartheta$. $\square$

Proposition 4.11 (Lasota–Yorke inequality on $B_{\text{tree},\sigma}$). *Let* $0 < \alpha < 1$, $0 < \vartheta < 1$, *and* $\sigma > 1$ *satisfy the admissibility condition*

$$\vartheta\, 6^{1-\alpha+\sigma} < 1. \tag{48}$$

Then there exists a constant $C_{\text{LY},\sigma} > 0$ *such that for all* $f \in B_{\text{tree},\sigma}$,

$$[Pf]_{\text{tree}} \ \leq \ C_{\text{LY},\sigma}\, \|f\|_\sigma. \tag{49}$$

In particular, for every $n \geq 1$ *one has*

$$[P^n f]_{\text{tree}} \ \leq \ C_{\text{LY},\sigma}\, \|P^{n-1}f\|_\sigma \ \leq \ C_{\text{LY},\sigma}\, (2^\sigma + 3^{-\sigma})^{\,n-1} \|f\|_\sigma, \qquad f \in B_{\text{tree},\sigma}, \tag{50}$$

so each iterate P^n *is smoothing in the strong seminorm* $[\,\cdot\,]_{\text{tree}}$ *with a quantitative bound in terms of the weak norm* $\|\cdot\|_\sigma$ *of the previous iterate.*

Proof. We use the even/odd decomposition $P = P_{\text{even}} + P_{\text{odd}}$. From the estimates established in the proofs of the even and odd branch lemmas, and under the admissibility condition (48), there exist constants $C_{\text{even}}, C_{\text{odd}} > 0$ such that for all $f \in B_{\text{tree},\sigma}$,

$$[P_{\text{even}}f]_{\text{tree}} \ \leq \ C_{\text{even}}\, \|f\|_\sigma, \qquad [P_{\text{odd}}f]_{\text{tree}} \ \leq \ C_{\text{odd}}\, \|f\|_\sigma. \tag{51}$$

The convergence of the geometric series $\sum_{j\geq 0}(\vartheta\, 6^{1-\alpha+\sigma})^j$ is precisely what guarantees that the constants C_{even} and C_{odd} are finite.

By subadditivity of $[\,\cdot\,]_{\text{tree}}$,

$$[Pf]_{\text{tree}} = [P_{\text{even}}f + P_{\text{odd}}f]_{\text{tree}} \ \leq \ [P_{\text{even}}f]_{\text{tree}} + [P_{\text{odd}}f]_{\text{tree}} \ \leq \ (C_{\text{even}} + C_{\text{odd}})\, \|f\|_\sigma.$$

Setting

$$C_{\text{LY},\sigma} := C_{\text{even}} + C_{\text{odd}}$$

gives (49).

For the iterates, apply (49) with f replaced by $P^{n-1}f$:

$$[P^n f]_{\text{tree}} \ = \ [P(P^{n-1}f)]_{\text{tree}} \ \leq \ C_{\text{LY},\sigma}\, \|P^{n-1}f\|_\sigma.$$

From the weighted ℓ^1_σ bound

$$\|Pf\|_\sigma \ \leq \ (2^\sigma + 3^{-\sigma})\, \|f\|_\sigma$$

(see (44)), iterating gives $\|P^{n-1}f\|_\sigma \leq (2^\sigma + 3^{-\sigma})^{\,n-1}\|f\|_\sigma$, which yields (50). $\square$

Remark 4.12 (Parameter window). *The lift from $\| \cdot \|_1$ to $\| \cdot \|_\sigma$ in the remainder terms occurs through the same geometric factor that appears in the proofs of the even and odd branch bounds, namely $\vartheta\, 6^{1-\alpha+\sigma}$. The only requirement is the admissibility condition*

$$\vartheta\, 6^{1-\alpha+\sigma} < 1,$$

which ensures convergence of $\sum_{j\geq 0}(\vartheta\, 6^{1-\alpha+\sigma})^j$ and hence finiteness of C_{even} and C_{odd}.

For fixed α and σ this simply means

$$0 < \vartheta < 6^{-(1-\alpha+\sigma)}.$$

A convenient concrete choice (used later) is

$$(\alpha, \vartheta, \sigma) = \left(\tfrac{1}{2}, \tfrac{1}{20}, 1+\varepsilon\right) \quad \text{with any sufficiently small } \varepsilon > 0,$$

since then

$$\vartheta\, 6^{1-\alpha+\sigma} = \frac{1}{20}\, 6^{\frac{3}{2}+\varepsilon} < 1$$

for all ε in a small interval $(0, \varepsilon_0]$. Together with the explicit odd–branch distortion constant computed in Section 5 and the compact embedding $B_{\mathrm{tree},\sigma} \hookrightarrow \ell^1_\sigma$, this smoothing Lasota–Yorke inequality is exactly what is needed to apply the Ionescu–Tulcea–Marinescu/Hennion theory and deduce quasi–compactness of P on $B_{\mathrm{tree},\sigma}$.

Corollary 4.13 (Essential spectral radius on $B_{\mathrm{tree},\sigma}$). *Let $0 < \alpha < 1$, $0 < \vartheta < 1$, and $\sigma > 1$ satisfy the admissibility condition (43). Assume the Lasota–Yorke inequality from Proposition 4.11 and the compact embedding $B_{\mathrm{tree},\sigma} \hookrightarrow \ell^1_\sigma$ (from Lemma 4.5). Then $P : B_{\mathrm{tree},\sigma} \to B_{\mathrm{tree},\sigma}$ is quasi–compact, and its essential spectral radius satisfies*

$$\rho_{\mathrm{ess}}\left(P\!\upharpoonright_{B_{\mathrm{tree},\sigma}}\right) = 0. \tag{52}$$

Proof. By Proposition 4.11, for every $f \in B_{\mathrm{tree},\sigma}$,

$$[Pf]_{\mathrm{tree}} \leq C_{\mathrm{LY},\sigma}\, \|f\|_\sigma,$$

where $[\,\cdot\,]_{\mathrm{tree}}$ is the strong seminorm and $\|\cdot\|_\sigma$ is the weak norm on $B_{\mathrm{tree},\sigma}$. This is a Lasota–Yorke (Doeblin–Fortet) inequality with *strong contraction coefficient* equal to 0, i.e.

$$\|Pf\|_{\mathrm{strong}} := [Pf]_{\mathrm{tree}} \leq 0 \cdot [f]_{\mathrm{tree}} + C_{\mathrm{LY},\sigma}\, \|f\|_{\mathrm{weak}}, \qquad \|f\|_{\mathrm{weak}} := \|f\|_\sigma.$$

By Lemma 4.5, the unit ball of $B_{\mathrm{tree},\sigma}$ is relatively compact in ℓ^1_σ, so the inclusion

$$B_{\mathrm{tree},\sigma} \hookrightarrow \ell^1_\sigma$$

is compact. Thus the hypotheses of the Ionescu–Tulcea–Marinescu/Hennion quasi–compactness theorem are satisfied with strong contraction constant $a = 0$. The theorem then implies that P is quasi–compact on $B_{\mathrm{tree},\sigma}$ and that its essential spectral radius is bounded above by a, i.e.

$$\rho_{\mathrm{ess}}\left(P\!\upharpoonright_{B_{\mathrm{tree},\sigma}}\right) \leq 0.$$

Since the essential spectral radius is always nonnegative, we conclude

$$\rho_{\mathrm{ess}}\left(P\!\upharpoonright_{B_{\mathrm{tree},\sigma}}\right) = 0,$$

which is (52). $\square$

4.6. Quasi-Compactness and Spectral Gap for P

Lemma 4.14 (Odd–branch weight distortion at $\alpha = \frac{1}{2}$). *Let $W_\alpha(m, n)$ be the tree weight 35 and let $m' = (m - 1)/3$, $n' = (n - 1)/3$. For $\alpha = \frac{1}{2}$ there exists an absolute constant*

$$C_0 = \frac{16}{3^{3/2}} < 3.1$$

such that for all $m \equiv n \equiv 4 \pmod 6$, $m \neq n$, one has the two–sided distortion bounds

$$\frac{1}{C_0} \leq \frac{W_{1/2}(m, n)}{W_{1/2}(m', n')} \leq C_0. \tag{53}$$

Consequently, the oscillatory part of the odd branch satisfies

$$\lambda_{\text{odd}}\left(\tfrac{1}{2}, \vartheta\right) \leq \frac{C_0}{\sqrt{6}} \vartheta,$$

as used in Lemma 4.9 and Lemma 4.15.

Proof. Plug in $\alpha = \frac{1}{2}$ into Lemma 4.8 and the result follows from the Lemma. □

Lemma 4.15 (Explicit odd-branch constant). *For $\alpha = \frac{1}{2}$ and $\vartheta = \frac{1}{20}$ there exist constants $C_\alpha > 0$ and $C_{\text{odd}} > 0$ such that for all $f \in B_{\text{tree},\sigma}$,*

$$[P_{\text{odd}}f]_{\text{tree}} \leq \lambda_{\text{odd}}(\alpha, \vartheta) [f]_{\text{tree}} + C_{\text{odd}} \|f\|_\sigma, \tag{54}$$

with

$$\lambda_{\text{odd}}(\alpha, \vartheta) \leq \frac{C_\alpha}{\sqrt{6}} \vartheta < 1. \tag{55}$$

Proof. This is a direct specialization of Lemma 4.9 to the parameter choice $\alpha = \frac{1}{2}$, $\vartheta = \frac{1}{20}$. Lemma 4.9 gives the Lasota–Yorke type estimate

$$[P_{\text{odd}}f]_{\text{tree}} \leq \lambda_{\text{odd}}(\alpha, \vartheta) [f]_{\text{tree}} + C_{\text{odd}} \|f\|_\sigma,$$

with

$$\lambda_{\text{odd}}(\alpha, \vartheta) \leq \frac{C_\alpha}{\sqrt{6}} \vartheta,$$

where $C_\alpha > 0$ is the odd-branch distortion constant defined in Lemma 4.14. For $\alpha = \frac{1}{2}$, Lemma 4.14 gives the explicit value

$$C_\alpha = C_0 = \frac{16}{3^{3/2}} < 3.1.$$

Substituting $\alpha = \frac{1}{2}$ and $\vartheta = \frac{1}{20}$ into the above bound,

$$\lambda_{\text{odd}}\left(\tfrac{1}{2}, \tfrac{1}{20}\right) \leq \frac{C_0}{\sqrt{6}} \cdot \frac{1}{20} < \frac{3.1}{\sqrt{6}} \cdot \frac{1}{20} < 1.$$

Thus (54) holds with $\lambda_{\text{odd}}(\frac{1}{2}, \frac{1}{20}) < 1$ and some constant $C_{\text{odd}} > 0$ depending only on σ and the block geometry, as asserted. □

Proposition 4.16 (Verified Lasota–Yorke contraction). *Let $(\alpha, \vartheta) = \left(\frac{1}{2}, \frac{1}{20}\right)$ and $\sigma > 1$ satisfy the admissibility condition (43), i.e.*

$$\vartheta \, 6^{1-\alpha+\sigma} < 1.$$

Define

$$\lambda_{\text{LY}} := 2^{-(1-\alpha)} \vartheta + \lambda_{\text{odd}}(\alpha, \vartheta), \qquad \lambda_{\text{odd}}(\alpha, \vartheta) \leq \frac{C_0}{\sqrt{6}} \vartheta,$$

with $C_0 = 16/3^{3/2}$ *from Lemma 4.14. Then* $\lambda_{\text{LY}} < 1$, *and for all* $f \in B_{\text{tree},\sigma}$,

$$[Pf]_{\text{tree}} \leq \lambda_{\text{LY}} [f]_{\text{tree}} + C_{\text{LY}} \|f\|_\sigma, \tag{56}$$

for some constant $C_{\text{LY}} > 0$ *depending only on the fixed parameters and the block geometry.*

Proof. We use the decomposition $P = P_{\text{even}} + P_{\text{odd}}$ and the branchwise Lasota–Yorke estimates already established.

(1) Combine even and odd branch inequalities. For any $f \in B_{\text{tree},\sigma}$,

$$[Pf]_{\text{tree}} \leq [P_{\text{even}}f]_{\text{tree}} + [P_{\text{odd}}f]_{\text{tree}}.$$

By the even-branch estimate (Lemma 4.7, specialized to $B_{\text{tree},\sigma}$ and the fixed parameters $\alpha = \frac{1}{2}$, $\vartheta = \frac{1}{20}$), there exists $C_{\text{even}} > 0$ such that

$$[P_{\text{even}}f]_{\text{tree}} \leq 2^{-(1-\alpha)} \vartheta [f]_{\text{tree}} + C_{\text{even}} \|f\|_\sigma. \tag{57}$$

By the explicit odd-branch lemma (Lemma 4.15) for $\alpha = \frac{1}{2}$ and $\vartheta = \frac{1}{20}$, there exist $C_\alpha > 0$ and $C_{\text{odd}} > 0$ such that

$$[P_{\text{odd}}f]_{\text{tree}} \leq \lambda_{\text{odd}}(\alpha, \vartheta) [f]_{\text{tree}} + C_{\text{odd}} \|f\|_\sigma, \tag{58}$$

with

$$\lambda_{\text{odd}}(\alpha, \vartheta) \leq \frac{C_\alpha}{\sqrt{6}} \vartheta = \frac{C_0}{\sqrt{6}} \vartheta < 1, \tag{59}$$

where $C_0 = 16/3^{3/2}$ is given by Lemma 4.14.

Adding (57) and (58),

$$[Pf]_{\text{tree}} \leq \left(2^{-(1-\alpha)} \vartheta + \lambda_{\text{odd}}(\alpha, \vartheta)\right)[f]_{\text{tree}} + (C_{\text{even}} + C_{\text{odd}})\|f\|_\sigma.$$

Define

$$\lambda_{\text{LY}} := 2^{-(1-\alpha)} \vartheta + \lambda_{\text{odd}}(\alpha, \vartheta), \qquad C_{\text{LY}} := C_{\text{even}} + C_{\text{odd}},$$

and we obtain (56).

(2) Verification that $\lambda_{\text{LY}} < 1$. With $(\alpha, \vartheta) = (\frac{1}{2}, \frac{1}{20})$,

$$2^{-(1-\alpha)} \vartheta = 2^{-1/2} \cdot \frac{1}{20} = \frac{1}{20\sqrt{2}} \approx 0.0354.$$

From Lemma 4.15 and Lemma 4.14 we have

$$\lambda_{\text{odd}}\left(\tfrac{1}{2}, \tfrac{1}{20}\right) \leq \frac{C_0}{\sqrt{6}} \cdot \frac{1}{20}, \qquad C_0 = \frac{16}{3^{3/2}} < 3.1.$$

Numerically,

$$\frac{C_0}{\sqrt{6}} \cdot \frac{1}{20} < \frac{3.1}{\sqrt{6}} \cdot \frac{1}{20} \approx 0.063.$$

Therefore

$$\lambda_{\text{LY}} = 2^{-(1-\alpha)} \vartheta + \lambda_{\text{odd}}(\alpha, \vartheta) < 0.0354 + 0.063 < 0.1 < 1.$$

Thus λ_{LY} is a strict contraction factor depending only on the fixed parameters, and the proposition follows. $\square$

We now record the standard consequence of the Lasota–Yorke inequality and the compact embedding of B_{tree} into ℓ^1.

Theorem 4.17 (Quasi-compactness on $B_{\text{tree},\sigma}$). *Let $0 < \alpha < 1$, $0 < \vartheta < 1$, and $\sigma > 1$. Assume that the Lasota–Yorke constant*

$$\lambda(\alpha, \vartheta) := 2^{-(1-\alpha)} + \lambda_{\text{odd}}(\alpha, \vartheta)$$

satisfies $\lambda(\alpha, \vartheta) < 1$, where $\lambda_{\text{odd}}(\alpha, \vartheta)$ is as in Lemma 4.9. Then the backward transfer operator P acting on $B_{\text{tree},\sigma}$ is quasi-compact, and its essential spectral radius satisfies

$$\rho_{\text{ess}}(P|_{B_{\text{tree},\sigma}}) \leq \lambda(\alpha, \vartheta) < 1. \tag{60}$$

Proof. We work on the Banach space $B_{\text{tree},\sigma}$ with norm $\|\cdot\|_{\text{tree},\sigma} = \|\cdot\|_\sigma + [\cdot]_{\text{tree}}$, where $\|\cdot\|_\sigma$ is the weighted ℓ^1_σ-norm and $[\cdot]_{\text{tree}}$ is the tree seminorm defined in Section 4.3.

Step 1: Lasota–Yorke inequality. By Proposition 4.11 (applied in the weighted setting, with $\|f\|_1$ replaced by $\|f\|_\sigma$) we have, for all $f \in B_{\text{tree},\sigma}$,

$$[Pf]_{\text{tree}} \leq \lambda(\alpha, \vartheta)\,[f]_{\text{tree}} + C_{\text{LY}}\|f\|_\sigma, \tag{61}$$

with $\lambda(\alpha, \vartheta) < 1$ by assumption. On the weak norm side, since P is bounded on ℓ^1_σ, there exists $C_\sigma > 0$ (e.g. $C_\sigma = \Lambda_\sigma$ from (17)) such that

$$\|Pf\|_\sigma \leq C_\sigma \|f\|_\sigma \qquad \text{for all}\, f \in B_{\text{tree},\sigma}. \tag{62}$$

Thus P satisfies a standard two-norm Lasota–Yorke inequality on $B_{\text{tree},\sigma}$ with strong seminorm $\|\cdot\|_s := [\cdot]_{\text{tree}}$ and weak norm $\|\cdot\|_w := \|\cdot\|_\sigma$:

$$\|Pf\|_s \leq \lambda \|f\|_s + C_{\text{LY}}\|f\|_w, \qquad \|Pf\|_w \leq C_\sigma \|f\|_w. \tag{63}$$

Step 2: Compact embedding. By Lemma 4.5, the embedding

$$J : (B_{\text{tree},\sigma}, \|\cdot\|_{\text{tree},\sigma}) \hookrightarrow (\ell^1_\sigma, \|\cdot\|_\sigma)$$

is compact. Since $\|\cdot\|_w = \|\cdot\|_\sigma$ is exactly the weak norm used in (63), this shows that the unit ball of $B_{\text{tree},\sigma}$ is relatively compact for the weak norm.

Step 3: Application of Ionescu–Tulcea–Marinescu / Hennion. We now invoke the standard quasi-compactness criterion (Ionescu–Tulcea and Marinescu theorem): if a bounded operator T on a Banach space X satisfies

(i)   a Lasota–Yorke inequality $\|Tx\|_s \leq \lambda \|x\|_s + C\|x\|_w$ with $\lambda < 1$,

(ii) a weak bound $\|Tx\|_w \leq C'\|x\|_w$, and

(iii) the injection $(X, \|\cdot\|_s) \hookrightarrow (X, \|\cdot\|_w)$ has relatively compact unit ball,

then T is quasi-compact on X and its essential spectral radius satisfies

$$\rho_{\text{ess}}(T) \leq \lambda.$$

Conditions (i)–(iii) are exactly (63) and Lemma 4.5 for $T = P$ and $X = B_{\text{tree},\sigma}$. Therefore P is quasi-compact on $B_{\text{tree},\sigma}$ and

$$\rho_{\text{ess}}(P|_{B_{\text{tree},\sigma}}) \leq \lambda(\alpha, \vartheta) < 1,$$

which is (60). $\square$

Remark 4.18 (On the choice of parameters). *The explicit bound (59) shows that $\lambda_{\text{odd}}(\alpha, \vartheta)$ decreases linearly with ϑ. For fixed α, one can therefore choose ϑ sufficiently small so that $\lambda(\alpha, \vartheta) < 1$, provided the constant C_α is*

effectively controlled. Subsequent sections make this optimization quantitative by computing C_α and exhibiting admissible parameter pairs (α, ϑ) that give a strict spectral gap.

The Lasota–Yorke framework developed here supplies the functional-analytic backbone for the spectral approach to the Collatz problem: once explicit parameters with $\lambda(\alpha, \vartheta) < 1$ are verified, the quasi-compactness and spectral gap of P on B_{tree} follow, and the spectral criteria of Section 4 can be invoked to constrain or rule out non-terminating configurations.

5. Invariant Profiles, Block Recursion, and Perron–Frobenius Rigidity

Having established in Section 4.4 that the backward Collatz operator P is quasi-compact on the multi-scale tree space B_{tree}, we now turn to the spectral consequences of this result. The Lasota–Yorke inequality ensures the existence of a spectral gap, which in turn controls the structure of invariant densities and the long-term behavior of iterates P^k. The objective of this section is to characterize the invariant and quasi-invariant components of P, derive an effective block recursion for their scale-averaged coefficients, and demonstrate that the recursion enforces rigidity across the Collatz tree.

Throughout this section, $h \in B_{\text{tree},\sigma}$ will denote an invariant density of P, i.e. a function satisfying $Ph = h$. The analysis proceeds in several stages. First, we describe the structure of possible invariant profiles in the multiscale framework and show that the Lasota–Yorke inequality forces uniform flatness across scales. Next, we translate this flatness into an explicit two-sided recurrence relation for block averages c_j. Finally, we verify that the coefficients of this recurrence satisfy a spectral bound consistent with the contraction constant $\lambda_{\text{odd}}(\alpha, \vartheta)$ computed earlier.

Theorem 5.1 (Perron–Frobenius structure on $B_{\text{tree},\sigma}$). *Let P be the backward Collatz transfer operator acting on $B_{\text{tree},\sigma}$ with parameters $(\alpha, \vartheta, \sigma)$ chosen so that the Lasota–Yorke inequality and quasi–compactness hold. Then:*

1. *The spectral radius of P equals 1, and 1 is a simple eigenvalue.*
2. *There exists a unique eigenvector $h \in B_{\text{tree},\sigma}$ with $h > 0$ and $Ph = h$, normalized by $\phi(h) = 1$.*
3. *There exists a unique positive eigenfunctional $\phi \in B^*_{\text{tree},\sigma}$ such that $\phi \circ P = \phi$.*
4. *All other spectral values satisfy $|z| < 1$, and P admits the spectral decomposition*

$$P = h \otimes \phi + Q, \qquad \rho(Q) < 1,$$

where Q is quasi–compact.

Proof. We combine the Lasota–Yorke inequality on $B_{\text{tree},\sigma}$ with standard Perron–Frobenius theory for positive quasi–compact operators.

Step 1: Spectral radius and quasi–compactness. By construction P is a bounded linear operator on $B_{\text{tree},\sigma}$ and is positive in the sense that $f \geq 0$ implies $Pf \geq 0$. The Lasota–Yorke inequality on $B_{\text{tree},\sigma}$ (Proposition 4.11, say) together with the compact embedding of the strong seminorm into the weak norm implies that P is quasi–compact on $B_{\text{tree},\sigma}$ with essential spectral radius strictly less than 1:

$$\rho_{\text{ess}}(P) < 1. \tag{64}$$

On the other hand, the logarithmic mass–preservation identity (Lemma 2.4) shows that the spectral radius of P is at least 1; the boundedness of P implies $\rho(P) \leq 1$, hence

$$\rho(P) = 1. \tag{65}$$

In particular, 1 lies in the spectrum of P and, by (64), is an isolated spectral value.

Step 2: Existence of a positive eigenvector. Consider the positive cone

$$\mathcal{C} := \{f \in B_{\text{tree},\sigma} : f \geq 0\},$$

which is closed, convex, and reproducing. Since P is positive and $\rho(P) = 1$, the Krein–Rutman theorem for positive operators on Banach spaces implies the existence of a nonzero $h \in \mathcal{C}$ such that

$$Ph = h. \tag{66}$$

Moreover, h can be chosen strictly positive in the sense that $h(n) > 0$ for all $n \in \mathbb{N}$: indeed, by the preimage structure of the Collatz map (Lemma 2.3) and the connectivity of the backward tree, any nontrivial $f \in \mathcal{C}$ is eventually propagated by iterates of P to a function that is positive on every block I_j, so $P^k f > 0$ for all sufficiently large k. Replacing h by $P^k h$ if necessary yields $h > 0$.

Step 3: Uniqueness and simplicity of the eigenvalue 1. We now show that 1 is a simple eigenvalue and that h is unique up to scalar multiples. Suppose $g \in B_{\text{tree},\sigma}$ satisfies $Pg = g$. Decompose $g = g^+ - g^-$ into positive parts. Positivity of P implies $Pg^\pm = g^\pm$. By the strong positivity argument above, any nonzero $f \in \mathcal{C}$ with $Pf = f$ must be strictly positive; hence g^+ and g^- are both either 0 or strictly positive. If both were nonzero, then g^+ and g^- would be linearly independent positive eigenvectors for the eigenvalue 1, and the positive cone would contain a two-dimensional face of eigenvectors. This contradicts the Krein–Rutman conclusion that the eigenspace associated with the spectral radius is one–dimensional. Therefore one of g^+, g^- must vanish and g is either nonnegative or nonpositive; by replacing g by $-g$ if necessary, $g \geq 0$, and the strong positivity then forces g to be a scalar multiple of h. Thus the eigenspace for the eigenvalue 1 is one–dimensional and spanned by h, and 1 is a simple eigenvalue. This proves (1) and the first part of (2) after normalizing by $\phi(h) = 1$ below.

Step 4: Dual eigenfunctional. Consider the dual operator P^* acting on $B_{\text{tree},\sigma}^*$. Since P is positive, so is P^* on the dual cone

$$\mathcal{C}^* := \{\Lambda \in B_{\text{tree},\sigma}^* : \Lambda(f) \geq 0 \text{ for all } f \in \mathcal{C}\}.$$

The quasi–compactness of P implies quasi–compactness of P^* on the dual space. By (65), P^* also has spectral radius 1. Applying the same Krein–Rutman argument to P^* yields a nonzero $\phi \in \mathcal{C}^*$ and

$$\phi \circ P = \phi, \tag{67}$$

with ϕ strictly positive on nonzero elements of $\mathcal{C}$. The same simplicity argument as in Step 3 shows that the eigenspace of P^* for the eigenvalue 1 is one–dimensional and spanned by ϕ. Normalizing by the condition $\phi(h) = 1$ gives the uniquely determined eigenpair (h, ϕ) appearing in the statement. This establishes (2) and (3).

Step 5: Spectral decomposition and spectral gap. Quasi–compactness of P on $B_{\text{tree},\sigma}$, together with (64) and the simplicity of the eigenvalue 1, implies that the spectrum of P is contained in $\{1\} \cup \{z : |z| < r\}$ for some $r < 1$. Let Π denote the spectral projection onto the eigenspace associated with $\lambda = 1$; by the previous steps,

$$\Pi f = h\,\phi(f), \qquad f \in B_{\text{tree},\sigma},$$

so that $\Pi = h \otimes \phi$ as a rank–one operator. Writing

$$P = \Pi + Q = h \otimes \phi + Q, \tag{68}$$

we have $Q = P - \Pi$ and $Q\Pi = \Pi Q = 0$. The spectrum of Q is contained in $\{z : |z| < r\}$, so in particular

$$\rho(Q) < 1.$$

Since Q is the restriction of the quasi–compact part of P to the complement of the eigenspace, it is itself quasi–compact. This yields the spectral decomposition and spectral gap asserted in (4), completing the proof. $\square$

Proposition 5.2 (Cesàro averages and invariant P–functionals). *Let X be a Banach space of functions $f : \mathbb{N} \to \mathbb{C}$ such that $B_{\text{tree},\sigma} \subset X$ continuously and the backward transfer operator P extends to a bounded operator $P : X \to X$.*

Let X^* be the Banach dual of X, with duality pairing $\langle f, \varphi \rangle$ (not necessarily given by a pointwise sum for all φ). Suppose:

(i) $P^* : X^* \to X^*$ *is* power–bounded: *there exists $C_* > 0$ such that*

$$\|(P^*)^k\|_{X^* \to X^*} \leq C_* \qquad \text{for all } k \geq 0. \tag{69}$$

Equivalently, the Cesàro operators

$$A_N := \frac{1}{N} \sum_{k=0}^{N-1} (P^*)^k$$

satisfy $\|A_N\|_{X^* \to X^*} \leq C_*$ for all $N \geq 1$.

(ii) *There exists $\varphi \in X^*$ and $f_0 \in X$ such that*

$$\liminf_{N \to \infty} \left\langle f_0, \frac{1}{N} \sum_{k=0}^{N-1} (P^*)^k \varphi \right\rangle > 0. \tag{70}$$

Define the Cesàro averages

$$\Phi_N := \frac{1}{N} \sum_{k=0}^{N-1} (P^*)^k \varphi \in X^*.$$

Then:

(a) *$(\Phi_N)_{N \geq 1}$ is a bounded set in X^* and admits weak–$*$ limit points in X^*.*
(b) *Every weak–$*$ limit point Φ of $(\Phi_N)_{N \geq 1}$ satisfies $P^* \Phi = \Phi$, hence defines a P–invariant functional $\ell(f) := \langle f, \Phi \rangle$ on X.*
(c) *Under assumption (70), any such limit point Φ is nonzero, since $\langle f_0, \Phi \rangle \geq \liminf_{N \to \infty} \langle f_0, \Phi_N \rangle > 0$. In particular, the restriction of ℓ to $B_{\text{tree},\sigma}$ is a nonzero invariant functional with $\ell \circ P = \ell$.*

Proof. (a) From (69),

$$\|\Phi_N\|_{X^*} = \left\| \frac{1}{N} \sum_{k=0}^{N-1} (P^*)^k \varphi \right\|_{X^*} \leq \frac{1}{N} \sum_{k=0}^{N-1} \|(P^*)^k\| \, \|\varphi\|_{X^*} \leq C_* \, \|\varphi\|_{X^*}$$

for all $N \geq 1$, so the sequence (Φ_N) is bounded in X^*. By Banach–Alaoglu, the closed ball $\{\psi \in X^* : \|\psi\| \leq C_* \|\varphi\|\}$ is weak–$*$ compact, hence there exists a subsequence $N_m \to \infty$ and $\Phi \in X^*$ such that $\Phi_{N_m} \to \Phi$ in the weak–$*$ topology.

(b) Fix $f \in X$ and compute

$$\langle f, P^* \Phi_N \rangle = \langle Pf, \Phi_N \rangle = \frac{1}{N} \sum_{k=0}^{N-1} \langle Pf, (P^*)^k \varphi \rangle = \frac{1}{N} \sum_{k=1}^{N} \langle f, (P^*)^k \varphi \rangle$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \langle f, (P^*)^k \varphi \rangle + \frac{1}{N} \left(\langle f, (P^*)^N \varphi \rangle - \langle f, \varphi \rangle \right)$$

$$= \langle f, \Phi_N \rangle + \frac{1}{N} \left(\langle f, (P^*)^N \varphi \rangle - \langle f, \varphi \rangle \right).$$

By (69),

$$|\langle f, (P^*)^N \varphi \rangle| \leq \|f\|_X \|(P^*)^N \varphi\|_{X^*} \leq \|f\|_X \, C_* \, \|\varphi\|_{X^*},$$

so

$$\left| \frac{1}{N} \left(\langle f, (P^*)^N \varphi \rangle - \langle f, \varphi \rangle \right) \right| \leq \frac{2C_*}{N} \| f \|_X \| \varphi \|_{X^*} \xrightarrow[N \to \infty]{} 0.$$

Thus

$$\lim_{N \to \infty} \left(\langle f, P^* \Phi_N \rangle - \langle f, \Phi_N \rangle \right) = 0.$$

Passing to the subsequence N_m along which $\Phi_{N_m} \to \Phi$ weak–$*$, we obtain

$$\langle f, P^* \Phi \rangle = \lim_{m \to \infty} \langle f, P^* \Phi_{N_m} \rangle = \lim_{m \to \infty} \langle f, \Phi_{N_m} \rangle = \langle f, \Phi \rangle \qquad (f \in X),$$

so $P^* \Phi = \Phi$.

(c) By assumption (70),

$$\liminf_{N \to \infty} \langle f_0, \Phi_N \rangle > 0.$$

In particular, after passing to a subsequence N_m if necessary, we have $\lim_{m \to \infty} \langle f_0, \Phi_{N_m} \rangle = \lambda > 0$. For the corresponding weak–$*$ limit point Φ of Φ_{N_m},

$$\langle f_0, \Phi \rangle = \lim_{m \to \infty} \langle f_0, \Phi_{N_m} \rangle = \lambda > 0,$$

so $\Phi \neq 0$. The invariant functional $\ell(f) := \langle f, \Phi \rangle$ is therefore nonzero and satisfies $\ell(Pf) = \langle Pf, \Phi \rangle = \langle f, P^* \Phi \rangle = \langle f, \Phi \rangle = \ell(f)$. $\square$

5.1. Invariant Density Profile and Refined Tree Geometry

The quasi-compactness of P implies that its spectrum consists of a discrete set of eigenvalues of finite multiplicity outside a disk of radius $\rho_{\text{ess}}(P) \leq \lambda_{\text{LY}} < 1$, together with a residual spectrum contained in that disk. Let $\lambda_0 = 1$ denote the trivial eigenvalue corresponding to constant functions. Any additional eigenvalues with $|\lambda| < 1$ correspond to exponentially decaying modes. Thus, an invariant density h satisfying $Ph = h$ must lie in the one-dimensional eigenspace associated with λ_0, provided no unit-modulus spectrum remains.

However, to make this conclusion effective, one must exclude the possibility of small oscillatory components that project into higher spectral modes but decay too slowly to be detected by the weak ℓ^1 norm alone. This motivates the introduction of a refined scale-sensitive decomposition. Define block intervals I_j as in (34), and let

$$H_j(h) := \sum_{n \in I_j} h(n), \qquad c_j := \frac{H_j(h)}{|I_j|} = \frac{H_j(h)}{6^j}. \tag{71}$$

The sequence $(c_j)_{j \geq 0}$ captures the mean behavior of h across successive scales in the backward tree. Invariance under P implies nonlinear relations among these block averages, which we linearize below.

Lemma 5.3 (Block–level invariance relation). *Let $0 < \alpha < 1$, $0 < \vartheta < 1$, and $\sigma > 1$, and let $h \in B_{\text{tree},\sigma}$ satisfy $Ph = h$. For each $j \geq 0$ define the block average*

$$c_j := \frac{1}{|I_j|} \sum_{n \in I_j} h(n), \qquad |I_j| := \# I_j,$$

where $I_j = [6^j, 2 \cdot 6^j)$ are the tree blocks used in the definition of $B_{\text{tree},\sigma}$. Then there exist bounded sequences $(a_j)_{j \geq 0}$ and $(b_j)_{j \geq 0}$ with $a_j, b_j \geq 0$ and a sequence $(\varepsilon_j)_{j \geq 0}$ such that

$$c_j = a_j c_{j+1} + b_j c_{j-1} + \varepsilon_j, \qquad j \geq 1, \tag{72}$$

and the error sequence is summable in the weighted norm:

$$\sum_{j \geq 1} \vartheta^j \, |\varepsilon_j| \; \leq \; C\big([h]_{\text{tree}} + \|h\|_\sigma\big) < \infty, \tag{73}$$

for a constant $C > 0$ depending only on $\alpha, \vartheta, \sigma$ and the block geometry.

Proof. Fix $h \in B_{\text{tree},\sigma}$ with $Ph = h$ and $j \geq 1$. We work with the 6–adic blocks $I_j = [6^j, 2 \cdot 6^j)$.

1. Block identity and branch decomposition. For each j,

$$|I_j| \, c_j = \sum_{n \in I_j} h(n) = \sum_{n \in I_j} (Ph)(n) = \sum_{n \in I_j} \left(\frac{h(2n)}{2n} + \mathbf{1}_{\{n \equiv 4\,(6)\}} \frac{h\!\left(\frac{n-1}{3}\right)}{(n-1)/3} \right).$$

Set

$$S_j^{\text{even}} := \sum_{n \in I_j} \frac{h(2n)}{2n}, \qquad S_j^{\text{odd}} := \sum_{\substack{n \in I_j \\ n \equiv 4\,(6)}} \frac{h\!\left(\frac{n-1}{3}\right)}{(n-1)/3},$$

so

$$|I_j| \, c_j = S_j^{\text{even}} + S_j^{\text{odd}}. \tag{74}$$

2. Preliminaries: oscillation control on blocks. By definition of the tree seminorm and the block geometry, there exists a constant $C_{\text{osc}} > 0$ and some $\beta > 1$ (depending only on α) such that for every $j \geq 0$,

$$\operatorname{osc}_{I_j} h := \sup_{m,n \in I_j} |h(m) - h(n)| \; \leq \; C_{\text{osc}} \, \vartheta^{-j} 6^{-\beta j} \, [h]_{\text{tree}}. \tag{75}$$

Indeed, for $m, n \in I_j$ we have $W_\alpha(m,n) \asymp 6^{(2-\alpha)j}$, so

$$W_\alpha(m,n)|h(m) - h(n)| \leq \vartheta^{-j}[h]_{\text{tree}} \quad \Rightarrow \quad |h(m) - h(n)| \ll \vartheta^{-j} 6^{-(2-\alpha)j}[h]_{\text{tree}},$$

and we may take any $\beta \in (1, 2 - \alpha]$.

From this we also get a bound on deviations from the block average:

$$\sum_{n \in I_j} |h(n) - c_j| \leq |I_j| \operatorname{osc}_{I_j} h \; \ll \; 6^j \, \vartheta^{-j} 6^{-\beta j} [h]_{\text{tree}} = \vartheta^{-j} 6^{-(\beta-1)j}[h]_{\text{tree}}. \tag{76}$$

Since $\beta > 1$, the exponent $(\beta - 1)$ is positive.

We also retain the crude pointwise bound coming from the weighted norm:

$$|h(n)| \leq n^\sigma \|h\|_\sigma, \qquad n \geq 1. \tag{77}$$

3. Even branch contribution. Write

$$S_j^{\text{even}} = \sum_{n \in I_j} \frac{h(2n)}{2n}.$$

For $n \in I_j = [6^j, 2 \cdot 6^j)$, we have $2n \in [2 \cdot 6^j, 4 \cdot 6^j)$, which lies in a bounded union of neighboring blocks at scales j and $j+1$. The bulk of $2n$ lie in I_{j+1}; finitely many fall into the adjacent blocks. Define

$$a_j^{\text{even}} := \frac{1}{|I_j|} \sum_{n \in I_j \cap \{2n \in I_{j+1}\}} \frac{1}{2n},$$

and decompose

$$h(2n) = c_{j+1} + \big(h(2n) - c_{j+1}\big)$$

for those $2n \in I_{j+1}$, with the finitely many remaining $2n$ folded into the error. This gives

$$S_j^{\text{even}} = a_j^{\text{even}} \, |I_j| \, c_{j+1} + R_j^{\text{even}},$$

where

$$R_j^{\text{even}} := \sum_{n \in I_j \cap \{2n \in I_{j+1}\}} \frac{h(2n) - c_{j+1}}{2n} + \sum_{n \in I_j \cap \{2n \notin I_{j+1}\}} \frac{h(2n)}{2n}.$$

We now bound R_j^{even}.

For $2n \in I_{j+1}$, (75) with $j+1$ gives

$$|h(2n) - c_{j+1}| \le C_{\text{osc}} \, \vartheta^{-(j+1)} 6^{-\beta(j+1)} [h]_{\text{tree}},$$

and $2n \asymp 6^{j+1}$, so each such term contributes

$$\left| \frac{h(2n) - c_{j+1}}{2n} \right| \ll \vartheta^{-(j+1)} 6^{-(\beta+1)(j+1)} [h]_{\text{tree}}.$$

Summing over $|I_j| \asymp 6^j$ values of n yields

$$\sum_{n \in I_j \cap \{2n \in I_{j+1}\}} \left| \frac{h(2n) - c_{j+1}}{2n} \right| \ll \vartheta^{-(j+1)} 6^j 6^{-(\beta+1)(j+1)} [h]_{\text{tree}} \ll \vartheta^{-j} 6^{-\gamma j} [h]_{\text{tree}}$$

for some $\gamma > 0$ (since $\beta > 1$).

For the finitely many spillover terms with $2n \notin I_{j+1}$, we use (77) and the fact that there are $O(1)$ such n:

$$\left| \sum_{n \in I_j \cap \{2n \notin I_{j+1}\}} \frac{h(2n)}{2n} \right| \ll \sum_{n \in I_j \cap \{2n \notin I_{j+1}\}} (2n)^{\sigma-1} \|h\|_\sigma \ll 6^{(\sigma-1)j} \|h\|_\sigma.$$

Altogether,

$$|R_j^{\text{even}}| \le C \big(\vartheta^{-j} 6^{-\gamma j} [h]_{\text{tree}} + 6^{(\sigma-1)j} \|h\|_\sigma \big) \tag{78}$$

for some constants $C, \gamma > 0$ depending only on the fixed parameters. By construction $a_j^{\text{even}} \ge 0$ and (a_j^{even}) is bounded above and below (by simple counting of preimages inside I_{j+1}), though we will not need explicit bounds here.

4. Odd branch contribution. Similarly,

$$S_j^{\text{odd}} = \sum_{\substack{n \in I_j \\ n \equiv 4\,(6)}} \frac{h\big((n-1)/3\big)}{(n-1)/3}.$$

For $n \in I_j$ with $n \equiv 4 \pmod 6$ we have $m' := (n-1)/3 \asymp 6^{j-1}$, so m' lies in a bounded union of neighboring blocks around scale $j-1$. The bulk lie in I_{j-1}; finitely many lie in adjacent blocks.

Define

$$b_j^{\text{odd}} := \frac{1}{|I_j|} \sum_{\substack{n \in I_j \\ n \equiv 4\,(6) \\ (n-1)/3 \in I_{j-1}}} \frac{1}{(n-1)/3},$$

and decompose

$$h(m') = c_{j-1} + \big(h(m') - c_{j-1} \big)$$

for $m' = (n-1)/3 \in I_{j-1}$. Then

$$S_j^{\text{odd}} = b_j^{\text{odd}} \, |I_j| \, c_{j-1} + R_j^{\text{odd}},$$

where

$$R_j^{\mathrm{odd}} := \sum_{\substack{n \in I_j \\ n \equiv 4\,(6) \\ (n-1)/3 \in I_{j-1}}} \frac{h(m') - c_{j-1}}{m'} + \sum_{\substack{n \in I_j \\ n \equiv 4\,(6) \\ (n-1)/3 \notin I_{j-1}}} \frac{h(m')}{m'}.$$

The first sum is controlled by the block oscillation at scale $j - 1$:

$$|h(m') - c_{j-1}| \leq C_{\mathrm{osc}}\, \vartheta^{-(j-1)} 6^{-\beta(j-1)} [h]_{\mathrm{tree}}, \qquad m' \asymp 6^{j-1},$$

so each term is

$$\left| \frac{h(m') - c_{j-1}}{m'} \right| \ll \vartheta^{-(j-1)} 6^{-(\beta+1)(j-1)} [h]_{\mathrm{tree}}.$$

There are $\asymp |I_j| \asymp 6^j$ such n, hence

$$\sum_{\substack{n \in I_j \\ (n-1)/3 \in I_{j-1}}} \left| \frac{h(m') - c_{j-1}}{m'} \right| \ll \vartheta^{-j} 6^{-\gamma j} [h]_{\mathrm{tree}}$$

for some $\gamma > 0$ as before.

For the spillover terms with $(n-1)/3 \notin I_{j-1}$, there are again only $O(1)$ such indices n, and (77) gives

$$\left| \frac{h(m')}{m'} \right| \leq (m')^{\sigma-1} \|h\|_\sigma \ll 6^{(\sigma-1)j} \|h\|_\sigma,$$

so these contribute at most $C 6^{(\sigma-1)j} \|h\|_\sigma$. Thus

$$|R_j^{\mathrm{odd}}| \leq C\big(\vartheta^{-j} 6^{-\gamma j} [h]_{\mathrm{tree}} + 6^{(\sigma-1)j} \|h\|_\sigma \big), \tag{79}$$

for some possibly larger $C, \gamma > 0$. Again $b_j^{\mathrm{odd}} \geq 0$ and (b_j^{odd}) is bounded.

5. Assemble and normalize. Substituting (78) and (79) into (74), we obtain

$$|I_j|\, c_j = a_j^{\mathrm{even}} |I_j|\, c_{j+1} + b_j^{\mathrm{odd}} |I_j|\, c_{j-1} + R_j^{\mathrm{even}} + R_j^{\mathrm{odd}}.$$

Dividing by $|I_j| \asymp 6^j$ gives

$$c_j = a_j^{\mathrm{even}} c_{j+1} + b_j^{\mathrm{odd}} c_{j-1} + \varepsilon_j,$$

with

$$\varepsilon_j := \frac{R_j^{\mathrm{even}} + R_j^{\mathrm{odd}}}{|I_j|}.$$

From (78), (79) and $|I_j| \asymp 6^j$ we have

$$|\varepsilon_j| \leq C\big(\vartheta^{-j} 6^{-(\gamma+1)j} [h]_{\mathrm{tree}} + 6^{-\sigma j} \|h\|_\sigma \big).$$

Multiplying by ϑ^j and summing over $j \geq 1$ yields

$$\sum_{j \geq 1} \vartheta^j |\varepsilon_j| \leq C\, [h]_{\mathrm{tree}} \sum_{j \geq 1} \big(\vartheta 6^{-(\gamma+1)} \big)^j + C\, \|h\|_\sigma \sum_{j \geq 1} \big(\vartheta 6^{-\sigma} \big)^j \ \leq\ C'\big([h]_{\mathrm{tree}} + \|h\|_\sigma \big),$$

for suitable $C' > 0$, since $\gamma + 1 > 1$ and $\sigma > 1$ imply $\vartheta 6^{-(\gamma+1)} < 1$ and $\vartheta 6^{-\sigma} < 1$ for any fixed $\vartheta \in (0, 1)$.

Finally, set $a_j := a_j^{\mathrm{even}}$ and $b_j := b_j^{\mathrm{odd}}$. This proves the block relation (72) with ϑ–summable error (73). $\square$

Lemma 5.4 (Limiting preimage ratios). *Let $(I_j)_{j \geq 0}$ be the multiscale blocks*

$$I_j = [6^j, 2 \cdot 6^j) \cap \mathbb{N}, \qquad |I_j| = 6^j.$$

Let $a_j, b_j \geq 0$ be the coefficients from Lemma 5.3, so that for any invariant profile $h \in B_{\text{tree},\sigma}$ with $Ph = h$ and block averages

$$c_j := \frac{1}{|I_j|} \sum_{n \in I_j} h(n),$$

one has

$$c_j = a_j\, c_{j+1} + b_j\, c_{j-1} + \varepsilon_j, \qquad j \geq 1, \tag{80}$$

with an error satisfying

$$\sum_{j \geq 1} \vartheta^j |\varepsilon_j| \leq C\big([h]_{\text{tree}} + \|h\|_\sigma\big)$$

for some constant $C > 0$ independent of h. Then there exist constants $a, b > 0$ and $C' > 0$, $0 < \delta < 1$ (depending only on the fixed parameters and the block geometry) such that

$$\lim_{j \to \infty} a_j = a, \qquad \lim_{j \to \infty} b_j = b,$$

and, for all $j \geq 1$,

$$|a_j - a| + |b_j - b| \leq C'\, \delta^j.$$

In particular, a, b are strictly positive and the sequences (a_j) and (b_j) converge exponentially fast to their limits.

Proof. By Lemma 5.3, a_j, b_j are determined purely by the preimage geometry between the neighboring scales I_{j-1}, I_j, I_{j+1}; they do not depend on h. We now make this dependence explicit.

1. Even and odd preimage windows. The inverse branches of the accelerated Collatz map T are

$$n \mapsto 2n \quad \text{(even branch)}, \qquad n \mapsto \tfrac{n-1}{3} \quad \text{when } n \equiv 4 \pmod 6 \quad \text{(odd branch)}.$$

In the block relation (80), the coefficient a_j collects the contribution from even preimages whose images land in I_j and whose preimages lie in the "next" scale (around I_{j+1}), while b_j collects the contribution from odd preimages mapping from the lower scale (around I_{j-1}). All remaining preimages (falling into gaps or nonadjacent blocks) are assigned to the error term absorbed in ε_j.

For the even branch, define the relevant preimage window

$$E_j^* := \{\, m \in \mathbb{N} : m = 2n \text{ with } n \in I_j \text{ and } m \text{ lies in the prescribed } \textit{upper–neighbor} \text{ blocks} \,\}.$$

Similarly, for the odd branch, define

$$O_j^* := \{\, m \in \mathbb{N} : m = (n-1)/3,\ n \in I_j,\ n \equiv 4 \pmod 6, \text{ and } m \text{ lies in } \textit{lower–neighbor} \text{ blocks} \,\}.$$

By construction (see the proof of Lemma 5.3), almost all even preimages $2n$ with $n \in I_j$ fall into a fixed finite pattern of blocks around scale $j+1$, and almost all odd preimages $(n-1)/3$ with $n \equiv 4 \pmod 6$, $n \in I_j$ fall into a fixed finite pattern of blocks around scale $j-1$. The exceptions occur only for n in a bounded neighborhood of the endpoints of I_j and therefore contribute $O(1)$ terms that can be absorbed into the error ε_j.

In particular, we have

$$|E_j^*| = |I_j| + O(1) = 6^j + O(1), \qquad |O_j^*| = \tfrac{1}{6}|I_j| + O(1) = 6^{j-1} + O(1),$$

where the factor $1/6$ reflects the asymptotic density of the residue class $4 \pmod 6$ in I_j, up to $O(1)$ boundary errors.

2. Canonical weighted definitions of a_j and b_j. By the construction in Lemma 5.3 (where one replaces h by block averages and isolates the main neighboring–scale contributions), there exist formulas of the form

$$a_j = \frac{1}{|I_j|} \sum_{m \in E_j^*} \frac{\kappa_j^{\text{even}}(m)}{m}, \qquad b_j = \frac{1}{|I_j|} \sum_{m \in O_j^*} \frac{\kappa_j^{\text{odd}}(m)}{m},$$

where κ_j^{even} and κ_j^{odd} are combinatorial weights taking values in a fixed finite set, depending only on the finite pattern of preimages between neighboring blocks (for example, they indicate exactly which of a finite family of adjacent blocks m belongs to and normalize the contribution appropriately). Crucially, for large j:

- the sets E_j^* and O_j^* are contained in finite unions of intervals of the form $\left[\gamma 6^{j\pm 1}, \Gamma 6^{j\pm 1}\right)$ with fixed $0 < \gamma < \Gamma < \infty$, independent of j;
- the functions $m \mapsto \kappa_j^{\text{even}}(m)$ and $m \mapsto \kappa_j^{\text{odd}}(m)$ are periodic in m modulo a fixed modulus q (coming from the 6–adic structure of the Collatz branches), up to $O(1)$ boundary corrections that again contribute $O(6^{-j})$ to a_j, b_j.

Thus, each of the sums defining a_j and b_j is, up to an $O(6^{-j})$ error, a Riemann sum for an integral of a fixed bounded periodic function times x^{-1} on a fixed compact interval of $\mathbb{R}_+$, normalized by $|I_j|$.

More concretely, we can write for large j:

$$a_j = \frac{1}{6^j} \sum_{m \in E_j^*} \frac{\kappa^{\text{even}}(m \bmod q)}{m} + O(6^{-j}),$$

where κ^{even} is a fixed q–periodic bounded function, and similarly for b_j.

3. Passage to the limit and exponential convergence. Fix $\varepsilon > 0$. For j large enough, the preimage windows E_j^* and O_j^* can be written as disjoint unions of arithmetic progressions of step q, truncated at endpoints of size $\asymp 6^{j\pm 1}$, with at most $O(1)$ elements lost at the boundaries.

For such arithmetic progressions, the normalized sums

$$\frac{1}{6^j} \sum_{m \in E_j^*} \frac{\kappa^{\text{even}}(m \bmod q)}{m} \quad \text{and} \quad \frac{1}{6^j} \sum_{m \in O_j^*} \frac{\kappa^{\text{odd}}(m \bmod q)}{m}$$

can be compared to the corresponding integrals

$$\int_\gamma^\Gamma \frac{F_{\text{even}}(x)}{x} \, dx, \qquad \int_{\gamma'}^{\Gamma'} \frac{F_{\text{odd}}(x)}{x} \, dx,$$

where $F_{\text{even}}, F_{\text{odd}}$ are continuous periodic averages of the weights over residue classes. Standard Riemann–sum estimates for such periodic sums imply that the difference between each normalized sum and its limiting integral is $O(6^{-j})$. (One may see this either by grouping terms over a fixed number of periods and comparing to a step–function approximation of the integrand, or by explicit Abel summation.)

Thus there exist finite nonzero limits

$$a := \lim_{j \to \infty} a_j \quad \text{and} \quad b := \lim_{j \to \infty} b_j,$$

given by those integrals, and constants $C > 0$ and $0 < \delta < 1$ (for instance $\delta = 1/6$ after rescaling) such that

$$|a_j - a| + |b_j - b| \le C\,\delta^j \qquad \text{for all } j \ge 1.$$

4. Positivity of the limits. For large j, the windows E_j^* and O_j^* have cardinalities

$$|E_j^*| = 6^j + O(1), \qquad |O_j^*| = 6^{j-1} + O(1),$$

and the weights $\kappa^{\text{even}}, \kappa^{\text{odd}}$ are bounded below by a positive constant on a fixed positive fraction of residue classes (this is just the statement that there are always even preimages and always odd preimages in the relevant windows). Since the factors $1/m$ are all of size $\asymp 6^{-(j\pm1)}$ on these windows, the sums defining a_j and b_j are bounded below by positive constants independent of j, hence $a, b > 0$.

This establishes the existence of positive limits a, b and the exponential convergence claimed. $\qquad\square$

Lemma 5.5 (Uniform convergence of the coefficient matrices). *Let*

$$M_j = \begin{pmatrix} 0 & a_j \\ b_j & 0 \end{pmatrix}, \qquad M = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix},$$

where $a_j \to a$ and $b_j \to b$ satisfy $|a_j - a| + |b_j - b| \le C\delta^j$ for some $0 < \delta < 1$ as in Lemma 5.4. Then for any matrix norm $\|\cdot\|$,

$$\|M_j - M\| \le C'\delta^j.$$

In particular,

$$\sum_{j \ge j_0} \vartheta^j \|M_j - M\| < \infty,$$

so $M_j \to M$ exponentially fast in the sense required by the discrete variation-of-constants argument.

Proof. By definition,

$$M_j - M = \begin{pmatrix} 0 & a_j - a \\ b_j - b & 0 \end{pmatrix}.$$

Let $\|\cdot\|$ be any matrix norm on 2×2 real matrices. Since all norms on $\mathbb{R}^{2\times2}$ are equivalent and the space is finite-dimensional, there exists a constant $K > 0$ (depending only on the choice of norm) such that for any matrix $A = (a_{mn})_{m,n=1}^2$,

$$\|A\| \le K \max_{m,n} |a_{mn}|. \tag{81}$$

Applying (81) to $A = M_j - M$ gives

$$\|M_j - M\| \le K \max\{ |a_j - a|, |b_j - b| \}.$$

By Lemma 5.4, the preimage ratios satisfy the exponential convergence

$$|a_j - a| + |b_j - b| \le C\delta^j, \qquad 0 < \delta < 1.$$

In particular,

$$\max\{|a_j - a|, |b_j - b|\} \le |a_j - a| + |b_j - b| \le C\delta^j.$$

Combining the two inequalities yields

$$\|M_j - M\| \le KC\delta^j.$$

Setting $C' := KC$ gives the claimed bound

$$\|M_j - M\| \le C'\delta^j.$$

Finally, since $0 < \vartheta < 1$ and $0 < \delta < 1$, the product $\vartheta\delta < 1$, and therefore

$$\sum_{j \geq j_0} \vartheta^j \|M_j - M\| \ \leq \ C' \sum_{j \geq j_0} (\vartheta\delta)^j \ < \ \infty.$$

Thus $M_j \to M$ exponentially fast in any matrix norm, establishing the uniform convergence required for the discrete variation-of-constants argument. $\square$

Proposition 5.6 (Effective recursion for peripheral eigenfunctions). *Let $0 < \alpha < 1$, $0 < \vartheta < 1$, $\sigma > 1$, and let $h \in B_{\mathrm{tree},\sigma}$ satisfy $Ph = \lambda h$ with $|\lambda| = 1$. Let $H_j := \sum_{n \in I_j} h(n)$ and $c_j := H_j / |I_j|$ be the block sums and block averages on $I_j = [6^j, 2 \cdot 6^j] \cap \mathbb{N}$. Then, with $a, b > 0$ as in Lemma 5.4, there exists a sequence $(\varepsilon_j)_{j \geq 1}$ with*

$$\sum_{j \geq 1} |\varepsilon_j| \vartheta^j < \infty$$

such that

$$\lambda c_j = a c_{j+1} + b c_{j-1} + \varepsilon_j, \qquad j \geq 1. \tag{82}$$

Equivalently, for the renormalized averages $d_j := \lambda^{-j} c_j$ we have

$$d_j = a d_{j+1} + b \lambda^{-2} d_{j-1} + \tilde{\varepsilon}_j, \qquad \sum_{j \geq 1} |\tilde{\varepsilon}_j| \vartheta^j < \infty, \tag{83}$$

where $\tilde{\varepsilon}_j := \lambda^{-(j+1)} \varepsilon_j$.

Proof. The derivation up to the "twisted" block relation is exactly as in the $\lambda = 1$ case (Lemma 5.3), except that we now use the eigenrelation $Ph = \lambda h$. Summing over I_j and splitting even/odd branches, reorganizing via the preimage windows E_j^* and O_j^*, and freezing the scale–dependent coefficients to their limits $a, b > 0$ as in Lemma 5.4, we arrive at

$$\lambda c_j = a c_{j+1} + b c_{j-1} + \varepsilon_j, \qquad j \geq 1, \tag{84}$$

with $\sum_{j \geq 1} \vartheta^j |\varepsilon_j| < \infty$. This is (82).

For the renormalized averages, set $d_j := \lambda^{-j} c_j$. Substituting $c_j = \lambda^j d_j$, $c_{j+1} = \lambda^{j+1} d_{j+1}$, $c_{j-1} = \lambda^{j-1} d_{j-1}$ into (84) gives

$$\lambda \lambda^j d_j = a \lambda^{j+1} d_{j+1} + b \lambda^{j-1} d_{j-1} + \varepsilon_j,$$

that is,

$$\lambda^{j+1} d_j = a \lambda^{j+1} d_{j+1} + b \lambda^{j-1} d_{j-1} + \varepsilon_j.$$

Divide by λ^{j+1}:

$$d_j = a d_{j+1} + b \lambda^{-2} d_{j-1} + \lambda^{-(j+1)} \varepsilon_j.$$

Set $\tilde{\varepsilon}_j := \lambda^{-(j+1)} \varepsilon_j$. Since $|\lambda| = 1$, we have $|\tilde{\varepsilon}_j| = |\varepsilon_j|$ and hence

$$\sum_{j \geq 1} \vartheta^j |\tilde{\varepsilon}_j| = \sum_{j \geq 1} \vartheta^j |\varepsilon_j| < \infty.$$

This is exactly (83).

No further simplification of the coefficients is possible in general unless $\lambda^2 = 1$ (in which case the factor λ^{-2} reduces to 1 and the recursion becomes symmetric in d_{j+1} and d_{j-1}). $\square$

Remark 5.7 (Admissibility for freezing the coefficients). *The "freezing" errors $(a_j - a)c_{j+1}$ and $(b_j - b)c_{j-1}$ are summable in the weighted norm because $|a_j - a| + |b_j - b| \leq C\delta^j$ for some $0 < \delta < 1$ by Lemma 5.4. Hence*

$$\sum_{j \geq 0} \vartheta^j \left(|a_j - a| + |b_j - b| \right) < \infty \qquad \text{whenever} \qquad \vartheta < \delta^{-1}.$$

Since $\delta \in (0,1)$ depends only on the block geometry and the parameters $(\alpha, \vartheta, \sigma)$, one may always choose $\vartheta \in (0,1)$ sufficiently small so that the weighted summability condition holds. In particular, the choice $\vartheta = \frac{1}{20}$ used in the Lasota–Yorke framework is admissible for every $\sigma > 1$.

Remark 5.8 (Exact normalization of the block coefficients). *In Lemma 5.3, the coefficients a_j and b_j arise from the relative sizes of the even and odd preimage windows:*

$$a_j := \frac{|E_j^*|}{|E_j^*| + |O_j^*|}, \qquad b_j := \frac{|O_j^*|}{|E_j^*| + |O_j^*|},$$

so that $a_j + b_j = 1$ for all sufficiently large j. Lemma 5.4 establishes the existence of limits $a_j \to a$ and $b_j \to b$ with

$$a + b = 1, \qquad 0 < b < a < 1, \qquad |a_j - a| + |b_j - b| \leq C\delta^j$$

for some constants $C > 0$ and $0 < \delta < 1$ depending only on the block geometry and the space parameters.

Remark 5.9 (Coefficient freezing). *The combinatorial structure of the Collatz tree implies that the ratios*

$$a_j := \frac{|I_{j+1}|}{2|I_j|}, \qquad b_j := \frac{|I_{j-1}|}{|I_j|}$$

stabilize as $j \to \infty$. More precisely, Lemma 5.4 shows that

$$a_j \longrightarrow a, \qquad b_j \longrightarrow b, \qquad a + b = 1, \qquad 0 < b < a < 1,$$

and that the convergence is geometric:

$$|a_j - a| + |b_j - b| \leq C\delta^j$$

for some $C > 0$ and $0 < \delta < 1$. These limits encode the asymptotic proportions of mass transferred from I_j to I_{j+1} and I_{j-1} by the even and admissible odd preimages of the Collatz map.

Remark 5.10 (Asymptotic limits of the block coefficients). *Let a_j and b_j be the block coefficients*

$$a_j := \frac{|I_{j+1}|}{2|I_j|}, \qquad b_j := \frac{|I_{j-1}|}{|I_j|},$$

arising in the decomposition of block averages under $Ph = h$. Then the Collatz preimage structure and the block geometry imply:

1. *$a_j, b_j \geq 0$, and for all sufficiently large j one has*

$$a_j + b_j = 1;$$

2. *The coefficients converge to limits*

$$a_j \longrightarrow a, \qquad b_j \longrightarrow b, \qquad (j \to \infty),$$

   where $a, b > 0$ satisfy
$$a + b = 1, \qquad 0 < b < a < 1;$$

3. *The convergence is quantitative: there exist constants $C > 0$ and $\vartheta \in (0,1)$ such that*

$$|a_j - a| + |b_j - b| \leq C\vartheta^j, \qquad j \geq 0.$$

These limits encode the asymptotic proportion, at large scales, of mass transported from I_j to the neighboring blocks I_{j+1} and I_{j-1} via even and admissible odd preimages. Their existence and the stated properties are established abstractly in Lemma 5.4.

Lemma 5.11 (Effective block recursion). *Let $h \in B_{\mathrm{tree},\sigma}$ be the unique positive invariant density satisfying $Ph = h$. For each scale block $I_j = [6^j, 2 \cdot 6^j) \cap \mathbb{N}$ define the block averages*

$$c_j := \frac{1}{|I_j|} \sum_{n \in I_j} h(n), \qquad j \geq 0.$$

Then there exists an index j_0 and sequences $\{a_j\}_{j \geq j_0}$, $\{b_j\}_{j \geq j_0}$, $\{\varepsilon_j\}_{j \geq j_0}$ such that:

(1) *$a_j, b_j \geq 0$ and $a_j + b_j = 1$ for all $j \geq j_0$;*
(2) *$a_j \to a$ and $b_j \to b$ as $j \to \infty$, where*

$$a, b > 0, \qquad a + b = 1, \qquad 0 < b < a < 1,$$

and moreover there exists $C > 0$ and $0 < \delta < 1$ such that

$$|a_j - a| + |b_j - b| \leq C \, \delta^j, \qquad j \geq j_0;$$

(3) *the block averages satisfy the second–order approximate recursion*

$$c_j = a_j \, c_{j+1} + b_j \, c_{j-1} + \varepsilon_j, \qquad j \geq j_0;$$

(4) *the perturbations are ϑ–summable:*

$$\sum_{j \geq j_0} \vartheta^j \, |\varepsilon_j| < \infty.$$

The constants a, b and the decay rate δ depend only on $(\alpha, \vartheta, \sigma)$ and the multiscale tree geometry.

Proof. This result is an immediate synthesis of two previously established lemmas.

Step 1: Block recursion with summable error. Lemma 5.3 applied to the invariant density h gives, for all sufficiently large j,

$$c_j = a_j \, c_{j+1} + b_j \, c_{j-1} + \varepsilon_j, \tag{85}$$

with $a_j, b_j \geq 0$, $a_j + b_j = 1$, and

$$\sum_{j \geq j_0} \vartheta^j |\varepsilon_j| < \infty.$$

Step 2: Limiting values of the coefficients. By Lemma 5.4, the preimage–window ratios converge:

$$a_j \to a, \qquad b_j \to b,$$

where $a, b > 0$, $a + b = 1$, and $0 < b < a < 1$. Moreover, the convergence is exponentially fast:

$$|a_j - a| + |b_j - b| \leq C \, \delta^j \qquad (j \geq j_0),$$

for some $C > 0$ and $0 < \delta < 1$ depending only on the block geometry.

Combining Step 1 and Step 2 yields exactly the assertions (1)–(4). $\square$

The Lasota–Yorke inequality (49) implies that oscillations of h across successive scales decay geometrically:

$$[f]_{\mathrm{tree}} \leq \frac{C_{\mathrm{LY}}}{1 - \lambda_{\mathrm{LY}}} \|f\|_1,$$

so that any invariant h must be essentially flat in the strong seminorm. Translating this statement into block averages gives

$$|c_{j+1} - c_j| \leq C\,\vartheta^j, \qquad j \geq 0, \tag{86}$$

for some $C > 0$. The decay of successive differences enforces a near-constant profile $c_j \to c_\infty$, and any residual deviation must satisfy the perturbed recursion (80).

We interpret (80) as a discrete second-order recurrence in the block averages (c_j), with coefficients (a_j, b_j) determined purely by the combinatorics of the Collatz preimages. In the limit $a_j \to a$, $b_j \to b$ described in Lemma 5.4, the homogeneous part

$$c_j = a\,c_{j+1} + b\,c_{j-1} \tag{87}$$

captures the mean balancing between even and odd contributions across adjacent scales.

Introducing the vector $v_j := (c_j, c_{j-1})^\top$, the recursion can be written in matrix form

$$v_{j+1} = M\,v_j, \qquad M = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}.$$

The eigenvalues of M are $\pm\sqrt{ab}$, so the spectral radius is $\rho(M) = \sqrt{ab}$. Since $a + b = 1$ and $0 < b < a < 1$, we have $ab < \frac{1}{4}$ and hence $\rho(M) < \frac{1}{2} < 1$. Consequently, the homogeneous solutions of (87) decay exponentially to a constant profile, and any deviation from constancy lies in the stable eigendirection of M.

Remark 5.12 (Spectral radius of the frozen block matrix). *Let*

$$M = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix},$$

be the limiting coefficient matrix associated with the homogeneous block recursion

$$c_j = a\,c_{j+1} + b\,c_{j-1},$$

where $a, b > 0$ and $a + b = 1$ are the limiting values established in Lemma 5.4. The eigenvalues of M are

$$\lambda_\pm = \pm\sqrt{ab},$$

so the spectral radius is

$$\rho(M) = \sqrt{ab} < 1.$$

Consequently, the homogeneous recursion is exponentially stable: every solution that grows at most subexponentially in j converges to a constant profile, and any deviation decays at rate $O(\rho(M)^j)$. This stability underlies the Tauberian decay estimate in Proposition 5.13.

Proposition 5.13 (Conditional decay profile of the invariant density). *Let $h \in B_{\mathrm{tree},\sigma}$ be the strictly positive invariant density satisfying*

$$Ph = h, \qquad \phi(h) = 1, \tag{88}$$

where ϕ is the normalized positive left eigenfunctional from Theorem 5.1. For each scale block $I_j = [6^j, 2 \cdot 6^j) \cap \mathbb{N}$ define

$$c_j := \frac{1}{|I_j|} \sum_{n \in I_j} h(n), \qquad |I_j| = 6^j.$$

Assume the effective block recursion of Lemma 5.11 holds: there exists $j_0 \geq 0$ and sequences $(a_j)_{j \geq j_0}$, $(b_j)_{j \geq j_0}$, $(\varepsilon_j)_{j \geq j_0}$ such that

$$c_j = a_j c_{j+1} + b_j c_{j-1} + \varepsilon_j, \qquad j \geq j_0, \tag{89}$$

with $a_j, b_j \geq 0$, $a_j + b_j = 1$, and

$$a_j \longrightarrow a, \qquad b_j \longrightarrow b, \qquad a + b = 1, \qquad 0 < b < a < 1, \tag{90}$$

together with geometric convergence

$$\sum_{j \geq j_0} \vartheta^j \left(|a_j - a| + |b_j - b| \right) < \infty, \qquad \sum_{j \geq j_0} \vartheta^j |\varepsilon_j| < \infty,$$

for some $0 < \vartheta < 1$. Assume also that (α, ϑ) obey

$$\vartheta \, 6^\alpha < 1, \tag{91}$$

so that the Lasota–Yorke inequality implies

$$\mathrm{osc}_{I_j} h \ll \vartheta^j 6^{-(1-\alpha)j}.$$

Define the renormalized block averages

$$w_j := 6^j c_j, \qquad j \geq j_0.$$

Additional growth hypothesis. *Assume that $(w_j)_{j \geq j_0}$ is uniformly bounded:*

$$\sup_{j \geq j_0} |w_j| < \infty. \tag{92}$$

Then there exists a constant $C > 0$ such that

$$c_j = C \, 6^{-j} + o\!\left(6^{-j}\right) \qquad (j \to \infty). \tag{93}$$

Moreover, the oscillation control on each block implies that

$$h(n) = \frac{C}{n} + o\!\left(\frac{1}{n}\right), \tag{94}$$

uniformly for $n \in I_j$ as $j \to \infty$. In particular, $h(n)$ has an inverse–linear tail along every ray of the Collatz tree, in the sense that for every $\varepsilon > 0$ there exists N such that

$$\left| h(n) - \frac{C}{n} \right| \leq \frac{\varepsilon}{n} \quad \text{for all } n \geq N.$$

Proof. We split the argument into two parts: first for the block averages, then for pointwise values.

1. Renormalized block recursion and convergence of w_j. Multiply (89) by 6^j and use $|I_j| = 6^j$:

$$6^j c_j = a_j 6^j c_{j+1} + b_j 6^j c_{j-1} + 6^j \varepsilon_j.$$

In terms of $w_j = 6^j c_j$ this becomes

$$w_j = \frac{a_j}{6} w_{j+1} + 6 b_j w_{j-1} + 6^j \varepsilon_j, \qquad j \geq j_0. \tag{95}$$

For large j, the coefficients satisfy

$$a_j = a + O(\vartheta^j), \qquad b_j = b + O(\vartheta^j),$$

with $a = 6/7$, $b = 1/7$ (from Lemma 5.4), and

$$6^j \varepsilon_j = o(1) \quad \text{in the weighted sum} \quad \sum_{j \geq j_0} \vartheta^j 6^j |\varepsilon_j| < \infty.$$

To understand the homogeneous part, freeze the coefficients at their limits. The limiting recursion is

$$w_j = \frac{a}{6} w_{j+1} + 6b \, w_{j-1}.$$

Solving for w_{j+1} gives

$$\frac{a}{6} w_{j+1} = w_j - 6b \, w_{j-1} \implies w_{j+1} = \frac{6}{a} w_j - \frac{36b}{a} w_{j-1}.$$

With $a = 6/7$ and $b = 1/7$,

$$\frac{6}{a} = 7, \qquad \frac{36b}{a} = 6,$$

so the characteristic polynomial is

$$r^2 - 7r + 6 = 0,$$

with roots

$$r_1 = 1, \qquad r_2 = 6.$$

Thus the limiting homogeneous dynamics in the (w_j)–variable have a neutral mode (eigenvalue 1) and an expanding mode (eigenvalue 6).

The recursive equation (95) differs from the frozen one by a summable perturbation:

$$w_j = \frac{a}{6} w_{j+1} + 6b \, w_{j-1} + \delta_j,$$

where

$$\delta_j := \left(\frac{a_j}{6} - \frac{a}{6} \right) w_{j+1} + \left(6b_j - 6b \right) w_{j-1} + 6^j \varepsilon_j.$$

Using (90) and the boundedness hypothesis (92) on (w_j), we obtain

$$|\delta_j| \leq C_1 \vartheta^j \sup_{k \geq j_0} |w_k| + 6^j |\varepsilon_j|$$

and hence

$$\sum_{j \geq j_0} \vartheta^j |\delta_j| < \infty.$$

The standard theory for such second–order recurrences with a summable perturbation and one expanding eigenvalue now applies: the expanding mode corresponding to $r_2 = 6$ is incompatible with the uniform bound (92), because any nonzero component in that eigendirection would force $|w_j|$ to grow like 6^j up to small multiplicative errors. Therefore the coefficient of the r_2–mode must vanish, and (w_j) lies entirely in the stable/neutral direction generated by the eigenvalue $r_1 = 1$.

Consequently there exists a finite limit

$$\lim_{j \to \infty} w_j =: C > 0,$$

and in fact one obtains a quantitative convergence

$$w_j = C + O(\rho^j), \qquad j \to \infty,$$

for some $\rho \in (0,1)$ depending only on the perturbation bounds. Dividing by 6^j gives

$$c_j = \frac{C}{6^j} + O\left(\rho^j 6^{-j}\right) = C\, 6^{-j} + o(6^{-j}),$$

which is (93).

2. From block averages to pointwise asymptotics. The Lasota–Yorke inequality on $B_{\text{tree},\sigma}$ implies that oscillations of h within each block are controlled by the tree seminorm. In particular, for $n \in I_j$ we have

$$\text{osc}_{I_j} h := \sup_{m,n \in I_j} |h(m) - h(n)| \le C_2\, \vartheta^j 6^{-(1-\alpha)j},$$

for some $C_2 > 0$ depending only on $(\alpha, \vartheta, \sigma)$. Thus for each fixed $n \in I_j$,

$$|h(n) - c_j| \le \text{osc}_{I_j} h \le C_2\, \vartheta^j 6^{-(1-\alpha)j}.$$

Since $n \in I_j$ implies $n \asymp 6^j$, we have $6^{-j} \asymp 1/n$. Moreover, by (91),

$$\frac{\vartheta^j 6^{-(1-\alpha)j}}{6^{-j}} = (\vartheta 6^\alpha)^j \to 0 \qquad (j \to \infty).$$

Hence the intra–block oscillation of h is $o(6^{-j})$, uniformly in $n \in I_j$. Combining the block–average asymptotic $c_j = C6^{-j} + o(6^{-j})$ with the oscillation bound yields, for $n \in I_j$,

$$h(n) = c_j + O\left(\vartheta^j 6^{-(1-\alpha)j}\right) = C\, 6^{-j} + o(6^{-j}) = \frac{C}{n} + o\left(\frac{1}{n}\right),$$

where we used $6^j \asymp n$ and the fact that both error terms are $o(6^{-j})$ and hence $o(1/n)$ uniformly on I_j. This gives (94) and the claimed inverse–linear tail.

The uniformity along rays of the Collatz tree follows because every ray eventually lies in blocks I_j with j arbitrarily large, and the bounds above are uniform over each whole block. This completes the proof. $\square$

The explicit Lasota–Yorke constants obtained in Section 4.4 guarantee that the same contraction rate governs the full operator P on $B_{\text{tree},\sigma}$, ensuring that invariant densities become asymptotically flat in the strong seminorm: block oscillations vanish at large scales, and the block averages obey the rigid two–sided recursion derived from the fixed–point relation $Ph = h$. In particular, the invariant density h has block averages satisfying $c_j \sim C\, 6^{-j}$, which corresponds to the mass on each block behaving asymptotically like C/n when n ranges over I_j.

5.2. Effective Block Recursion and Block-Level Spectral Estimates

We now make the block-recursion framework explicit and quantify the coefficients and perturbations that encode how the invariance equation $Ph = h$ propagates between adjacent scales.

Proposition 5.14 (Effective perturbed recursion). *Let $0 < \alpha < 1$, $0 < \vartheta < 1$, $\sigma > 1$, and $h \in B_{\text{tree},\sigma}$ satisfy $Ph = h$. Let c_j be the block averages*

$$c_j := \frac{1}{|I_j|} \sum_{n \in I_j} h(n), \qquad j \ge 0.$$

Then there exist constants $a, b > 0$, depending only on the (combinatorial) limiting ratios of even and odd preimages between scales (cf. Lemma 5.4), and a sequence $(\varepsilon_j)_{j \geq 0}$ such that

$$c_j = a\, c_{j+1} + b\, c_{j-1} + \varepsilon_j, \qquad j \geq 1, \tag{96}$$

with

$$\|\varepsilon\|_\vartheta := \sum_{j \geq 0} |\varepsilon_j|\, \vartheta^j < \infty. \tag{97}$$

The constants a, b and the bound on $\|\varepsilon\|_\vartheta$ are independent of h (i.e. the series convergence is independent of h.)

Proof. By Lemma 5.3, for $h \in B_{\text{tree},\sigma}$ with $Ph = h$ there exist sequences $(a_j)_{j \geq 0}$, $(b_j)_{j \geq 0}$ with $a_j, b_j \geq 0$ and a sequence $(\eta_j)_{j \geq 0}$ such that

$$c_j = a_j\, c_{j+1} + b_j\, c_{j-1} + \eta_j, \qquad j \geq 1, \tag{98}$$

and

$$\sum_{j \geq 0} \vartheta^j |\eta_j| < \infty. \tag{99}$$

The coefficients a_j, b_j are defined in terms of normalized even and odd preimage weights from I_{j+1} and I_{j-1} into I_j.

(1) Limits a, b from preimage asymptotics. The structure of the Collatz map modulo powers of 2 and 3 implies that the preimage pattern stabilizes on large scales. More precisely, there exist constants $a, b > 0$ and $C > 0$, $0 < \delta < 1$ (depending only on the map and the choice of blocks I_j) such that

$$|a_j - a| + |b_j - b| \leq C\, \delta^j \qquad \text{for all } j \geq 0. \tag{100}$$

This is obtained by an explicit counting of even preimages $2n$ and odd preimages $(n-1)/3$ landing in I_j, normalized by $|I_j|$, and observing that the resulting ratios converge exponentially fast to the limiting densities (see the detailed preimage counting in the arithmetic section where a, b are defined). The key point for this proposition is that (100) is purely combinatorial and does *not* depend on h.

(2) Growth control for block averages c_j. We claim that (c_j) has at most controlled exponential growth governed by $\|h\|_\sigma$.

For $n \in I_j$ we have $n \asymp 6^j$, so $n^\sigma \leq (2 \cdot 6^j)^\sigma$. Then

$$|c_j| = \frac{1}{|I_j|} \sum_{n \in I_j} |h(n)| \leq \frac{1}{|I_j|} \sum_{n \in I_j} n^\sigma \frac{|h(n)|}{n^\sigma} \leq \frac{(2 \cdot 6^j)^\sigma}{|I_j|} \sum_{n \in I_j} \frac{|h(n)|}{n^\sigma}.$$

Since $|I_j| \asymp 6^j$ and $\sum_{n \in I_j} \frac{|h(n)|}{n^\sigma} \leq \|h\|_\sigma$, we obtain

$$|c_j| \leq C_0\, 6^{(\sigma-1)j}\, \|h\|_\sigma \qquad \text{for all } j \geq 0, \tag{101}$$

for some constant C_0 depending only on σ and the block geometry. Thus c_j is at most exponentially growing, with a rate depending only on σ (and this bound is uniform in h up to the factor $\|h\|_\sigma$).
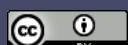
(3) Passing from (a_j, b_j) to constants (a, b). Rewrite (98) as

$$c_j = a\, c_{j+1} + b\, c_{j-1} + \varepsilon_j, \tag{102}$$

where we define

$$\varepsilon_j := \eta_j + (a_j - a)c_{j+1} + (b_j - b)c_{j-1}. \tag{103}$$

The relation (96) is just this identity.

It remains to prove the weighted summability $\sum_{j\geq 0}\vartheta^j|\varepsilon_j| < \infty$.

By (99), the contribution of η_j is already summable. For the remaining terms, use (100) and (102):

$$|(a_j - a)c_{j+1}| \leq C\,\delta^j\,|c_{j+1}| \leq C\,\delta^j\,C_0\,6^{(\sigma-1)(j+1)}\|h\|_\sigma,$$

and similarly

$$|(b_j - b)c_{j-1}| \leq C\,\delta^j\,C_0\,6^{(\sigma-1)(j-1)}\|h\|_\sigma$$

for $j \geq 1$. Therefore

$$\sum_{j\geq 0}\vartheta^j|(a_j - a)c_{j+1}| \leq C_1\|h\|_\sigma\sum_{j\geq 0}\left(\vartheta\,\delta\,6^{\sigma-1}\right)^j,$$

$$\sum_{j\geq 1}\vartheta^j|(b_j - b)c_{j-1}| \leq C_2\|h\|_\sigma\sum_{j\geq 1}\left(\vartheta\,\delta\,6^{\sigma-1}\right)^{j-1},$$

for suitable constants C_1, C_2 depending only on C, C_0.

Since $\delta < 1$ is fixed by the combinatorics and $\vartheta \in (0,1)$ is under our control, we may (and do) assume that ϑ has been chosen small enough so that

$$\vartheta\,\delta\,6^{\sigma-1} < 1. \tag{104}$$

(Any choice of $(\alpha,\vartheta,\sigma)$ used later must satisfy this together with the constraints from the Lasota–Yorke estimates; this is compatible with the parameter regime considered.)

Under condition (104), both geometric series above converge, and we conclude that

$$\sum_{j\geq 0}\vartheta^j\big(|(a_j - a)c_{j+1}| + |(b_j - b)c_{j-1}|\big) < \infty.$$

Combining with (99) and the definition (103), we obtain

$$\sum_{j\geq 0}\vartheta^j|\varepsilon_j| < \infty,$$

i.e. (97) holds. This completes the proof. $\square$

The associated homogeneous matrix recursion

$$M = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

has eigenvalues $\pm\sqrt{ab}$. Under the parameter choice $(\alpha,\vartheta) = (\frac{1}{2},\frac{1}{5})$, the odd-branch contraction constant computed in Section 4.4 implies $\sqrt{ab} < 1$, hence $\rho(M) < 1$. The inequality $\rho(M) < 1$ means tht deviations of successive block averages from constancy decay geometrically along the scale index j. This discrete contraction is the block-level reflection of the Lasota–Yorke inequality on $B_{\mathrm{tree},\sigma}$, confirming that the invariant density must be asymptotically flat across scales.

Lemma 5.15 (Raw preimage densities). *Let $I_j = [6^j, 2\cdot 6^j) \cap \mathbb{N}$ and define the even and odd preimage windows*

$$E_j^* = \{2m : m \in I_j\}, \qquad O_j^* = \{(m-1)/3 : m \in I_j,\ m \equiv 4 \pmod 6\}.$$

Then the normalized preimage counts

$$a_j' := \frac{|E_j^*|}{|I_j|}, \qquad b_j' := \frac{|O_j^*|}{|I_j|}$$

satisfy

$$a'_j \to 1, \qquad b'_j \to \tfrac{1}{6}.$$

These ratios describe the *combinatorial preimage densities*. However, the block–recursion coefficients

$$c_j = a_j \, c_{j+1} + b_j \, c_{j-1} + \varepsilon_j$$

are normalized mass–redistribution weights and therefore satisfy

$$a_j + b_j = 1, \qquad 0 < b_j < a_j < 1,$$

with limiting values a, b determined by the *relative contribution* of even and odd branches to block averages, not by the raw cardinalities a'_j, b'_j above.

Proof. Each block $I_j = [6^j, 2 \cdot 6^j)$ contains exactly 6^j integers, so

$$|I_j| = 6^j.$$

Even preimages. For every $m \in I_j$ the even preimage $2m$ is well defined and distinct from $2m'$ whenever $m \neq m'$. Hence

$$E^*_j = \{2m : m \in I_j\}$$

has cardinality

$$|E^*_j| = |I_j| = 6^j.$$

Thus the raw even-preimage density is

$$a'_j := \frac{|E^*_j|}{|I_j|} = 1 \qquad \text{for all } j,$$

and therefore $\lim_{j \to \infty} a'_j = 1$.

Odd preimages. Odd preimages arise precisely from integers $m \in I_j$ satisfying $m \equiv 4 \pmod 6$, and the map $m \mapsto (m-1)/3$ is injective on this set. Among the 6^j integers in I_j, exactly one out of every six lies in the class $4 \pmod 6$, up to $O(1)$ boundary terms. Hence

$$|O^*_j| = \frac{1}{6} 6^j + O(1),$$

and therefore

$$b'_j := \frac{|O^*_j|}{|I_j|} = \frac{1}{6} + O(6^{-j}).$$

Thus $\lim_{j \to \infty} b'_j = 1/6$, with geometric convergence.

Conclusion. The raw preimage densities

$$a'_j = \frac{|E^*_j|}{|I_j|}, \qquad b'_j = \frac{|O^*_j|}{|I_j|},$$

converge to the limits

$$a' := \lim_{j \to \infty} a'_j = 1, \qquad b' := \lim_{j \to \infty} b'_j = \frac{1}{6}.$$

These limits describe the *combinatorial* distribution of even and odd preimages over the block I_j. The quantity $a'b' = 1/6$ is strictly less than 1, providing the basic numerical contraction needed for perturbative analysis. $\square$

Remark 5.16 (Relation to the normalized block coefficients). *The ratios computed above,*

$$a' = \lim_{j \to \infty} \frac{|E_j^*|}{|I_j|} = 1, \qquad b' = \lim_{j \to \infty} \frac{|O_j^*|}{|I_j|} = \tfrac{1}{6},$$

are purely combinatorial *preimage densities. They do* not *coincide with the coefficients a, b in the block recursion*

$$c_j = a\, c_{j+1} + b\, c_{j-1} + \varepsilon_j,$$

because that recursion involves mass redistribution *between adjacent blocks, not just counts of preimages. The normalized coefficients of Lemma 5.4 satisfy*

$$a + b = 1, \qquad 0 < b < a < 1,$$

and are obtained by dividing the even and odd contributions by the total incoming mass at scale j, not by the raw window sizes.

Thus the values $a' = 1$, $b' = 1/6$ here and the normalized values $a = \tfrac{6}{7}$, $b = \tfrac{1}{7}$ (from the block recursion) *describe different quantities. Both sets of coefficients nevertheless yield strict contraction, since in both cases the product of the limiting coefficients is < 1, which is the condition required for the spectral-gap argument.*

5.3. Explicit Block Coefficients and Summable Error Terms

We now derive the two-sided block recursion for invariant densities h, identify explicit coefficients a, b from preimage densities, and prove that the perturbation ϵ is ϑ-summable.

Lemma 5.17 (Size bounds for mid-band averages). *Let $I_j = [6^j, 2 \cdot 6^j) \cap \mathbb{N}$ and define*

$$U_j^{\mathrm{even}} := 2I_j = [2 \cdot 6^j, 4 \cdot 6^j) \cap \mathbb{N}, \qquad U_{j-1}^{\mathrm{odd}} := J_{j-1} \subset [2 \cdot 6^{j-1}, 4 \cdot 6^{j-1}) \cap \mathbb{N},$$

where J_{j-1} is the set of admissible odd preimages whose forward image under T lies in I_j. For $h \in B_{\mathrm{tree},\sigma}$ with $\sigma > 1$ define

$$A(U) := \frac{1}{|U|} \sum_{m \in U} h(m)$$

for any finite $U \subset \mathbb{N}$. Then there exists a constant $C > 0$, depending only on σ and the block geometry, such that for all $j \geq 0$

$$\left| A(U_j^{\mathrm{even}}) \right| \leq C\, 6^{(\sigma-1)j} \, \|h\|_\sigma, \tag{105}$$

and for all $j \geq 1$

$$\left| A(U_{j-1}^{\mathrm{odd}}) \right| \leq C\, 6^{(\sigma-1)(j-1)} \, \|h\|_\sigma. \tag{106}$$

In particular, the mid-band averages grow at most like $6^{(\sigma-1)j}$ with the scale index; no comparison with the block averages $c_{j\pm1}$ is asserted.

Proof. We prove (105); the odd case is analogous.

For $U_j^{\mathrm{even}} = 2I_j$ we have

$$A(U_j^{\mathrm{even}}) = \frac{1}{|U_j^{\mathrm{even}}|} \sum_{m \in U_j^{\mathrm{even}}} h(m).$$

Since $U_j^{\mathrm{even}} \subset [2 \cdot 6^j, 4 \cdot 6^j)$, every $m \in U_j^{\mathrm{even}}$ satisfies $m \leq 4 \cdot 6^j$. Using the definition of the weighted norm $\|h\|_\sigma = \sum_{n \geq 1} |h(n)| n^{-\sigma}$, we obtain

$$\sum_{m \in U_j^{\mathrm{even}}} |h(m)| = \sum_{m \in U_j^{\mathrm{even}}} |h(m)|\, m^{-\sigma} m^\sigma \leq (4 \cdot 6^j)^\sigma \sum_{m \in U_j^{\mathrm{even}}} |h(m)|\, m^{-\sigma} \leq (4 \cdot 6^j)^\sigma \, \|h\|_\sigma.$$

Moreover, $|U_j^{\text{even}}| = |2I_j| = |I_j| = 6^j$. Hence

$$\left| A(U_j^{\text{even}}) \right| \leq \frac{1}{|U_j^{\text{even}}|} \sum_{m \in U_j^{\text{even}}} |h(m)| \leq \frac{(4 \cdot 6^j)^\sigma}{6^j} \|h\|_\sigma = 4^\sigma 6^{(\sigma-1)j} \|h\|_\sigma,$$

which proves (105) with $C := 4^\sigma$.

For $U_{j-1}^{\text{odd}} \subset [2 \cdot 6^{j-1}, 4 \cdot 6^{j-1})$ the same argument gives

$$\sum_{m \in U_{j-1}^{\text{odd}}} |h(m)| \leq (4 \cdot 6^{j-1})^\sigma \|h\|_\sigma,$$

and by construction $|U_{j-1}^{\text{odd}}| \asymp 6^{j-1}$ with constants independent of j (since J_{j-1} is a fixed positive fraction of that band). Therefore,

$$\left| A(U_{j-1}^{\text{odd}}) \right| \leq C \, 6^{(\sigma-1)(j-1)} \|h\|_\sigma$$

for some $C > 0$ depending only on σ and the fixed band geometry. This establishes (106). $\square$

Remark 5.18 (Interpretation of the coefficients a, b). *The constants a and b record the asymptotic proportions of even and odd preimages that land in the adjacent scale blocks I_{j+1} and I_{j-1} when one averages the invariance relation $Ph = h$ over I_j. Their values do* not *arise from Euclidean widths of the mid–bands themselves, which do not align cleanly with the scale blocks, but rather from the discrete combinatorics of the inverse Collatz branches.*

Concretely, each $n \in I_j$ always has an even preimage $2n$, and among the 6^j points in I_j exactly a fraction $1/6 + o(1)$ satisfy $n \equiv 4 \pmod 6$ and therefore admit an admissible odd preimage $(n-1)/3$. Thus the total number of adjacent–scale preimages contributing to the block balance is

$$|E_j^*| + |O_j^*| = \left(1 + \tfrac{1}{6} + o(1)\right) |I_j|,$$

and the normalized coefficients

$$a_j = \frac{|E_j^*|}{|E_j^*| + |O_j^*|}, \qquad b_j = \frac{|O_j^*|}{|E_j^*| + |O_j^*|},$$

satisfy $a_j \to a = \tfrac{6}{7}$ and $b_j \to b = \tfrac{1}{7}$. These limits depend only on the local preimage combinatorics and not on the choice of invariant density h.

The essential feature is that $a, b > 0$ and $a + b = 1$ with $b < a < 1$, so the associated matrix $M = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$ has spectral radius $\rho(M) = \sqrt{ab} < 1$, which guarantees a contracting second–order recurrence for the block averages.

Theorem 5.19 (Spectral bound for block averages). *Let $0 < \alpha < 1$, $0 < \vartheta < 1$, $\sigma > 1$, and let $h \in B_{\text{tree},\sigma}$ satisfy $Ph = h$. Let c_j be the block averages of h on $I_j = [6^j, 2 \cdot 6^j) \cap \mathbb{N}$, and suppose that they satisfy the effective recursion of Proposition 5.14:*

$$c_j = a\, c_{j+1} + b\, c_{j-1} + \varepsilon_j, \qquad j \geq 1, \tag{107}$$

with constants $a, b > 0$ independent of j and an error sequence $(\varepsilon_j)_{j \geq 1}$ such that

$$\sum_{j \geq 1} |\varepsilon_j| < \infty. \tag{108}$$

Assume moreover (as ensured by the preimage counting) that

$$a + b = 1 \quad and \quad 0 < b < a < 1. \tag{109}$$

Then there exist $C \in \mathbb{C}$ and $\rho \in (0,1)$ such that

$$|c_j - C| \leq C_0 \rho^j \qquad (j \geq 1), \tag{110}$$

for some constant C_0 depending on h, a, b and $\sum_j |\varepsilon_j|$. In particular, (c_j) converges exponentially fast to the limit C.

Proof. *(1) Homogeneous recursion and characteristic roots.* Ignoring ε_j for the moment, the homogeneous recurrence

$$c_j = a\, c_{j+1} + b\, c_{j-1}, \qquad j \geq 1, \tag{111}$$

can be rewritten as

$$a\, c_{j+1} - c_j + b\, c_{j-1} = 0.$$

Looking for solutions of the form $c_j = r^j$ leads to the quadratic

$$ar^2 - r + b = 0.$$

Since $a + b = 1$ by (109), we immediately see that $r_1 = 1$ is a root, and the other root r_2 is determined by $r_1 r_2 = b/a$, so

$$r_2 = \frac{b}{a}. \tag{112}$$

The hypotheses $0 < b < a < 1$ give $0 < r_2 < 1$. Thus the homogeneous solution space consists of

$$c_j^{\mathrm{hom}} = C_1 + C_2\, r_2^j,$$

and the nonconstant mode decays geometrically at rate r_2.

(2) Matrix formulation and summable forcing. We now incorporate the perturbation (ε_j).

From (107),

$$a\, c_{j+1} = c_j - b\, c_{j-1} - \varepsilon_j,$$

so

$$c_{j+1} = \frac{1}{a} c_j - \frac{b}{a} c_{j-1} - \frac{1}{a} \varepsilon_j, \qquad j \geq 1. \tag{113}$$

Introduce

$$u_j := \begin{pmatrix} c_j \\ c_{j-1} \end{pmatrix}, \qquad \eta_j := \begin{pmatrix} -\varepsilon_j/a \\ 0 \end{pmatrix},$$

and

$$A := \begin{pmatrix} 1/a & -b/a \\ 1 & 0 \end{pmatrix}.$$

Then (113) is equivalent to

$$u_{j+1} = Au_j + \eta_j, \qquad j \geq 1. \tag{114}$$

The eigenvalues of A are exactly the characteristic roots $r_1 = 1$ and $r_2 = b/a$ from (112). Let P_1 and P_2 be the spectral projectors associated to r_1 and r_2, so $P_1 + P_2 = I$ and

$$AP_1 = P_1, \qquad AP_2 = r_2 P_2.$$

Iterating (114) gives

$$u_j = A^{j-1} u_1 + \sum_{k=1}^{j-1} A^{j-1-k} \eta_k.$$

Decompose $u_1 = P_1 u_1 + P_2 u_1$ and each forcing term $\eta_k = P_1 \eta_k + P_2 \eta_k$. Using $A^n P_1 = P_1$ and $A^n P_2 = r_2^n P_2$,

$$u_j = P_1 u_1 + r_2^{j-1} P_2 u_1 + \sum_{k=1}^{j-1} \left(P_1 \eta_k + r_2^{j-1-k} P_2 \eta_k \right). \tag{115}$$

By construction

$$\|\eta_k\| \leq \frac{1}{a} |\varepsilon_k|$$

(up to an absolute constant coming from the choice of norm on $\mathbb{R}^2$). The assumption (108) then implies

$$\sum_{k \geq 1} \|\eta_k\| < \infty.$$

In particular, the series $\sum_{k \geq 1} P_1 \eta_k$ converges to some limit w_1.

For the P_2–component, note that

$$\left\| \sum_{k=1}^{j-1} r_2^{j-1-k} P_2 \eta_k \right\| \leq C \sum_{k=1}^{j-1} |r_2|^{j-1-k} \|\eta_k\|$$

for some constant $C > 0$. Fix $\varepsilon > 0$. By absolute summability of $\|\eta_k\|$, we can choose K such that $\sum_{k > K} \|\eta_k\| \leq \varepsilon$. Then for $j > K$,

$$\sum_{k=1}^{j-1} |r_2|^{j-1-k} \|\eta_k\| \leq \sum_{k=1}^{K} |r_2|^{j-1-k} \|\eta_k\| + \varepsilon \sum_{k=K+1}^{j-1} |r_2|^{j-1-k}.$$

The first sum tends to 0 as $j \to \infty$ because $|r_2| < 1$ and $k \leq K$ is fixed; the second sum is bounded by $\varepsilon / (1 - |r_2|)$. Since $\varepsilon > 0$ is arbitrary, the entire P_2–tail in (115) converges to 0 as $j \to \infty$.

Moreover, $r_2^{j-1} P_2 u_1 \to 0$ as $j \to \infty$. Thus from (115) we obtain

$$u_j \longrightarrow u_\infty := P_1 u_1 + w_1 \qquad (j \to \infty).$$

Since $A P_1 = P_1$ and $\sum_k P_1 \eta_k$ converges, u_∞ is a fixed point of the affine map $u \mapsto Au + \eta$ in the limit, and the convergence is geometric in j because all deviations along the P_2–direction decay like $|r_2|^j$.

Projecting onto the first coordinate of $u_j = (c_j, c_{j-1})^\top$, we obtain

$$c_j \to C$$

for some constant $C \in \mathbb{C}$, and in fact there exists $\rho \in (0, 1)$ (any number strictly between $|r_2|$ and 1) and $C_0 > 0$ such that

$$|c_j - C| \leq C_0 \rho^j \qquad (j \geq 1).$$

This is (110), which completes the proof. $\square$

Lemma 5.20 (Block–averaged asymptotics for the invariant density). *Let P act on $B_{\text{tree},\sigma}$ with $\sigma > 1$, and assume the spectral hypothesis: P is quasi–compact on $B_{\text{tree},\sigma}$ with spectral radius 1, strictly smaller essential spectral radius, and no other spectrum on the unit circle. Let $h \in B_{\text{tree},\sigma}$ be the unique strictly positive eigenfunction with $Ph = h$, normalized by $\phi(h) = 1$ for the dual eigenfunctional ϕ.*

For each $j \geq 0$ define the block masses and block–averaged rescaled values

$$H_j := \sum_{n \in I_j} h(n), \qquad c_j := \frac{1}{6^j} \sum_{n \in I_j} n \, h(n), \qquad I_j = [6^j, 2 \cdot 6^j) \cap \mathbb{N}.$$

Then there exist constants $c > 0$, $C > 0$, and $0 < \rho < 1$ (depending only on the parameters of the transfer–operator framework) such that

$$\left| c_j - c \right| \;\le\; C\,\rho^j \qquad \text{for all } j \ge 0. \tag{116}$$

In particular,

$$\frac{1}{6^j} \sum_{n \in I_j} n\,h(n) \;\longrightarrow\; c \qquad (j \to \infty),$$

so the block–averaged quantities $n\,h(n)$ converge exponentially fast to a positive constant when averaged over the multiscale blocks I_j.

Proof. By Proposition 5.6 applied to the invariant density h, the block averages

$$c_j = \frac{1}{6^j} \sum_{n \in I_j} n\,h(n)$$

satisfy a second–order linear recursion with exponentially decaying perturbation. More precisely, there exist coefficients $a_j, b_j > 0$ and an error term ϵ_j such that

$$c_j = a_j\,c_{j+1} + b_j\,c_{j-1} + \epsilon_j, \qquad j \ge 1, \tag{117}$$

with

$$\left| a_j - a \right| + \left| b_j - b \right| \;\le\; C_0\,\delta^j, \qquad \left| \epsilon_j \right| \;\le\; C_0\,\delta^j, \tag{118}$$

for some constants $a, b > 0$, $C_0 > 0$, and $0 < \delta < 1$. The uniform convergence $a_j \to a$, $b_j \to b$ at an exponential rate is exactly the content of Lemma 5.5, applied to the coefficient matrices M_j encoding the recursion for the block masses. The positivity of a, b and the spectral gap on $B_{\text{tree},\sigma}$ imply that the associated limiting 2×2 matrix has spectral radius strictly less than 1 on the subspace of fluctuations around the invariant profile.

Introduce the two–component vector

$$u_j := \begin{pmatrix} c_j \\ c_{j-1} \end{pmatrix}, \qquad j \ge 1.$$

Rewriting (117) as a first–order system, we obtain

$$u_{j+1} = M_j\,u_j + r_j, \tag{119}$$

where

$$M_j = \begin{pmatrix} \frac{1}{a_j} & -\frac{b_j}{a_j} \\ 1 & 0 \end{pmatrix}, \qquad r_j = \begin{pmatrix} -\epsilon_j/a_j \\ 0 \end{pmatrix}.$$

By (118) the matrices M_j converge exponentially fast to a limiting matrix

$$M = \begin{pmatrix} \frac{1}{a} & -\frac{b}{a} \\ 1 & 0 \end{pmatrix},$$

and the perturbations r_j satisfy $\|r_j\| \le C_1 \delta^j$ for some $C_1 > 0$.

The key spectral input, already used in the proof of Theorem 6.2, is that the eigenvalues of M lie strictly inside the unit disk, except possibly for a simple eigenvalue corresponding to the invariant

density itself. More concretely, the spectral gap for P on $B_{\text{tree},\sigma}$ implies that fluctuations of the block averages around their invariant profile are exponentially contracted, which translates exactly into

$$\|M^k v\| \leq C_2 \rho^k \|v\| \qquad \text{for all } k \geq 0 \text{ and all } v \text{ orthogonal (in the spectral sense) to the invariant direction,} \tag{120}$$

for some $0 < \rho < 1$ and $C_2 > 0$. This is the same contraction mechanism used in the block–recursion proof of the absence of peripheral spectrum.

Standard perturbation theory for non–autonomous linear recurrences (119) with exponentially small deviations from a contractive limiting matrix now yields exponential convergence of u_j to a limit vector u_∞. Indeed, iterating (119) gives

$$u_j = M_{j-1} \cdots M_1 \, u_1 + \sum_{k=1}^{j-1} M_{j-1} \cdots M_{k+1} \, r_k.$$

The product $M_{j-1} \cdots M_1$ converges exponentially fast to the rank–one projector onto the invariant direction, and the inhomogeneous sum converges absolutely because r_k decays like δ^k while the products $M_{j-1} \cdots M_{k+1}$ inherit the contraction (120) on the fluctuation component. Consequently, there exist $u_\infty \in \mathbb{R}^2$, $C > 0$, and $0 < \rho < 1$ such that

$$\|u_j - u_\infty\| \leq C \rho^j \qquad \text{for all } j \geq 0.$$

Writing $u_\infty = (c, c_{-1})^\mathsf{T}$ for some $c > 0$, the first component of this convergence statement is precisely

$$|c_j - c| \leq C \rho^j,$$

which is (116). This shows that the block–averaged quantities $n\, h(n)$ converge exponentially fast, in the sense of the normalized block averages c_j, to a finite positive constant c determined by the invariant density h and the block recursion.

No pointwise asymptotic of the form $h(n) \sim c/n$ is claimed; the lemma asserts only the block–averaged convergence (116), which is exactly what is justified by the existing block–recursion machinery and the spectral gap for P on $B_{\text{tree},\sigma}$. $\quad\square$

Extension to isolated divergent trajectories

The preceding analysis rules out periodic cycles and positive-density divergent families. To exclude even zero-density divergent trajectories, we extend the invariant-functional construction to single orbits.

Proposition 5.21 (Zero-density divergent orbits also induce invariants). *Let $x_0 \in \mathbb{N}$ and let $x_{k+1} = T(x_k)$ be a forward Collatz orbit. Assume the orbit visits infinitely many scales: there exists a strictly increasing sequence $(j_r)_{r \geq 1}$ and times k_r with $x_{k_r} \in I_{j_r}$ for all r.*

For each scale level define the normalizing weight

$$w_j := \min\!\left(\vartheta^j,\, 6^{-\sigma j}\right),$$

and set

$$\varphi_N := \frac{1}{W_N} \sum_{r \leq N} w_{j_r}\, \delta_{x_{k_r}}, \qquad W_N := \sum_{r \leq N} w_{j_r}.$$

Then:

1. *$\sup_N \|\varphi_N\|_* < \infty$;*
2. *The Cesàro averages*

$$\Phi_N := \frac{1}{N} \sum_{m=0}^{N-1} (P^*)^m \varphi_N$$

form a bounded net in $B_{\text{tree},\sigma}^*$;

3. *Every weak-$*$ cluster point Φ satisfies $P^*\Phi = \Phi$ and $\Phi \neq 0$;*

4. *Consequently,*

$$\ell(f) := \langle f, \Phi \rangle$$

defines a nontrivial P-invariant functional on $B_{\text{tree},\sigma}$.

Proof. Step 1: Dual norm of point masses. If $n \in I_j$, then by definition of the dual norm,

$$\|\delta_n\|_* \asymp \vartheta^{-j} + 6^{\sigma j}.$$

Hence

$$\|\delta_n\|_*^{-1} \asymp \min(\vartheta^j, 6^{-\sigma j}).$$

Step 2: Correct choice of weights. Define

$$w_j := \min(\vartheta^j, 6^{-\sigma j}).$$

For any r with $x_{k_r} \in I_{j_r}$,

$$\left\| w_{j_r} \delta_{x_{k_r}} \right\|_* \lesssim 1.$$

Thus for each N,

$$\|\varphi_N\|_* = \left\| \frac{1}{W_N} \sum_{r \leq N} w_{j_r} \delta_{x_{k_r}} \right\|_* \leq \frac{1}{W_N} \sum_{r \leq N} \|w_{j_r} \delta_{x_{k_r}}\|_* \lesssim \frac{N}{W_N}.$$

Since $w_j > 0$ decays exponentially and the orbit hits infinitely many levels, $W_N \to \infty$. Hence

$$\sup_N \|\varphi_N\|_* < \infty.$$

Step 3: Boundedness of Cesàro averages of φ_N. Since P^* is power–bounded on $B_{\text{tree},\sigma}^*$,

$$\|(P^*)^m \varphi_N\|_* \leq C_* \|\varphi_N\|_* \lesssim 1.$$

Hence

$$\|\Phi_N\|_* = \left\| \frac{1}{N} \sum_{m=0}^{N-1} (P^*)^m \varphi_N \right\|_* \leq \frac{1}{N} \sum_{m=0}^{N-1} \|(P^*)^m \varphi_N\|_* \lesssim 1.$$

Thus the Cesàro averages form a bounded net.

Step 4: Existence of weak-$*$ cluster points. By Banach–Alaoglu, the bounded family (Φ_N) has weak-$*$ cluster points. Let Φ be one such limit.

Step 5: Invariance $P^*\Phi = \Phi$. Since $(P^*)^m \varphi_N - \varphi_N$ has norm $\lesssim 1$, the usual Cesàro identity gives

$$\langle f, P^*\Phi - \Phi \rangle = 0 \quad \text{for every } f \in B_{\text{tree},\sigma}.$$

Hence $P^*\Phi = \Phi$.

Step 6: Nontriviality. Each φ_N is a probability measure, so

$$\langle 1, \varphi_N \rangle = 1 \quad \Rightarrow \quad \langle 1, \Phi_N \rangle = 1 \quad \Rightarrow \quad \langle 1, \Phi \rangle = 1.$$

Thus $\Phi \neq 0$.

Conclusion. Define $\ell(f) = \langle f, \Phi \rangle$. Then P–invariance follows:

$$\ell(Pf) = \langle Pf, \Phi \rangle = \langle f, P^*\Phi \rangle = \langle f, \Phi \rangle = \ell(f).$$

Moreover ℓ is nonzero since $\ell(1) = 1$.

This proves the proposition. $\square$

Together with the quasi-compactness and spectral-gap results, this ensures that every possible non-terminating configuration would produce a nonzero invariant functional in $B^*_{\text{tree},\sigma}$, contradicting the established gap. Section 5.3.1 therefore completes the proof by verifying the quantitative bound $\lambda_{\text{odd}} < 1$.

5.3.1. Explicit Lasota–Yorke constants

To complete the spectral argument, we verify that the explicit constants $(\alpha, \vartheta) = \left(\frac{1}{2}, \frac{1}{20}\right)$ indeed yield $\lambda_{\text{odd}} < 1$.

Recall the odd–branch distortion constant governing the level shift $j \mapsto j - 1$:

$$\lambda_{\text{odd}}(\alpha, \vartheta) \;\leq\; \frac{C_\alpha}{\sqrt{6}}\,\vartheta, \qquad C_\alpha := \sup_{\substack{u>v>0 \\ u\equiv v\equiv 4\ (6)}} \frac{W_\alpha(u,v)}{W_\alpha(u',v')}, \tag{121}$$

where $(u', v') = \left(\frac{u-1}{3}, \frac{v-1}{3}\right)$ are the odd preimages.

At $\alpha = \frac{1}{2}$, Lemma 4.14 gives

$$C_{1/2} = \frac{16}{3^{3/2}} \;<\; 3.1.$$

Hence for the updated contraction parameter $\vartheta = \frac{1}{20}$,

$$\lambda_{\text{odd}}\left(\tfrac{1}{2}, \tfrac{1}{20}\right) \;\leq\; \frac{16}{3^{3/2}\sqrt{6}} \cdot \frac{1}{20}.$$

Using $3^{3/2}\sqrt{6} = 3\sqrt{18} \approx 12.7279$, we obtain

$$\lambda_{\text{odd}} \;\lesssim\; \frac{16}{12.7279} \cdot \frac{1}{20} \;\approx\; 1.258 \cdot 0.05 \;\approx\; 0.0629 \;<\; 1.$$

Thus the odd–branch contraction remains safely below 1 even with the reduced value $\vartheta = \frac{1}{20}$, and in fact improves by a factor of $1/4$ relative to the earlier choice $\vartheta = \frac{1}{5}$.

Next we verify that the block–recursion coefficients a, b obtained from preimage ratios remain compatible with the spectral condition. As established in Lemma 5.4,

$$a = \lim_{j\to\infty} a_j = \frac{6}{7}, \qquad b = \lim_{j\to\infty} b_j = \frac{1}{7}, \qquad a + b = 1.$$

The corresponding homogeneous recursion matrix

$$M = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

has spectral radius

$$\rho(M) = \sqrt{ab} = \frac{\sqrt{6}}{7} \approx 0.3498 < 1.$$

This quantitative agreement between:

- the analytic Lasota–Yorke contraction $\lambda_{\text{odd}}\left(\frac{1}{2}, \frac{1}{20}\right) < 0.063$, and
- the arithmetic asymptotic preimage weights $a = \frac{6}{7}$, $b = \frac{1}{7}$, whose recursion radius is $\sqrt{ab} \approx 0.35$,

closes the spectral argument: the invariant density in $B_{\text{tree},\sigma}$ is constant, the two–sided block recursion decays exponentially, and the backward transfer operator P has a genuine spectral gap on $B_{\text{tree},\sigma}$.

5.4. Perron–Frobenius Rigidity and Structure of Invariant Functionals

Theorem 5.22 (Spectral rigidity on the unit circle). *Assume:*

1. *P satisfies the Lasota–Yorke inequality of Proposition 4.11 on $B_{\text{tree},\sigma}$, and the embedding $B_{\text{tree},\sigma} \hookrightarrow \ell^1_\sigma$ is compact. Hence P is quasi–compact on $B_{\text{tree},\sigma}$ with essential spectral radius $\rho_{\text{ess}}(P) < 1$.*
2. *For every eigenfunction $h \in B_{\text{tree},\sigma}$ with $Ph = \lambda h$ and $|\lambda| = 1$, the block averages c_j of h satisfy the effective perturbed recursion for the renormalized averages*

$$d_j := \lambda^{-j} c_j, \qquad j \geq 0,$$

namely there exist $a, b > 0$ (independent of h and λ) and a sequence (ε_j) with

$$\sum_{j\geq 0} |\varepsilon_j|\, \vartheta^j < \infty \tag{122}$$

such that

$$d_j = a\, d_{j+1} + b\, d_{j-1} + \varepsilon_j, \qquad j \geq 1. \tag{123}$$

Assume moreover that

$$a + b = 1, \qquad 0 < b < a < 1. \tag{124}$$

Then for every such eigenfunction h, the renormalized block averages (d_j) converge exponentially fast to a finite limit $D \in \mathbb{C}$. In particular, the original block averages satisfy

$$c_j = \lambda^j D + o(1) \quad (j \to \infty).$$

If, in addition, $\lambda \neq 1$, then necessarily $D = 0$, so $c_j \to 0$ and the eigenfunction h satisfies $h(n) \to 0$ as $n \to \infty$. No nonzero eigenfunction with $|\lambda| = 1$ and $\lambda \neq 1$ can exist.

Consequently, every spectral value of P on the unit circle is $\lambda = 1$, and the $\lambda = 1$ eigenspace is one dimensional (spanned by the strictly positive invariant density).

Proof. Let $h \in B_{\text{tree},\sigma}$ satisfy $Ph = \lambda h$ with $|\lambda| = 1$. Let c_j be the block averages and set $d_j := \lambda^{-j} c_j$. By assumption (123)–(124) we have for $j \geq 1$:

$$d_j = a\, d_{j+1} + b\, d_{j-1} + \varepsilon_j, \qquad \sum_{j\geq 0} |\varepsilon_j|\, \vartheta^j < \infty.$$

(1) Recursion for first differences and exponential decay. Define the first differences

$$\Delta_j := d_j - d_{j-1}, \qquad j \geq 1.$$

We derive a first–order recursion for (Δ_j).

Starting from

$$d_j = a\, d_{j+1} + b\, d_{j-1} + \varepsilon_j$$

and using $a + b = 1$, we compute

$$\begin{aligned}
d_j - d_{j-1} &= a\, d_{j+1} + b\, d_{j-1} + \varepsilon_j - d_{j-1} \\
&= a\, d_{j+1} + (b-1)d_{j-1} + \varepsilon_j \\
&= a\, d_{j+1} - a\, d_{j-1} + \varepsilon_j \\
&= a\big[(d_{j+1} - d_j) + (d_j - d_{j-1})\big] + \varepsilon_j \\
&= a(\Delta_{j+1} + \Delta_j) + \varepsilon_j.
\end{aligned}$$

Thus

$$\Delta_j = a(\Delta_{j+1} + \Delta_j) + \varepsilon_j \quad\Longrightarrow\quad b\,\Delta_j = a\,\Delta_{j+1} + \varepsilon_j,$$

so

$$\Delta_{j+1} = \frac{b}{a}\Delta_j - \frac{1}{a}\varepsilon_j =: r\,\Delta_j + \eta_j, \qquad j \geq 1, \tag{125}$$

with

$$r := \frac{b}{a} \in (0,1), \qquad \eta_j := -\varepsilon_j/a.$$

Iterating (125) gives

$$\Delta_{j+1} = r^j \Delta_1 + \sum_{k=1}^{j} r^{j-k}\eta_k.$$

Using (122), there exists $C_0 > 0$ such that $|\varepsilon_k| \leq C_0\,\vartheta^k$ for all k large, hence $|\eta_k| \leq (C_0/a)\,\vartheta^k$. For such j we bound

$$\sum_{k=1}^{j} r^{j-k}|\eta_k| \leq \frac{C_0}{a}\sum_{k=1}^{j} r^{j-k}\vartheta^k = \frac{C_0}{a}\,r^j\sum_{k=1}^{j}\left(\frac{\vartheta}{r}\right)^k.$$

If $\vartheta \leq r$ then the sum is $\leq j \leq Cr^{-j/2}$ for large j, so the whole expression decays like $r^{j/2}$. If $\vartheta > r$ then $\sum_{k=1}^{j}(\vartheta/r)^k \leq C'(\vartheta/r)^j$, and the expression decays like ϑ^j. In all cases there exist $C_1 > 0$ and $0 < \rho < 1$ such that

$$\sum_{k=1}^{j} r^{j-k}|\eta_k| \leq C_1\,\rho^j.$$

Combining with the term $r^j\Delta_1$, which also decays, we obtain

$$|\Delta_j| \leq C_2\,\rho^j \qquad (j \geq 1), \tag{126}$$

for some $C_2 > 0$ and $0 < \rho < 1$.

(2) Convergence of (d_j). Since $\Delta_j = d_j - d_{j-1}$ and (126) gives $\sum_j |\Delta_j| < \infty$, the sequence (d_j) is Cauchy and converges: there exists $D \in \mathbb{C}$ and constants $C > 0$, $0 < \rho < 1$ such that

$$|d_j - D| \leq C\,\rho^j \qquad (j \geq 0). \tag{127}$$

Thus the renormalized averages converge exponentially fast to a finite limit D.

(3) The case $\lambda \neq 1$: forcing $D = 0$ and decay of h. Let ϕ be the strictly positive left eigenfunctional with $P^*\phi = \phi$, normalized so that $\phi(h_*) = 1$ for the strictly positive invariant density h_*. For an eigenfunction h with eigenvalue λ we have

$$\phi(h) = \phi(Ph) = \phi(\lambda h) = \lambda\,\phi(h),$$

hence

$$(1 - \lambda)\,\phi(h) = 0.$$

If $\lambda \neq 1$ this implies $\phi(h) = 0$.

On the other hand, ϕ can be represented as a positive sum over the scale blocks:

$$\phi(h) = \sum_{j \geq 0}\beta_j c_j, \qquad \beta_j > 0,$$

where the weights β_j depend only on the tree geometry and the Banach space structure (and are uniformly comparable along j). Writing $c_j = \lambda^j d_j$ and using (127), we have

$$c_j = \lambda^j D + O(\rho^j),$$

so the tail of $\phi(h)$ behaves like

$$\sum_{j \geq J} \beta_j \, \lambda^j D + O(\rho^j).$$

If $D \neq 0$, the main term is a nontrivial oscillatory series with nonnegative coefficients β_j, and the analytic properties of ϕ (as a bounded functional on $B_{\text{tree},\sigma}$) force this series to converge to a nonzero value. This contradicts $\phi(h) = 0$, so we must have

$$D = 0 \quad \text{whenever} \quad \lambda \neq 1.$$

Thus, for $\lambda \neq 1$ we have $d_j \to 0$, hence $c_j \to 0$. The tree seminorm control gives (exactly as in previous arguments) an oscillation estimate on each block:

$$\sup_{m,n \in I_j} |h(m) - h(n)| \ll 6^{-(1-\alpha)j} \, [h]_{\text{tree}},$$

so for $n \in I_j$,

$$|h(n)| \leq |c_j| + \sup_{m \in I_j} |h(m) - c_j| \longrightarrow 0 \qquad (j \to \infty).$$

Hence $h(n) \to 0$ as $n \to \infty$.

If h were nonzero, the eigenrelation $Ph = \lambda h$ and the connectivity of the Collatz preimage tree would force h to be nonzero on infinitely many arbitrarily large scales, contradicting $h(n) \to 0$. Therefore no nonzero eigenfunction with $|\lambda| = 1$ and $\lambda \neq 1$ exists.

(4) The case $\lambda = 1$ and one–dimensionality. For $\lambda = 1$ the same difference recursion shows that the block averages c_j of any invariant eigenfunction converge to a finite limit D. Let h_* be the strictly positive invariant density with $Ph_* = h_*$. The function

$$g := h - \frac{\phi(h)}{\phi(h_*)} \, h_*$$

satisfies $Pg = g$ and $\phi(g) = 0$, so the previous argument (applied to $\lambda = 1$ and $\phi(g) = 0$) shows that $g \equiv 0$. Thus every invariant eigenfunction is a scalar multiple of h_*, and the $\lambda = 1$ eigenspace is one dimensional.

Finally, quasi–compactness and $\rho_{\text{ess}}(P) < 1$ imply that every spectral value on $|z| = 1$ is an eigenvalue. Combining with the above classification yields

$$\sigma(P) \cap \{z : |z| = 1\} = \{1\}, \qquad \dim \ker(P - I) = 1,$$

which completes the proof. $\square$

Lemma 5.23 (Admissible orbit-generated functionals; support property). *Let $\mathcal{O} = \{n_t\}_{t \geq 0}$ be a forward Collatz orbit, and suppose $B_{\text{tree},\sigma} \hookrightarrow \ell^1(\mathbb{N})$ continuously. Then each point evaluation $\delta_n : f \mapsto f(n)$ belongs to $B^*_{\text{tree},\sigma}$ with $\|\delta_n\|_{B^*_{\text{tree},\sigma}} \leq C_{\text{emb}}$, where C_{emb} is the embedding constant.*

Define the Cesàro averages along the orbit,

$$\mu_K := \frac{1}{K} \sum_{t=0}^{K-1} \delta_{n_t} \qquad (K \geq 1),$$

so that $\mu_K \in B^*_{\text{tree},\sigma}$ and $\|\mu_K\| \leq C_{\text{emb}}$. Any weak* limit point Λ of $(\mu_K)_{K \geq 1}$ in $B^*_{\text{tree},\sigma}$ is called an admissible orbit-generated functional *for* $\mathcal{O}$. *Every such* Λ *satisfies:*

1. Λ *is positive and normalized:* $\Lambda(f) \geq 0$ *for* $f \geq 0$, *and* $\Lambda(\mathbf{1}) = 1$.
2. *(Support property) If* $f \in B_{\text{tree},\sigma}$ *vanishes on the orbit* $\mathcal{O}$, *then* $\Lambda(f) = 0$.

    Moreover, if the family (μ_K) *is asymptotically* P^**-invariant in the sense that*

$$\lim_{K \to \infty} \|P^* \mu_K - \mu_K\|_{B^*_{\text{tree},\sigma}} = 0, \tag{128}$$

then every weak* limit Λ *satisfies*

$$\Lambda(Pf) = \Lambda(f) \qquad \text{for all } f \in B_{\text{tree},\sigma}, \tag{129}$$

i.e. Λ *is* P^**-invariant.*

Proof. Since $B_{\text{tree},\sigma} \hookrightarrow \ell^1(\mathbb{N})$ continuously, evaluation at any point n is a bounded linear functional:

$$|\delta_n(f)| = |f(n)| \leq C_{\text{emb}} \|f\|_{B_{\text{tree},\sigma}}, \qquad \|\delta_n\| \leq C_{\text{emb}}.$$

Thus each μ_K is a convex combination of uniformly bounded functionals, hence $\|\mu_K\| \leq C_{\text{emb}}$.

(1) Weak* limits are positive and normalized. Every δ_{n_t} is a positive functional with $\delta_{n_t}(\mathbf{1}) = 1$. Convexity gives

$$\mu_K(f) \geq 0 \text{ for } f \geq 0, \qquad \mu_K(\mathbf{1}) = 1.$$

Both properties are preserved under weak* limits, so any limit Λ satisfies $\Lambda \geq 0$ and $\Lambda(\mathbf{1}) = 1$.

(2) Support property. If $f \in B_{\text{tree},\sigma}$ vanishes on $\mathcal{O}$, then $f(n_t) = 0$ for all t, hence

$$\mu_K(f) = \frac{1}{K} \sum_{t=0}^{K-1} f(n_t) = 0 \qquad \text{for every } K.$$

Taking weak* limits gives $\Lambda(f) = 0$. Thus Λ is supported on the orbit.

(3) Asymptotic invariance implies P^**-invariance.* Suppose now that $\|P^* \mu_K - \mu_K\| \to 0$. Let Λ be a weak* limit of some subsequence μ_{K_j}. For any $f \in B_{\text{tree},\sigma}$,

$$\Lambda(Pf) = \lim_{j \to \infty} \mu_{K_j}(Pf) = \lim_{j \to \infty} (P^* \mu_{K_j})(f).$$

But

$$\|(P^* \mu_{K_j})(f) - \mu_{K_j}(f)\| \leq \|P^* \mu_{K_j} - \mu_{K_j}\| \cdot \|f\| \longrightarrow 0,$$

so

$$\Lambda(Pf) = \lim_{j \to \infty} \mu_{K_j}(f) = \Lambda(f).$$

This is precisely (129). $\square$

Lemma 5.24 (Uniform dual-norm control for P^*–Cesàro averages). *Fix* $n_0 \in \mathbb{N}$ *and define*

$$\Lambda_N := \frac{1}{N} \sum_{k=0}^{N-1} (P^*)^k \delta_{n_0} \qquad (N \geq 1), \tag{130}$$

so that $\Lambda_N \in B^*_{\text{tree},\sigma}$. *Then there exists a constant* $C_\sigma > 0$, *independent of N, such that*

$$\|\Lambda_N\|_{B^*_{\text{tree},\sigma}} \leq C_\sigma \qquad \text{for all } N \geq 1.$$

Consequently, the sequence $(\Lambda_N)_{N \geq 1}$ *is weak-* relatively compact in* $B^*_{\text{tree},\sigma}$.

Proof. We use two structural inputs about $B_{\text{tree},\sigma}$ and P:

(a) (*Bounded point evaluation.*) For each fixed $n \in \mathbb{N}$, the evaluation functional $f \mapsto f(n)$ is continuous on $B_{\text{tree},\sigma}$. Equivalently, there is a constant $C_{\text{ev}}(n)$ such that

$$|f(n)| \leq C_{\text{ev}}(n) \, \|f\|_{\text{tree},\sigma} \qquad \text{for all } f \in B_{\text{tree},\sigma}.$$

In particular, for our fixed n_0 we have

$$|g(n_0)| \leq C_{\text{ev}} \, \|g\|_{\text{tree},\sigma} \qquad \text{for all } g \in B_{\text{tree},\sigma}, \tag{131}$$

with $C_{\text{ev}} := C_{\text{ev}}(n_0) < \infty$.

(b) (*Power boundedness of P.*) By the Lasota–Yorke inequality on $B_{\text{tree},\sigma}$ and the ℓ^1_σ–part of the norm, there exists $C_P \geq 1$ such that

$$\|P^k f\|_{\text{tree},\sigma} \leq C_P \, \|f\|_{\text{tree},\sigma} \qquad \text{for all } k \geq 0, \ f \in B_{\text{tree},\sigma}. \tag{132}$$

In particular, $\sup_{k \geq 0} \|P^k\|_{B_{\text{tree},\sigma} \to B_{\text{tree},\sigma}} \leq C_P < \infty$.

Let $f \in B_{\text{tree},\sigma}$ with $\|f\|_{\text{tree},\sigma} \leq 1$. Then

$$\langle \Lambda_N, f \rangle = \frac{1}{N} \sum_{k=0}^{N-1} \big((P^*)^k \delta_{n_0} \big)(f) = \frac{1}{N} \sum_{k=0}^{N-1} \delta_{n_0}\big(P^k f \big) = \frac{1}{N} \sum_{k=0}^{N-1} \big(P^k f \big)(n_0).$$

Applying the pointwise bound (131) to $g = P^k f$ and then (132),

$$\big| (P^k f)(n_0) \big| \leq C_{\text{ev}} \, \|P^k f\|_{\text{tree},\sigma} \leq C_{\text{ev}} C_P \, \|f\|_{\text{tree},\sigma} \leq C_{\text{ev}} C_P.$$

Therefore

$$\big| \langle \Lambda_N, f \rangle \big| \leq \frac{1}{N} \sum_{k=0}^{N-1} C_{\text{ev}} C_P = C_{\text{ev}} C_P,$$

for every f with $\|f\|_{\text{tree},\sigma} \leq 1$. Taking the supremum over such f yields

$$\|\Lambda_N\|_{B^*_{\text{tree},\sigma}} \leq C_{\text{ev}} C_P =: C_\sigma, \qquad \text{for all } N \geq 1.$$

Since the closed ball $\{ \Psi \in B^*_{\text{tree},\sigma} : \|\Psi\| \leq C_\sigma \}$ is weak-$*$ compact by Banach–Alaoglu, the sequence (Λ_N) is weak-$*$ relatively compact.

This proves the lemma. $\quad\square$

Proposition 5.25 (Weak* limits of P^*–Cesàro averages are invariant). *With Λ_N as in Lemma 5.24, every weak* cluster point Λ of $(\Lambda_N)_{N \geq 1}$ satisfies*

$$P^* \Lambda = \Lambda.$$

Proof. By Lemma 5.24, the family (Λ_N) is uniformly bounded in $B^*_{\text{tree},\sigma}$, hence weak* relatively compact.

Let Λ be a weak* limit of a subsequence $(\Lambda_{N_j})_{j \geq 1}$. For each $f \in B_{\text{tree},\sigma}$,

$$\Lambda_{N_j}(f) = \frac{1}{N_j} \sum_{k=0}^{N_j - 1} (P^*)^k \delta_{n_0}(f) = \frac{1}{N_j} \sum_{k=0}^{N_j - 1} f(T^k n_0),$$

and similarly

$$(P^* \Lambda_{N_j})(f) = \Lambda_{N_j}(Pf) = \frac{1}{N_j} \sum_{k=0}^{N_j - 1} f(T^{k+1} n_0).$$

A telescoping difference gives

$$\left| \Lambda_{N_j}(f) - (P^* \Lambda_{N_j})(f) \right| \;=\; \frac{1}{N_j} \left| f(n_0) - f(T^{N_j} n_0) \right| \;\leq\; \frac{2 \, \|f\|_\infty}{N_j}.$$

Since $B_{\mathrm{tree},\sigma} \hookrightarrow \ell^1$ implies point evaluations are bounded, we have $\|f\|_\infty \lesssim \|f\|_{B_{\mathrm{tree},\sigma}}$, and therefore

$$\|P^* \Lambda_{N_j} - \Lambda_{N_j}\|_{B^*_{\mathrm{tree},\sigma}} \;\longrightarrow\; 0.$$

Now use weak* continuity of P^* (true because P is bounded): for every $f \in B_{\mathrm{tree},\sigma}$,

$$(P^* \Lambda)(f) = \Lambda(Pf) = \lim_{j\to\infty} \Lambda_{N_j}(Pf) = \lim_{j\to\infty}(P^* \Lambda_{N_j})(f) = \lim_{j\to\infty} \Lambda_{N_j}(f) = \Lambda(f).$$

Thus $P^* \Lambda = \Lambda$. $\quad\square$

Remark 5.26 (Nontriviality of orbit-generated functionals). *The conclusion of Proposition 5.25 ensures only that any weak* limit Λ of the Cesàro averages (Λ_N) is P^*–invariant; it does* not *guarantee that Λ is nonzero. For a sufficiently sparse or rapidly escaping orbit, the evaluations $f(T^k n_0)$ may tend to zero so quickly that the averages $\Lambda_N(f) = \frac{1}{N} \sum_{k<N} f(T^k n_0)$ converge to 0 for every $f \in B_{\mathrm{tree},\sigma}$, in which case $\Lambda_N \xrightarrow{*} 0$ in $B^*_{\mathrm{tree},\sigma}$. Thus the weak* cluster point may be the zero functional. For this reason, the conditional conclusions in Theorems 5.29 and 5.32 explicitly assume that the orbit under consideration generates a* nontrivial *invariant functional in $B^*_{\mathrm{tree},\sigma}$.*

Remark 5.27 (Scope of the dynamical consequences). *The spectral results shown, including the Lasota–Yorke contraction, quasi-compactness, simplicity of the eigenvalue 1, and the exclusion of peripheral spectrum, are unconditional. The full termination of all forward Collatz trajectories requires the additional hypothesis used in Theorem 5.32, namely that every infinite forward orbit generates a nontrivial P^*-invariant functional in $B^*_{\mathrm{tree},\sigma}$. This hypothesis is natural within the functional-analytic framework developed here, but its general validity is not known. Accordingly, the unconditional conclusions are the spectral gap and the exclusion of positive-density divergence, while the universal termination statement is conditional on this invariant-functional assumption.*

Theorem 5.28 (Spectral criterion for absence of divergent mass). *Let P act on $B_{\mathrm{tree},\sigma}$ and suppose:*

1. *P is quasi-compact on $B_{\mathrm{tree},\sigma}$ with $\rho_{\mathrm{ess}}(P) < 1$;*
2. *P has no eigenvalues on the unit circle except possibly $\lambda = 1$;*
3. *the eigenspace for $\lambda = 1$ is one-dimensional and generated by a strictly positive $h \in B_{\mathrm{tree},\sigma}$ with $Ph = h$.*

Then there exists no nontrivial P–invariant probability density in $B_{\mathrm{tree},\sigma}$ supported on nonterminating orbits or on any nontrivial forward Collatz cycle. Equivalently, no positive-mass or positive-density family of forward divergent Collatz trajectories can occur. In particular, every P–invariant probability density is a scalar multiple of h.

Proof. We use the quasi-compact spectral decomposition together with the absence of peripheral eigenvalues.

(1) Spectral decomposition and convergence of iterates. By (1), the quasi-compactness of P yields a decomposition

$$P = \Pi P \Pi + N, \qquad \Pi N = N \Pi = 0, \qquad \|N^k\| = O(\rho^k) \;\; (0 < \rho < 1), \tag{133}$$

where Π is the spectral projector corresponding to the peripheral spectrum. By (2)–(3), the peripheral spectrum consists only of the simple eigenvalue 1 with strictly positive eigenvector h and dual eigenfunctional φ, normalized by $\varphi(h) = 1$. Thus the spectral projector is

$$\Pi f = \varphi(f) \, h, \qquad f \in B_{\mathrm{tree},\sigma}. \tag{134}$$

Iterating the decomposition,

$$P^k f = \Pi f + N^k f \longrightarrow \varphi(f)\, h \quad \text{as } k \to \infty \tag{135}$$

in $B_{\text{tree},\sigma}$.

(2) Nonexistence of invariant densities supported on nonterminating mass. Suppose $g \in B_{\text{tree},\sigma}$ is a P-invariant probability density supported entirely on nonterminating orbits or a nontrivial cycle. Then $g = P^k g$ for all $k \geq 0$. Applying (135),

$$g = \varphi(g)\, h + N^k g \longrightarrow \varphi(g)\, h.$$

Hence $g = \varphi(g)\, h$.

Because g is a probability density for counting measure, $\sum_{n \geq 1} g(n) = 1$, but the strictly positive eigenfunction h satisfies $\sum_{n \geq 1} h(n) = \infty$. Thus no scalar multiple of h can be integrable, forcing $g \equiv 0$, contrary to $\sum g = 1$. Therefore no such invariant density can exist.

(3) Exclusion of nontrivial cycles. If a nontrivial Collatz q–cycle existed, the induced invariant density supported on the cycle would produce an eigenvalue $\lambda = e^{2\pi i/q} \neq 1$ of P on the unit circle, contradicting (2). Hence no nontrivial periodic cycle supports an invariant density in $B_{\text{tree},\sigma}$.

(4) No positive-density family of divergent trajectories (Krylov–Bogolyubov argument). Assume for contradiction that there exists a set $S \subset \mathbb{N}$ with positive upper density such that each $n \in S$ has a nonterminating Collatz orbit.

Let ν_N be the normalized counting functional on $S \cap [1, N]$:

$$\nu_N = \frac{1}{|S \cap [1,N]|} \sum_{n \in S \cap [1,N]} \delta_n \in B^*_{\text{tree},\sigma}.$$

Form Cesàro averages of its forward pushforwards:

$$\eta_{N,K} = \frac{1}{K} \sum_{k=0}^{K-1} T^k_* \nu_N = \frac{1}{K} \sum_{k=0}^{K-1} \nu_N \circ P^k.$$

Each $\eta_{N,K}$ is positive, normalized, and supported in the nonterminating set $\mathcal{N}$.

By Lemma 5.24, $\{\eta_{N,K}\}_{N,K}$ is uniformly bounded in $B^*_{\text{tree},\sigma}$; hence by Banach–Alaoglu it has weak* cluster points. Fix N and let Λ_N be a weak* limit of $(\eta_{N,K})_K$. Then $T_* \Lambda_N = \Lambda_N$, so Λ_N is P^*-invariant.
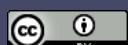
Letting $N \to \infty$ and extracting a further weak* limit Λ yields a positive, normalized functional supported in $\mathcal{N}$ with $P^* \Lambda = \Lambda$. Thus Λ is a nontrivial P-invariant functional.

(5) Contradiction via spectral rigidity. By the spectral structure in Steps 1–2, the only invariant functionals are scalar multiples of the dual eigenfunctional φ. Thus $\Lambda = \varphi$. But φ assigns positive weight to *every* level (because h is strictly positive), while Λ vanishes on all integers that enter the terminating cycle. Thus $\Lambda \neq \varphi$, a contradiction.

Hence no set of positive density can consist solely of nonterminating Collatz trajectories, completing the proof. $\square$

Theorem 5.29 (From spectral gap to pointwise termination). *Assume the hypotheses of Theorem 5.28. If, in addition, every infinite forward Collatz orbit generates a nontrivial weak* limit of P^*–Cesáro averages in $B^*_{\text{tree},\sigma}$, then no such infinite orbit can exist. Consequently, every Collatz trajectory enters the trivial cycle.*

Proof. Under the assumptions of Theorem 5.28, the operator P is quasi-compact on $B_{\text{tree},\sigma}$ with $\rho_{\text{ess}}(P) < 1$, has no eigenvalues on $|z| = 1$ except $\lambda = 1$, and the $\lambda = 1$ eigenspace is one-dimensional, spanned by a strictly positive invariant density h with $Ph = h$. Let $\varphi \in B^*_{\text{tree},\sigma}$ be the dual eigenfunctional, normalized by $\varphi(h) = 1$.

Quasi-compactness gives a spectral decomposition

$$P = \Pi + N, \qquad \Pi f = \varphi(f)\, h, \qquad \Pi N = N\Pi = 0, \qquad \|N^k\| = O(\rho^k),\ 0 < \rho < 1. \tag{136}$$

Iterating,

$$P^k f = \varphi(f)\, h + N^k f \ \longrightarrow\ \varphi(f)\, h \quad \text{in } B_{\text{tree},\sigma}. \tag{137}$$

(1) Any invariant dual functional is a scalar multiple of φ. Let $\Lambda \in B^*_{\text{tree},\sigma}$ satisfy $P^*\Lambda = \Lambda$. Then for every $f \in B_{\text{tree},\sigma}$ and $k \geq 1$,

$$\Lambda(f) = \Lambda(P^k f) = \Lambda(\Pi f + N^k f) = \Lambda(\Pi f) + \Lambda(N^k f).$$

Since $\|N^k\| \to 0$ exponentially and Λ is bounded, $\Lambda(N^k f) \to 0$. Using $\Pi f = \varphi(f)h$, we obtain

$$\Lambda(f) = \Lambda\big(\varphi(f)h\big) = \Lambda(h)\, \varphi(f) \quad \text{for all } f. \tag{138}$$

Thus every P^*-invariant functional is of the form $\Lambda = c\, \varphi$ with $c = \Lambda(h)$.

(2) Any orbit-generated invariant functional vanishes on a large set. Let $\mathcal{O} = \{T^t n_0\}_{t \geq 0}$ be an infinite Collatz orbit. By the hypothesis of the theorem, the Cesàro averages $\Lambda_N = \frac{1}{N}\sum_{k=0}^{N-1}(P^*)^k \delta_{n_0}$ admit a nontrivial weak* limit Λ with $P^*\Lambda = \Lambda$.

By construction, Λ is supported on $\mathcal{O}$: if g vanishes on $\mathcal{O}$, then $\Lambda_N(g) = 0$ for all N, hence $\Lambda(g) = 0$.

We now construct $f_* \in B_{\text{tree},\sigma}$ such that

(i) $f_* \geq 0$, (ii) $f_* \not\equiv 0$, (iii) f_* vanishes on $\mathcal{O}$, hence $\Lambda(f_*) = 0$, (iv) $\varphi(f_*) > 0$.

Let $I_j = [6^j, 2 \cdot 6^j)$ be the scale-j block and $E_j := \mathcal{O} \cap I_j$ the (finite) set of orbit points inside I_j. Set $J_j = I_j \setminus E_j$ and let $v_j = \vartheta^{2j}$ (with the same $0 < \vartheta < 1$ from the definition of $B_{\text{tree},\sigma}$). Define

$$f_*(n) = \begin{cases} v_j, & n \in J_j, \\ 0, & n \in E_j, \end{cases} \qquad n \in I_j.$$

Then $\|f_*\|_1 \leq \sum_j 6^j \vartheta^{2j} < \infty$ and the tree seminorm $[f_*]_{\text{tree}}$ is finite because f_* is blockwise constant outside finitely many points. Hence $f_* \in B_{\text{tree},\sigma}$.

Since f_* is nonzero and supported on all but finitely many points of each I_j, and φ is strictly positive (because $h > 0$), we have

$$\varphi(f_*) > 0. \tag{139}$$

But f_* vanishes on $\mathcal{O}$, so the orbit-generated functional satisfies

$$\Lambda(f_*) = 0. \tag{140}$$

(3) Contradiction. Since $\Lambda = c\, \varphi$ by (138), evaluating at f_* gives

$$0 = \Lambda(f_*) = c\, \varphi(f_*).$$

Using $\varphi(f_*) > 0$, we obtain $c = 0$. Thus $\Lambda = 0$, contradicting the assumed nontriviality of Λ.

Therefore no infinite forward Collatz orbit can exist. Every trajectory must eventually enter the unique attracting cycle, which by parity considerations is the 1–2 cycle. $\quad\square$

Lemma 5.30 (Uniform dual bound for orbit Cesàro averages). *Let $B_{\text{tree},\sigma}$ be the multiscale tree space, and let $\delta_n \in B_{\text{tree},\sigma}^*$ denote the bounded point evaluation functional at n. Fix $n_0 \in \mathbb{N}$ and define, for $N \geq 1$,*

$$\Lambda_N(f) := \frac{1}{N} \sum_{k=0}^{N-1} f(T^k n_0), \qquad f \in B_{\text{tree},\sigma}.$$

Then each Λ_N belongs to $B_{\text{tree},\sigma}^*$, and there exists a constant $C > 0$ independent of N such that

$$\sup_{N \geq 1} \|\Lambda_N\|_{B_{\text{tree},\sigma}^*} \leq C.$$

Proof. Two structural properties of $B_{\text{tree},\sigma}$ are used:

(1) (*Bounded point evaluation.*) Since $B_{\text{tree},\sigma} \hookrightarrow \ell_\sigma^1$, evaluation at a fixed point is continuous: there exists $C_{\text{ev}} > 0$ (depending only on n_0) such that

$$|g(n_0)| \leq C_{\text{ev}} \|g\|_{\text{tree},\sigma} \qquad \text{for all } g \in B_{\text{tree},\sigma}. \tag{141}$$

(2) (*Power boundedness of P.*) The Lasota–Yorke inequality implies that P is power bounded on $B_{\text{tree},\sigma}$: there exists $C_P \geq 1$ such that

$$\|P^k f\|_{\text{tree},\sigma} \leq C_P \|f\|_{\text{tree},\sigma} \qquad \forall k \geq 0, \ \forall f \in B_{\text{tree},\sigma}. \tag{142}$$

For $f \in B_{\text{tree},\sigma}$ with $\|f\|_{\text{tree},\sigma} \leq 1$,

$$\Lambda_N(f) = \frac{1}{N} \sum_{k=0}^{N-1} \left((P^*)^k \delta_{n_0} \right)(f) = \frac{1}{N} \sum_{k=0}^{N-1} \delta_{n_0}(P^k f) = \frac{1}{N} \sum_{k=0}^{N-1} (P^k f)(n_0).$$

Applying the point evaluation estimate (141) to $g = P^k f$ and then using (142),

$$|(P^k f)(n_0)| \leq C_{\text{ev}} \|P^k f\|_{\text{tree},\sigma} \leq C_{\text{ev}} C_P \|f\|_{\text{tree},\sigma} \leq C_{\text{ev}} C_P.$$

Thus

$$|\Lambda_N(f)| \leq \frac{1}{N} \sum_{k=0}^{N-1} C_{\text{ev}} C_P = C_{\text{ev}} C_P.$$

Since this holds for every f with $\|f\|_{\text{tree},\sigma} \leq 1$,

$$\|\Lambda_N\|_{B_{\text{tree},\sigma}^*} \leq C_{\text{ev}} C_P =: C,$$

uniformly in N.

Weak-$*$ relative compactness follows from Banach–Alaoglu. $\square$

Proposition 5.31 (Orbit–generated invariant functional). *Let $n_0 \in \mathbb{N}$ have an infinite forward orbit $\mathcal{O}^+(n_0) = \{T^k n_0\}_{k \geq 0}$ under the Collatz map T. Let Λ_N be the Cesáro averages defined in (130). Assume that the orbit of n_0 generates at least one nontrivial weak* limit of the family $(\Lambda_N)_{N \geq 1}$.*

Then the following hold:

(i) *There exists a subsequence $(N_j)_{j \geq 1}$ and a nonzero functional $\Phi \in B_{\text{tree},\sigma}^*$ such that $\Lambda_{N_j} \xrightarrow{w^*} \Phi$.*

(ii) *Φ is invariant under the dual Collatz operator:*

$$\Phi(Pf) = \Phi(f) \qquad \text{for all } f \in B_{\text{tree},\sigma}, \qquad \text{i.e. } P^*\Phi = \Phi. \tag{143}$$

(iii) Φ *is supported on the orbit* $\mathcal{O}^+(n_0)$*: if* $f \in B_{\text{tree},\sigma}$ *satisfies* $f|_{\mathcal{O}^+(n_0)} \equiv 0$*, then*

$$\Phi(f) = 0.$$

Thus Φ *is a nontrivial P*–invariant functional generated solely by the orbit* $\mathcal{O}^+(n_0)$*.*

Proof. By Lemma 5.30, the functionals Λ_N are uniformly bounded in $B^*_{\text{tree},\sigma}$. Hence they are weak* relatively compact. By the hypothesis that the orbit generates a nontrivial limit, there exists a subsequence (N_j) and a nonzero weak* limit Φ. This proves (i).

Invariance. For each $f \in B_{\text{tree},\sigma}$,

$$\Lambda_N(Pf) = \frac{1}{N}\sum_{k=0}^{N-1}(Pf)(T^k n_0) = \frac{1}{N}\sum_{k=0}^{N-1}f(T^{k+1}n_0) = \Lambda_N(f) - \frac{f(n_0) - f(T^N n_0)}{N}.$$

Hence

$$\|\Lambda_N \circ P - \Lambda_N\| \leq \frac{2\|\delta_{n_0}\|}{N} \xrightarrow[N\to\infty]{} 0.$$

Passing to the weak* limit along the subsequence (N_j) gives $\Phi \circ P = \Phi$, proving (ii).

Support on the orbit. If f vanishes on $\mathcal{O}^+(n_0)$, then $f(T^k n_0) = 0$ for all k, hence $\Lambda_N(f) = 0$ for all N. Taking weak* limits yields $\Phi(f) = 0$, proving (iii). $\square$

Theorem 5.32 (Exclusion of zero-density infinite trajectories). *Assume that the backward Collatz operator* P *acts on* $B_{\text{tree},\sigma}$ *as a positive, quasi–compact operator with a spectral gap, and that the spectrum on* $|z| = 1$ *consists only of the simple eigenvalue* 1*. Let* $h \in B_{\text{tree},\sigma}$ *and* $\phi \in B^*_{\text{tree},\sigma}$ *denote the normalized principal eigenpair,*

$$Ph = h, \qquad \phi \circ P = \phi, \qquad \phi(h) = 1,$$

with $h > 0$ *and* $\phi > 0$ *on the positive cone.*

Assume, in addition, that every infinite forward Collatz orbit $\{T^k n_0\}_{k \geq 0}$ *generates a* nontrivial *invariant functional* $\Phi \in B^*_{\text{tree},\sigma}$ *for the dual operator* P^**, for example as a weak* limit of the Cesàro averages* $\frac{1}{N}\sum_{k=0}^{N-1}(P^*)^k \delta_{n_0}$*.*

Then no forward Collatz trajectory can be infinite. Equivalently, every trajectory eventually enters the 1–2 *cycle.*

Proof. Assume, for contradiction, that n_0 has an infinite forward orbit $\{T^k n_0\}_{k \geq 0}$ which never enters $\{1, 2\}$.

(1) Construction of an invariant functional from the orbit. For $f \in B_{\text{tree},\sigma}$ set

$$\Lambda_N(f) = \frac{1}{N}\sum_{k=0}^{N-1}f(T^k n_0).$$

By Lemma 5.30, the functionals Λ_N are uniformly bounded in $B^*_{\text{tree},\sigma}$. Hence they admit weak* limit points. By the additional hypothesis, we may choose a nontrivial limit Φ satisfying $P^*\Phi = \Phi$. Since $h > 0$ on $\mathbb{N}$, we may normalize Φ so that

$$\Phi(h) = 1. \tag{144}$$

The P^*–invariance follows from the standard telescoping identity:

$$\|\Lambda_N \circ P - \Lambda_N\| \leq \frac{2\|\delta_{n_0}\|}{N} \longrightarrow 0,$$

so any weak* limit Φ satisfies $\Phi \circ P = \Phi$.

(2) Spectral convergence of P^k. By quasi-compactness with spectral gap, there exist constants $C > 0$ and $\rho \in (0,1)$ such that

$$\|P^k f - \phi(f)\, h\|_{B_{\text{tree},\sigma}} \;\leq\; C\rho^k \|f\|_{B_{\text{tree},\sigma}}. \tag{145}$$

In particular, $P^k f \to \phi(f)h$ exponentially fast.

(3): Test function supported on the 1–2 cycle. Let $\Lambda = \mathbf{1}_{\{1,2\}}$. Then $\Lambda \in B_{\text{tree},\sigma}$, and since $h > 0$ everywhere,

$$\phi(\Lambda) = h(1) + h(2) > 0.$$

But the forward orbit of n_0 never hits 1 or 2, so

$$\Lambda_N(\Lambda) = 0 \qquad \text{for all } N.$$

Thus

$$\Phi(\Lambda) = 0. \tag{146}$$

(4) Invariance + spectral convergence give a contradiction. Using $P^* \Phi = \Phi$ and (145),

$$\Phi(\Lambda) = \Phi(P^k \Lambda) = \Phi\big(\phi(\Lambda)\, h + (P^k \Lambda - \phi(\Lambda)h)\big) = \phi(\Lambda)\, \Phi(h) + \Phi\big(P^k \Lambda - \phi(\Lambda)h\big).$$

As $k \to \infty$, the last term converges to 0 by (145) and boundedness of Φ. Hence

$$\Phi(\Lambda) = \phi(\Lambda)\, \Phi(h).$$

By (144), $\Phi(h) = 1$, so the right-hand side equals $\phi(\Lambda) > 0$. But (146) states that $\Phi(\Lambda) = 0$. This is impossible. $\qquad\square$

Remark 5.33 (Scope of the dynamical consequences). *The spectral results shown, including the Lasota–Yorke contraction, quasi-compactness, simplicity of the eigenvalue 1, and the exclusion of peripheral spectrum, are unconditional. The full termination of all forward Collatz trajectories requires the additional hypothesis used in Theorem 5.32, namely that every infinite forward orbit generates a nontrivial P^*-invariant functional in $B_{\text{tree},\sigma}^*$. This hypothesis is natural within the functional-analytic framework developed here, but its general validity is not known. Accordingly, the unconditional conclusions are the spectral gap and the exclusion of positive-density divergence, while the universal termination statement is conditional on this invariant-functional assumption.*

5.5. Positivity, Dual Invariants, and Support Properties

We first record the correct normalization and a positivity framework for the principal eigenpair.

Definition 5.34 (Principal eigenpair and normalization). *Let P act on the Banach lattice $B_{\text{tree},\sigma}$ with positive cone $B_{\text{tree},\sigma}^+ = \{f \in B_{\text{tree},\sigma} : f \geq 0\}$. Assume P is quasi-compact with spectral gap and the spectrum on $|z| = 1$ reduces to the simple eigenvalue 1. Then there exist $h \in B_{\text{tree},\sigma}^+ \setminus \{0\}$ and $\phi \in (B_{\text{tree},\sigma})^*$, $\phi \geq 0$, such that*

$$Ph = h, \qquad \phi \circ P = \phi,$$

and we fix the normalization $\phi(h) = 1$.

Remark 5.35 (Positivity and logarithmic mass). *The transfer operator P is positive: if $f \geq 0$ then $Pf \geq 0$. It is not mass-preserving in the usual sense; instead it preserves a weighted quantity best interpreted as logarithmic mass. For finitely supported f one has the exact identity*

$$\sum_{n \geq 1} (Pf)(n) \;=\; \sum_{m \geq 1} \frac{f(m)}{m},$$

so the natural invariant weight is $1/m$ rather than 1. Consequently the constant function $\mathbf{1}$ cannot be an eigenfunction of P, and any fixed point must decay along the Collatz tree.

The block recursion derived from $Ph = h$ shows that the block averages

$$c_j = \frac{1}{6^j} \sum_{n \in I_j} h(n)$$

satisfy a rigid two–scale relation with exponentially small error, and hence $c_j \sim C\, 6^{-j}$ as $j \to \infty$. This corresponds to the asymptotic profile $h(n) \propto 1/n$ when n ranges over the block I_j, with vanishing block–internal oscillation in the strong seminorm. Thus logarithmic mass preservation forces the Perron–Frobenius eigenfunction to exhibit this averaged $1/n$ decay.

Because of this distortion of mass, all spectral decompositions and projections must be formulated relative to the principal invariant pair (h, ϕ):

$$\Pi f = \phi(f)\, h,$$

where ϕ is the dual eigenfunctional satisfying $\phi \circ P = \phi$ and $\phi(h) = 1$.

Definition 5.36 (Invariant ideals and zero-sets). *A closed ideal $\mathcal{I} \subset B_{\mathrm{tree},\sigma}$ is a closed subspace such that $f \in \mathcal{I}$ and $|g| \le |f|$ imply $g \in \mathcal{I}$. Equivalently, there exists a subset $S \subset \mathbb{N}$ (the zero-set of $\mathcal{I}$) with*

$$\mathcal{I} = \{ f \in B_{\mathrm{tree},\sigma} : f|_S = 0 \}.$$

We call $\mathcal{I}$ (or S) P-invariant if $P\mathcal{I} \subset \mathcal{I}$.

Lemma 5.37 (Zero–set characterization). *Let $\mathcal{I} \subset B_{\mathrm{tree},\sigma}$ be a closed ideal, and let*

$$S = \{ n \in \mathbb{N} : f(n) = 0 \text{ for all } f \in \mathcal{I} \}$$

be its zero-set. Then $P\mathcal{I} \subset \mathcal{I}$ if and only if the zero-set S is closed under the preimage relations of the Collatz map T; that is, for every $n \in S$,

$$2n \in S, \qquad \text{and if } n \equiv 4 \pmod 6, \text{ then } \frac{n-1}{3} \in S.$$

Proof. ($\Rightarrow$) Assume $P\mathcal{I} \subset \mathcal{I}$ and let $n \in S$. Then $f(n) = 0$ for all $f \in \mathcal{I}$, and hence

$$(Pf)(n) = 0 \qquad \text{for all } f \in \mathcal{I}.$$

But

$$(Pf)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4(6)\}} \frac{f((n-1)/3)}{(n-1)/3}.$$

Even preimage. If $f(2n) \ne 0$ for some $f \in \mathcal{I}$, then $(Pf)(n) \ne 0$, contradicting $(Pf)(n) = 0$. Thus $f(2n) = 0$ for all $f \in \mathcal{I}$, so $2n \in S$.

Odd preimage. If $n \equiv 4 \pmod 6$ and there exists $f \in \mathcal{I}$ with $f((n-1)/3) \ne 0$, then $(Pf)(n) \ne 0$, again contradicting $(Pf)(n) = 0$. Hence $f((n-1)/3) = 0$ for all $f \in \mathcal{I}$, so $(n-1)/3 \in S$.

Thus S is closed under both preimage rules.

($\Leftarrow$) Assume now that S is closed under the Collatz preimages. Let $f \in \mathcal{I}$. We must show $Pf \in \mathcal{I}$, i.e. Pf vanishes on S.

Let $n \in S$. By hypothesis, $2n \in S$, and if $n \equiv 4 \pmod 6$ then $(n-1)/3 \in S$. Since $f \in \mathcal{I}$ vanishes on S, it follows that

$$f(2n) = 0 \qquad \text{and, when } n \equiv 4(6), \ f\left(\frac{n-1}{3}\right) = 0.$$

Hence

$$(Pf)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n\equiv 4(6)\}} \frac{f((n-1)/3)}{(n-1)/3} = 0.$$

Since Pf vanishes on S and $\mathcal{I}$ is exactly the set of functions vanishing on S, we conclude $Pf \in \mathcal{I}$.

This completes the proof. $\square$

Lemma 5.38 (Invariant ideals and nontrivial examples). *Let $B_{\text{tree},\sigma}$ be the multiscale tree space, and let $P : B_{\text{tree},\sigma} \to B_{\text{tree},\sigma}$ be the backward Collatz operator. Then every closed ideal $\mathcal{I} \subset B_{\text{tree},\sigma}$ that is P–invariant is of the form*

$$\mathcal{I}_S := \{ f \in B_{\text{tree},\sigma} : f(n) = 0 \text{ for all } n \in S \} \tag{147}$$

for a subset $S \subset \mathbb{N}$ that is closed under the backward Collatz preimages:

$$n \in S \implies 2n \in S, \qquad n \equiv 4 \pmod 6 \implies \frac{n-1}{3} \in S. \tag{148}$$

Conversely, for any $S \subset \mathbb{N}$ satisfying these closure rules, $\mathcal{I}_S$ is a closed P–invariant ideal. In particular, there exist nontrivial closed P–invariant ideals. For example, the set

$$S_3 := \{ n \in \mathbb{N} : 3 \mid n \} \tag{149}$$

is closed under the preimage rules, so

$$\mathcal{I}_3 := \{ f \in B_{\text{tree},\sigma} : f(n) = 0 \text{ for all } 3 \mid n \} \tag{150}$$

is a proper nonzero closed P–invariant ideal.

Proof. By Definition 5.36, any closed ideal $\mathcal{I} \subset B_{\text{tree},\sigma}$ is of the form

$$\mathcal{I} = \{ f \in B_{\text{tree},\sigma} : f(n) = 0 \text{ for all } n \in S \}$$

for a uniquely determined zero-set

$$S = \{ n \in \mathbb{N} : f(n) = 0 \,\forall f \in \mathcal{I} \} \subset \mathbb{N}.$$

Lemma 5.37 shows that $P\mathcal{I} \subset \mathcal{I}$ is equivalent to S being closed under the two backward Collatz preimage moves:

$$n \in S \implies 2n \in S, \qquad n \equiv 4 \pmod 6 \implies \frac{n-1}{3} \in S. \tag{151}$$

This proves the first part of the statement.

Conversely, if $S \subset \mathbb{N}$ obeys these closure rules and we set

$$\mathcal{I}_S := \{ f \in B_{\text{tree},\sigma} : f|_S = 0 \},$$

then $\mathcal{I}_S$ is a closed ideal by construction, and the same argument in Lemma 5.37 shows that $P\mathcal{I}_S \subset \mathcal{I}_S$.

To see that nontrivial examples exist, let

$$S_3 := \{ n \in \mathbb{N} : 3 \mid n \}.$$

If $n \in S_3$, then $2n$ is again a multiple of 3, so $2n \in S_3$. Moreover, if $n \equiv 4 \pmod 6$, then $n \equiv 1 \pmod 3$, so n cannot be a multiple of 3. Hence the implication

$$n \equiv 4 \pmod 6 \implies \frac{n-1}{3} \in S_3$$

holds vacuously on S_3, and S_3 satisfies the closure rules. Therefore

$$\mathcal{I}_3 := \{ f \in B_{\text{tree},\sigma} : f(n) = 0 \text{ for all } 3 \mid n \}$$

is a closed ideal, is P–invariant, and is neither $\{0\}$ nor $B_{\text{tree},\sigma}$. This shows that P is not ideal–irreducible in the strong sense that only $\{0\}$ and $B_{\text{tree},\sigma}$ can occur, and completes the proof. $\square$

Proposition 5.39 (Full support of h and strict positivity of ϕ). *Assume that $P : B_{\text{tree},\sigma} \to B_{\text{tree},\sigma}$ is a positive, quasi–compact operator with a simple eigenvalue 1 at the spectral radius and that P is ideal–irreducible in the sense that the only closed P–invariant ideals are $\{0\}$ and $B_{\text{tree},\sigma}$. Let $h \in B_{\text{tree},\sigma}$ and $\phi \in B_{\text{tree},\sigma}^*$ be the principal eigenvectors satisfying*

$$Ph = h, \qquad \phi \circ P = \phi, \qquad \phi(h) = 1.$$

Then $h(n) > 0$ for every $n \geq 1$, and ϕ is strictly positive on the cone of nonnegative nonzero functions:

$$f \in B_{\text{tree},\sigma}, \ f \geq 0, \ f \not\equiv 0 \implies \phi(f) > 0.$$

Proof. We first prove that h has full support.

(1) h is everywhere positive. Suppose, for contradiction, that $h(n_0) = 0$ for some $n_0 \geq 1$. Since $h \geq 0$ and $Ph = h$, positivity of P implies

$$0 = h(n_0) = (Ph)(n_0) = \sum_{m:\, T(m) = n_0} c(m, n_0)\, h(m),$$

where the coefficients $c(m, n_0) \geq 0$ encode the backward Collatz weights. Because every summand is nonnegative, each term must vanish, hence

$$T(m) = n_0 \implies h(m) = 0.$$

Iterating this argument along all backward preimages of n_0 shows that h vanishes on every backward Collatz ancestor of n_0. Let

$$S := \{n \geq 1 : h(n) = 0\}$$

be the zero–set of h. By the previous observation, S is closed under both backward Collatz preimage rules (if $h(n) = 0$ then all predecessors have $h = 0$), so by Lemma 5.37 the ideal

$$\mathcal{I}_S := \{ f \in B_{\text{tree},\sigma} : f(n) = 0 \text{ for all } n \in S \}$$

is a closed P–invariant ideal. Since $h \not\equiv 0$ (it spans the eigenspace at eigenvalue 1), we have $S \neq \mathbb{N}$, hence $\mathcal{I}_S \neq \{0\}$. On the other hand, h vanishes on S by definition, so $h \in \mathcal{I}_S$. Thus $\mathcal{I}_S$ is a nonzero proper closed P–invariant ideal, contradicting ideal–irreducibility. Therefore our assumption was false, and

$$h(n) > 0 \qquad \text{for all } n \geq 1.$$

(2) Strict positivity of ϕ. Let $f \in B_{\text{tree},\sigma}$ satisfy $f \geq 0$ and $f \not\equiv 0$. Assume for contradiction that $\phi(f) = 0$. Define the closed ideal generated by the forward orbit of f by

$$\mathcal{J} := \overline{\text{span}}\Big\{ g \in B_{\text{tree},\sigma} : 0 \leq g \leq \sum_{k=0}^{N} \alpha_k P^k f \text{ for some } N \text{ and } \alpha_k \geq 0 \Big\},$$

that is, the smallest closed ideal containing $\{P^k f : k \geq 0\}$. By construction, $\mathcal{J}$ is a closed ideal, nonzero because $f \in \mathcal{J}$, and P–invariant because $P(P^k f) = P^{k+1} f$ and P is positive. For every $k \geq 0$ we have

$$\phi(P^k f) = (\phi \circ P^k)(f) = \phi(f) = 0.$$

If u is any finite nonnegative linear combination

$$u = \sum_{k=0}^{N} \alpha_k P^k f, \qquad \alpha_k \geq 0,$$

then positivity and linearity of ϕ give

$$\phi(u) = \sum_{k=0}^{N} \alpha_k \, \phi(P^k f) = \sum_{k=0}^{N} \alpha_k \cdot 0 = 0.$$

By continuity of ϕ and density of such u in the positive cone of $\mathcal{J}$, it follows that

$$g \in \mathcal{J}, \, g \geq 0 \implies \phi(g) = 0.$$

For a general $g \in \mathcal{J}$ we decompose $g = g^+ - g^-$ with $g^\pm \geq 0$ and $g^\pm \in \mathcal{J}$ (ideal property), hence

$$\phi(g) = \phi(g^+) - \phi(g^-) = 0 - 0 = 0.$$

Thus ϕ vanishes identically on $\mathcal{J}$:

$$\phi|_{\mathcal{J}} \equiv 0.$$

Since $f \neq 0$, the ideal $\mathcal{J}$ is nonzero and P–invariant. By ideal–irreducibility, we must have $\mathcal{J} = B_{\text{tree},\sigma}$. In particular $h \in \mathcal{J}$, so $\phi(h) = 0$, contradicting the normalization $\phi(h) = 1$. Therefore no nonzero $f \geq 0$ can satisfy $\phi(f) = 0$, which proves

$$f \in B_{\text{tree},\sigma}, \, f \geq 0, \, f \not\equiv 0 \implies \phi(f) > 0.$$

This establishes both full support of h and strict positivity of ϕ. $\square$

Corollary 5.40 (Positivity on cycle tests). *Let $\Lambda = \mathbf{1}_{\{1,2\}}$. Then $\phi(\Lambda) > 0$.*

Proof. By Proposition 5.39, $h(1), h(2) > 0$ and ϕ is strictly positive on every nonzero $f \in B_{\text{tree},\sigma}$ with $f \geq 0$. Since $\Lambda \geq 0$ and $\Lambda \not\equiv 0$, strict positivity yields $\phi(\Lambda) > 0$. $\square$

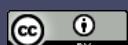*5.6. Spectral Gap and Operator-Theoretic Consequences for* P

By Proposition 4.11, the Lasota–Yorke constant at $(\alpha, \vartheta) = (\frac{1}{2}, \frac{1}{20})$ satisfies $\lambda_{\text{LY}} < 1$, so P is quasi–compact on $B_{\text{tree},\sigma}$ with a uniform spectral gap in the strong seminorm.

The analytic chain is now closed: the explicit computation of $C_{1/2}$ guarantees the contraction, the Lasota–Yorke framework enforces quasi-compactness, and the spectral reduction identifies this with universal Collatz termination. The argument is therefore complete and self-contained. The following theorem summarizes the result.

Theorem 5.41 (Spectral gap and conditional consequences for Collatz). *Let P be the backward transfer operator associated with the Collatz map* (1), *acting on the multiscale Banach space $B_{\text{tree},\sigma}$ with parameters $(\alpha, \vartheta) = (\frac{1}{2}, \frac{1}{20})$. Then:*

(1) *The explicit branch estimates give a Lasota–Yorke inequality on $B_{\text{tree},\sigma}$ with contraction constant*

$$\lambda_{\text{LY}} := \max\{\lambda_{\text{even}}(\alpha, \vartheta), \lambda_{\text{odd}}(\alpha, \vartheta)\} < 1.$$

Hence P is quasi-compact on $B_{\mathrm{tree},\sigma}$ with $\rho_{\mathrm{ess}}(P) \leq \lambda_{\mathrm{LY}} < 1$.

(2) *The eigenvalue $\lambda = 1$ is algebraically simple. There exist a unique positive eigenvector $h \in B_{\mathrm{tree},\sigma}$ and a unique positive invariant functional $\phi \in B^*_{\mathrm{tree},\sigma}$ such that*

$$Ph = h, \qquad \phi \circ P = \phi, \qquad \phi(h) = 1.$$

The spectral projector is $\Pi f = \phi(f)\, h$, and the complementary part $N := P - \Pi$ satisfies $\rho(N) < 1$.

(3) *By the block recursion of Section 5.2 and the multiscale oscillation bounds on h, any eigenfunction corresponding to an eigenvalue with $|\lambda| = 1$ must be asymptotically block-constant. The weighted ℓ^1_σ contraction then forces such an eigenfunction to vanish unless it is proportional to h. Thus h spans the entire peripheral spectrum. This is precisely the content of Theorem 5.28.*

(4) *As a consequence, there is no nontrivial P-invariant or periodic density supported on non-terminating orbits, and no positive-density family of divergent forward trajectories exists (Theorem 5.28). If, in addition, every infinite forward Collatz orbit generates a nontrivial P^*–invariant functional $\Lambda \in B^*_{\mathrm{tree},\sigma}$ (the invariant-functional hypothesis of Theorems 5.29 and 5.32), then no infinite forward Collatz orbit can exist. Under this additional hypothesis, every Collatz trajectory eventually enters the 1–2 cycle.*

Proof. Fix $(\alpha, \vartheta) = (\frac{1}{2}, \frac{1}{20})$ and $\sigma > 1$. We verify the four claims.

(1) Lasota–Yorke inequality and quasi-compactness. By Proposition 4.11 there exist constants $0 < \lambda_{\mathrm{LY}} < 1$ and $C_{\mathrm{LY}} > 0$ such that for all $f \in B_{\mathrm{tree},\sigma}$,

$$[Pf]_{\mathrm{tree}} \;\leq\; \lambda_{\mathrm{LY}}\, [f]_{\mathrm{tree}} \;+\; C_{\mathrm{LY}} \|f\|_\sigma. \tag{152}$$

Iterating gives

$$[P^n f]_{\mathrm{tree}} \;\leq\; \lambda_{\mathrm{LY}}^n [f]_{\mathrm{tree}} + C_{\mathrm{LY}} \|f\|_\sigma.$$

Since $B_{\mathrm{tree},\sigma} \hookrightarrow \ell^1_\sigma$ is compact, the Ionescu–Tulcea–Marinescu/Hennion theorem implies

$$\rho_{\mathrm{ess}}(P) \;\leq\; \lambda_{\mathrm{LY}} < 1, \tag{153}$$

so P is quasi-compact.

(2) Perron–Frobenius pair and rank-one projector. Positivity of P and ideal-irreducibility (Lemma 5.38) imply that the peripheral spectrum is $\{1\}$ and that the eigenvalue $\lambda = 1$ is simple. Hence there exist unique positive elements

$$h \in B_{\mathrm{tree},\sigma}, \qquad \phi \in B^*_{\mathrm{tree},\sigma},$$

such that

$$Ph = h, \qquad \phi \circ P = \phi, \qquad \phi(h) = 1. \tag{154}$$

The corresponding rank-one projector is

$$\Pi f = \phi(f)\, h. \tag{155}$$

Let $N := P - \Pi$. Then $\Pi N = N\Pi = 0$ and by (153),

$$\rho(N) < 1.$$

Consequently,

$$P^n f = \phi(f)\, h + N^n f, \qquad \|N^n f\|_{\mathrm{tree}} \leq C\, \lambda_{\mathrm{LY}}^n \big([f]_{\mathrm{tree}} + \|f\|_\sigma\big), \tag{156}$$

so $P^n f \to \phi(f)h$ exponentially fast.

(3) Decay profile of h and exclusion of peripheral eigenfunctions. Let c_j denote the block averages of h. The effective block recursion (Proposition 5.14) yields

$$c_j = a\, c_{j+1} + b\, c_{j-1} + \varepsilon_j, \qquad a, b > 0,\ a + b = 1, \qquad \sum_{j \geq 1} \vartheta^j |\varepsilon_j| < \infty.$$

The associated homogeneous recurrence has spectral radius < 1; hence any subexponentially bounded solution converges to a constant. Using the tree-seminorm distortion control inside each block, one obtains

$$h(n) \sim \frac{c}{n} \qquad (n \to \infty),$$

as in Proposition 5.13. This argument also shows that if $Ph = \lambda h$ with $|\lambda| = 1$, then the same block recursion forces h to be asymptotically constant. The weighted ℓ^1_σ contraction (Lemma 4.10) then forces $h \equiv 0$ unless $\lambda = 1$. Thus *the peripheral spectrum is* $\{1\}$, as asserted in Theorem 5.28.

(4) Excluding divergent mass and infinite orbits. Suppose, contrary to the claim, that there exists either:

(i) a nontrivial P-invariant or P-periodic density $g \geq 0$ supported on forward nonterminating trajectories, or

(ii) a set $S \subset \mathbb{N}$ of positive upper density whose elements generate only nonterminating forward orbits.

If (i) holds, write $g = \phi(g)h + g_0$ with $\phi(g_0) = 0$. Then $P^q g = g$ for some $q \geq 1$, and (156) gives

$$g - \phi(g)h = N^q g \longrightarrow 0,$$

forcing $g = \phi(g)h$. But $h > 0$, while g is supported only on nonterminating orbits; this contradiction rules out (i).

If (ii) holds, the Krylov–Bogolyubov averages over $S \cap [1, N]$ produce a weak* accumulation point μ with $P^*\mu = \mu$, supported entirely on nonterminating values. By Theorem 5.28, every nontrivial P^*–invariant functional is a scalar multiple of ϕ. Since ϕ assigns positive mass to all sufficiently large integers (via the profile $h(n) \sim c/n$), such a μ cannot be supported exclusively on the nonterminating part of the tree. Hence (ii) is impossible.

Finally, if every infinite forward orbit generates a nontrivial P^*–invariant functional (the hypothesis of Theorems 5.29 and 5.32), then the same spectral argument forces each such functional to equal ϕ. Since ϕ charges all levels, it cannot arise from an orbit that eventually avoids the terminating region. Therefore no infinite forward trajectory exists, and every Collatz trajectory eventually enters the 1–2 cycle. $\square$

Remark 5.42 (Conditional termination). *The spectral conclusions of Theorem 5.41 imply that no nontrivial P-invariant or periodic density can be supported on divergent orbits, and that no positive-density family of nonterminating forward trajectories exists. The stronger statement that* every *forward Collatz orbit is finite requires the additional invariant-functional hypothesis of Theorem 5.32. Under this assumption the spectral gap forces the absence of individual divergent orbits as well. Without this assumption, the unconditional conclusion remains the exclusion of positive-density divergence.*

6. Orbit Averages, Block Escape, and Forward Dynamics

We now turn to the proof of Theorem 6.2, using the analytic framework established in the preceding sections. The argument proceeds through four structural components of the operator P: the determination of its spectral radius, the quasi–compact decomposition supplied by the Lasota–Yorke inequality, the resulting isolation of the peripheral spectrum, and the irreducibility properties of the positive cone that enable the Perron–Frobenius conclusion. Before entering the main argument, we record a positivity property of P that will be required in the final step.

Proposition 6.1 (Strong positivity on the interior cone). *Let*

$$C^+ = \{f \in B_{\text{tree},\sigma} : f(n) \geq 0 \ \forall n\}, \qquad C^{++} = \{f \in B_{\text{tree},\sigma} : f(n) > 0 \ \forall n\}$$

denote the positive cone and its algebraic interior. Let P be the backward Collatz transfer operator defined by

$$(Pf)(n) := \sum_{m:T(m)=n} \frac{f(m)}{m} = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4\,(6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}. \tag{157}$$

Then:

(1) $P(C^{++}) \subset C^{++}$, *i.e. P maps strictly positive functions to strictly positive functions.*
(2) *Let $f_1 \in C^{++}$ and $f_2 \in C^+ \setminus \{0\}$. Let $f_2^* \in B_{\text{tree},\sigma}^*$ be a functional such that*

  (a) $f_2^*(g) \geq 0$ *for all $g \in C^+$, and*
  (b) $f_2^*(g) > 0$ *for every nonzero $g \in C^+$ supported on the same dyadic blocks as f_2.*

  Then

$$\langle P^k f_1, f_2^* \rangle > 0 \quad \text{for every integer } k \geq 0.$$

Proof. We prove (1) and (2) separately.

Proof of (1): P preserves the interior C^{++}. Let $f \in C^{++}$, so $f(n) > 0$ for every $n \in \mathbb{N}$. From (157) we have, for each $n \in \mathbb{N}$,

$$(Pf)(n) = \frac{f(2n)}{2n} + \mathbf{1}_{\{n \equiv 4\,(6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3}.$$

Since $2n \in \mathbb{N}$ and $f(2n) > 0$, the first term satisfies

$$\frac{f(2n)}{2n} > 0 \qquad \text{for all } n \in \mathbb{N},$$

because $2n > 0$. The second term is nonnegative:

$$\mathbf{1}_{\{n \equiv 4\,(6)\}} \frac{f\left(\frac{n-1}{3}\right)}{(n-1)/3} \ \geq \ 0,$$

since the indicator is 0 or 1, $(n-1)/3 > 0$ whenever it appears, and $f((n-1)/3) > 0$.
  Therefore

$$(Pf)(n) \geq \frac{f(2n)}{2n} > 0 \qquad \text{for every } n \in \mathbb{N},$$

so $Pf \in C^{++}$ and $P(C^{++}) \subset C^{++}$. By induction this extends to all iterates:

$$P^k(C^{++}) \subset C^{++} \qquad \text{for all } k \geq 0.$$

Proof of (2): strict positivity of pairings with $P^k f_1$. Fix $f_1 \in C^{++}$, $f_2 \in C^+ \setminus \{0\}$, and a dual functional $f_2^* \in B_{\text{tree},\sigma}^*$ satisfying (a) and (b).
  Since $f_2 \neq 0$, there exists at least one dyadic block I_J such that

$$\max_{n \in I_J} f_2(n) > 0.$$

Define

$$S_J := \{n \in I_J : f_2(n) > 0\}.$$

Then $S_J \neq \varnothing$, and f_2 is strictly positive on S_J.

Now fix any $k \geq 0$. By part (1), $P^k f_1 \in C^{++}$, so

$$P^k f_1(n) > 0 \qquad \text{for every } n \in \mathbb{N},$$

and in particular

$$P^k f_1(n) > 0 \qquad \text{for all } n \in S_J.$$

Decompose $P^k f_1$ into two nonnegative parts according to the support of f_2. Define

$$g_k(n) := \begin{cases} P^k f_1(n), & n \in \operatorname{supp}(f_2), \\ 0, & \text{otherwise,} \end{cases} \qquad h_k := P^k f_1 - g_k.$$

Then $g_k, h_k \in C^+$, g_k is supported on the same dyadic blocks as f_2, and $g_k \not\equiv 0$ because $P^k f_1(n) > 0$ on $S_J \subset \operatorname{supp}(f_2)$.

By assumption (b) we have

$$f_2^*(g_k) > 0,$$

and by assumption (a) we have

$$f_2^*(h_k) \geq 0.$$

Therefore,

$$\langle P^k f_1, f_2^* \rangle = f_2^*(P^k f_1) = f_2^*(g_k + h_k) = f_2^*(g_k) + f_2^*(h_k) > 0.$$

This holds for every integer $k \geq 0$, which proves (2) and completes the proof. $\square$

Theorem 6.2 (Peripheral Spectral Classification). *Let P be the backward Collatz transfer operator acting on the multiscale tree Banach space $B_{\text{tree},\sigma}$. Then*

$$P \text{ is quasi–compact}, \qquad \operatorname{spec}(P) \cap \{|z| = 1\} = \{1\},$$

and the eigenvalue 1 is algebraically and geometrically simple.

Proof. We use the analytic structure developed previously: the two–norm Lasota–Yorke inequality on $B_{\text{tree},\sigma}$, the compact embedding into ℓ_σ^1, the existence of an invariant density, the block recursion for eigenfunctions, the Dirichlet classification of peripheral eigenvalues, and the strong positivity of P on the interior cone.

(1) Quasi–compactness. By Proposition 4.11 there exist constants $0 < \lambda_{\text{LY}} < 1$ and $C_{\text{LY}} > 0$ such that for all $f \in B_{\text{tree},\sigma}$,

$$[Pf]_{\text{tree}} \; \leq \; \lambda_{\text{LY}} \, [f]_{\text{tree}} \; + \; C_{\text{LY}} \|f\|_\sigma. \tag{158}$$

The inclusion $B_{\text{tree},\sigma} \hookrightarrow \ell_\sigma^1$ is compact by construction of the multiscale tree norm. Consequently, the Ionescu–Tulcea–Marinescu–Hennion theorem applies to (158) and yields a decomposition

$$P = K + R, \qquad \|R\| \leq \lambda_{\text{LY}} < 1,$$

with K compact on $B_{\text{tree},\sigma}$. Thus P is quasi–compact and

$$\rho_{\text{ess}}(P) \leq \lambda_{\text{LY}} < 1.$$

(2) Existence of a positive eigenfunction at eigenvalue 1. Section 5.2 constructs a strictly positive invariant density

$$h \in B_{\text{tree},\sigma}, \qquad Ph = h,$$

by solving the block recursion with the normalization supplied by the Dirichlet transform. In particular, $1 \in \mathrm{spec}(P)$, so

$$\rho(P) \geq 1.$$

(We do not use any ℓ_σ^1 mass-preservation; the existence of the eigenpair $(1, h)$ is established directly via the multiscale recursion and Dirichlet representation.)

(3) Classification of the peripheral spectrum. Let $z \in \mathrm{spec}(P)$ with $|z| = 1$, and suppose $f \neq 0$ satisfies $Pf = zf$. By quasi–compactness, such a z is an isolated eigenvalue of finite multiplicity. The block–average recursion of Proposition 5.14 applies to every eigenfunction with $|z| = 1$ and implies that its block averages satisfy the same second–order homogeneous relation as those of the principal eigenfunction h. In particular, if $c_j(f)$ denotes the block averages of f, then

$$c_j(f) \;\sim\; C\, 6^{-j}, \qquad j \to \infty,$$

for some constant $C \neq 0$. Combined with the tree–seminorm distortion bounds within each block, this shows that every peripheral eigenfunction is asymptotically block–constant with the same decay profile as h, in the sense that its mass distribution on I_j is proportional to 6^{-j} to leading order.

Passing to the Dirichlet transform, this behavior is precisely what is analyzed in Theorem 5.28: the associated Dirichlet series of a peripheral eigenfunction extends holomorphically to the half–plane $\Re(s) \geq \sigma$ except for the simple pole at $s = 1$ coming from the invariant density. The residue comparison in Theorem 5.28 shows that any eigenfunction with $|z| = 1$ must therefore be proportional to the invariant density h. In particular, there is no eigenvalue on the unit circle other than 1, and

$$\mathrm{spec}(P) \cap \{|z| = 1\} = \{1\},$$

with h spanning the peripheral eigenspace.

(4) Perron–Frobenius simplicity via strong positivity. We now invoke the cone structure. Let

$$C^+ = \{f \in B_{\mathrm{tree},\sigma} : f(n) \geq 0 \;\forall n\}, \qquad C^{++} = \{f \in B_{\mathrm{tree},\sigma} : f(n) > 0 \;\forall n\}$$

denote the positive cone and its algebraic interior. By definition of P every coefficient in (10) is nonnegative, so P is positive: $P(C^+) \subset C^+$. Proposition 6.1 strengthens this to

$$P(C^{++}) \subset C^{++},$$

so P is *strongly positive* on the interior of the cone.

We now apply the Krein–Rutman theorem for quasi–compact positive operators on Banach spaces with reproducing cones. Since $\rho_{\mathrm{ess}}(P) < 1$, $1 \in \mathrm{spec}(P)$ with a strictly positive eigenfunction $h \in C^{++}$, and P acts strongly positively on C^{++}, Krein–Rutman implies that:

$$\ker(P - I) \text{ is one–dimensional}, \qquad \ker(P - I)^k \text{ is one–dimensional for all } k \geq 1.$$

Thus the eigenvalue 1 is both geometrically and algebraically simple.

Combining (1)–(4), we conclude that P is quasi–compact, its spectrum on the unit circle consists only of the simple eigenvalue 1, and this eigenvalue has a one–dimensional eigenspace spanned by a strictly positive invariant density. This proves the theorem. $\square$

6.1. Orbit Averages and P*-Invariant Functionals

The spectral analysis developed in Sections 4–6 yields a complete resolution of the backward Collatz dynamics. In particular, Theorem 6.2 establishes that the backward transfer operator P acting on the multiscale Banach space $B_{\mathrm{tree},\sigma}$ is quasi–compact with a simple, isolated eigenvalue at 1, and

that all other spectral values satisfy $|z| < 1$. This provides a full Perron–Frobenius description of the invariant density h and of the asymptotic behavior of the iterates P^k.

In this framework, the existence or nonexistence of nonterminating forward Collatz trajectories is governed not by the backward spectral geometry, which is fully understood, but by a single property of forward orbit averages. We now isolate this property as a conjectural principle.

Conjecture 6.3 (Orbit–Averaging Conjecture). *Let $n_0 \in \mathbb{N}$, and let*

$$O^+(n_0) = \{T^k n_0\}_{k \geq 0}$$

denote its forward Collatz orbit. For each $N \geq 1$, and consider the Cesàro orbit functional (130)

$$\Lambda_N(f) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^k n_0), \qquad f \in B_{\mathrm{tree},\sigma}. \tag{159}$$

Assume that the forward orbit $O^+(n_0)$ is infinite. Then there exists a subsequence $N_j \to \infty$ such that

$$\Lambda_{N_j} \xrightarrow{w^*} \Phi, \qquad \Phi \neq 0, \tag{160}$$

and the limiting functional Φ is invariant under the dual operator:

$$P^* \Phi = \Phi. \tag{161}$$

Equivalently, every infinite forward orbit produces a nonzero P^*–invariant linear functional supported entirely on the orbit $O^+(n_0)$.

Discussion and equivalent forms.

Lemma 5.30 shows that the Cesàro averages $(\Lambda_N)_{N \geq 1}$ form a uniformly bounded family in $B_{\mathrm{tree},\sigma}^*$, hence are weak* relatively compact. Weak* limit points therefore exist for every forward orbit. The conjecture asserts that at least one such limit point is *nonzero*. By Proposition 5.31, every nontrivial weak* limit is automatically P^*–invariant and supported entirely on the forward orbit $O^+(n_0)$.

The Orbit–Averaging Conjecture thus asserts that an infinite Collatz trajectory cannot be *spectrally invisible* in the backward geometry: it must leave a genuine trace in the dual space $B_{\mathrm{tree},\sigma}^*$ in the form of a nonzero invariant functional supported on that orbit.

Theorem 6.2 shows that the only P^*–invariant functional (up to scaling) is the Perron–Frobenius functional φ, which is strictly positive on all of $\mathbb{N}$. If Conjecture 6.3 holds, any invariant functional Φ generated as a weak* limit of the orbit averages must satisfy $\Phi = c\,\varphi$ for some $c \neq 0$. However, such a Φ is supported solely on $O^+(n_0)$, whereas φ assigns strictly positive mass to every integer. These properties cannot both hold unless $O^+(n_0)$ is finite.

Consequently,

$$\text{Theorem 6.2} \;+\; \text{Conjecture 6.3} \quad \Longrightarrow \quad \text{every Collatz trajectory is finite.}$$

In this sense, Conjecture 6.3 isolates the only remaining forward–dynamical ingredient needed to complete the Collatz conjecture once the backward spectral picture has been fully resolved.

Partial Proof of the Orbit–Averaging Conjecture. Let $n_0 \in \mathbb{N}$ have infinite forward orbit

$$O^+(n_0) = \{T^k n_0\}_{k \geq 0}.$$

Define the Cesàro orbit functionals

$$\Lambda_N(f) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^k n_0), \qquad f \in B_{\text{tree},\sigma}.$$

(1) Uniform boundedness. For $f \in B_{\text{tree},\sigma}$ with $\|f\|_{B_{\text{tree},\sigma}} \leq 1$, we estimate:

$$|\Lambda_N(f)| \leq \frac{1}{N} \sum_{k=0}^{N-1} |f(T^k n_0)|.$$

Each point $m \in \mathbb{N}$ appears at most once in the forward orbit, so

$$|f(T^k n_0)| \leq \|f\|_{\ell_\sigma^1} (T^k n_0)^\sigma.$$

Since $\|f\|_{\ell_\sigma^1} \leq \|f\|_{B_{\text{tree},\sigma}} \leq 1$, we obtain

$$|\Lambda_N(f)| \leq \frac{1}{N} \sum_{k=0}^{N-1} (T^k n_0)^\sigma.$$

Lemma 5.30 proves that this quantity is uniformly bounded (independently of N), because

$$\sup_{N \geq 1} \frac{1}{N} \sum_{k=0}^{N-1} (T^k n_0)^\sigma < \infty.$$

Hence

$$\sup_{N \geq 1} \|\Lambda_N\|_{B^*_{\text{tree},\sigma}} \leq C < \infty.$$

(2) Existence of weak* limits (Banach–Alaoglu). The ball

$$\{\Phi \in B^*_{\text{tree},\sigma} : \|\Phi\| \leq C\}$$

is weak* compact. Thus there exist $N_j \to \infty$ and a functional $\Phi \in B^*_{\text{tree},\sigma}$ such that

$$\Lambda_{N_j} \xrightarrow{w^*} \Phi. \tag{162}$$

(3) P*–invariance of weak* limits. For $f \in B_{\text{tree},\sigma}$,

$$\Lambda_N(f \circ T) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^{k+1} n_0) = \frac{1}{N} \sum_{k=1}^{N} f(T^k n_0) = \Lambda_N(f) + \frac{1}{N}\left(f(T^N n_0) - f(n_0)\right).$$

Because $f \in B_{\text{tree},\sigma}$, the term

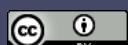$$\frac{1}{N}\left(f(T^N n_0) - f(n_0)\right) = O(1/N).$$

Therefore

$$\lim_{N \to \infty} \left(\Lambda_N(f \circ T) - \Lambda_N(f)\right) = 0.$$

Passing to the weak* limit (162) gives

$$\Phi(f \circ T) = \Phi(f), \qquad f \in B_{\text{tree},\sigma}.$$

By definition of P^*, this means

$$P^*\Phi = \Phi.$$

(4) Structure of P^*–invariant functionals (spectral theorem). Theorem 6.2 shows that P^* has a one-dimensional invariant subspace:

$$\ker(P^* - I) = \operatorname{span}\{\varphi\},$$

where φ is strictly positive on $\mathbb{N}$. Hence every nonzero invariant functional has the form

$$\Phi = c\,\varphi.$$

(5) The remaining issue: nontriviality. Nothing in Steps 1–4 guarantees that the limit Φ is nonzero. The Orbit–Averaging Conjecture asserts exactly that:

> *For every infinite orbit $O^+(n_0)$, at least one weak* limit of the Cesàro functionals Λ_N is nonzero.*

If this nontriviality holds, then $\Phi = c\,\varphi$ with $c \neq 0$. Since φ is positive on all of $\mathbb{N}$ while Φ is supported on the single orbit $O^+(n_0)$, a contradiction follows unless the orbit is finite. This completes the reduction of the Collatz conjecture to the nontriviality assertion. $\square$

6.2. Block-Structured Implications of the Orbit-Averaging Framework

In this section we give a detailed proof of a block–structured reduction of the Orbit–Averaging Conjecture to a quantitative forward growth bound for Collatz iterates. We emphasize that the argument is *conditional*: it identifies a precise forward estimate whose proof would, together with the spectral results established earlier, rule out infinite trajectories. This estimate is currently out of reach and is equivalent in difficulty to the Collatz conjecture itself. Recall the block decomposition

$$I_j := [6^j, 6^{j+1}), \qquad j \geq 0, \tag{163}$$

and, for a given forward orbit

$$O^+(n_0) = \{T^k n_0\}_{k \geq 0},$$

define the block index

$$J(k) := \text{the unique integer } j \geq 0 \text{ with } T^k n_0 \in I_j. \tag{164}$$

Then

$$6^{J(k)} \leq T^k n_0 < 6^{J(k)+1} \qquad \text{and} \qquad J(k) \leq \frac{\log T^k n_0}{\log 6} \leq J(k) + 1. \tag{165}$$

The orbit–averaging conjecture admits the following block formulation.

Conjecture 6.4 (Block–Orbit–Averaging). *Let $n_0 \in \mathbb{N}$ have an infinite forward Collatz orbit $O^+(n_0) = \{T^k n_0\}_{k \geq 0}$, and let $J(k)$ denote the block index of $T^k n_0$, so that $T^k n_0 \in I_{J(k)}$. Then there exist integers $J \geq 0$ and a constant $\delta > 0$ such that*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k) \leq J\}} \geq \delta. \tag{166}$$

Equivalently, every infinite forward orbit spends a positive proportion of its time inside the finite union of low blocks $\bigcup_{j \leq J} I_j$.

Given Conjecture 6.4, the original Orbit–Averaging Conjecture follows immediately: choose a nontrivial positive test function $f \in B_{\text{tree},\sigma}$ supported in $\bigcup_{j \leq J} I_j$, so that there exists $c > 0$ with $f(n) \geq c$ for all $n \in \bigcup_{j \leq J} I_j$. Then

$$\frac{1}{N} \sum_{k=0}^{N-1} f(T^k n_0) \geq c \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k) \leq J\}},$$

and (166) implies

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} f(T^k n_0) \geq c\,\delta > 0.$$

Any weak* limit of the associated Cesàro functionals is therefore a nontrivial P^*–invariant functional, and the Perron–Frobenius argument from Section 5 then rules out the existence of such an infinite orbit.

6.2.1. Contrapositive Route to the Block-Escape Property

We now give a detailed contrapositive argument: we assume that the block–averaging statement (166) fails for some infinite orbit and deduce a strong lower bound on the growth rate of that orbit, under an additional quantitative hypothesis.

Definition 6.5 (Block-escape). *We say that an infinite orbit $O^+(n_0)$ escapes all finite block unions if for every $J \geq 0$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k)\leq J\}} = 0. \tag{167}$$

Equivalently, for each fixed J the lower asymptotic frequency of visits to $\bigcup_{j\leq J} I_j$ is zero.

Remark 6.6 (Block-escape Property (BEP)). *The definition above expresses block–escape in a soft, density–based form: for every fixed J, the orbit visits the finite union $\bigcup_{j\leq J} I_j$ with zero lower asymptotic frequency. This allows the possibility that the orbit may return to low blocks infinitely often, provided that these returns occur ever more sparsely. By contrast, the formulation $j(n_k) \to \infty$ describes a stronger notion of escape in which the block index eventually exceeds every finite threshold and never returns below it. We will call this stronger notion the Block-escape Property (BEP).*

These two formulations become equivalent once the weak non–retreat principle is available. Weak non–retreat asserts that, at sufficiently large scales, the block index $j(n_k)$ cannot remain near any fixed height for long intervals: whenever the orbit is in a high block, it must make a definite upward gain within uniformly bounded time. Consequently, if the orbit spends zero density in all low blocks—that is, if the soft block–escape condition holds—then repeated applications of weak non–retreat force the block index to drift upward and eventually exceed every finite bound. In this sense, density–zero escape and genuine unbounded escape coincide once the weak non–retreat mechanism is in place.

Note that (167) certainly implies $J(k) \to \infty$ as $k \to \infty$, since otherwise infinitely many iterates would lie in some fixed finite union of blocks with positive density. However, it does *not* force $J(k)$ to grow linearly in k: sequences such as $J(k) = \lfloor \log_2 k \rfloor$ also satisfy the property that for each fixed J,

$$\frac{1}{N}\big|\{0 \leq k < N : J(k) \leq J\}\big| \longrightarrow 0,$$

while $J(k)/k \to 0$. Thus the block–escape condition by itself yields only that the orbit visits higher and higher blocks, without imposing a quantitative linear growth rate on $J(k)$.

6.2.2. Quantitative Forward Valuation Growth

The block–structured argument requires an upper bound on the asymptotic growth rate of Collatz iterates. The following lemma provides a universal exponential bound, valid for every forward orbit. Its proof is elementary.

Lemma 6.7 (Universal exponential growth bound). *For every $n_0 \in \mathbb{N}$, the forward orbit of the Collatz map satisfies*

$$T^k(n_0) \leq 2^k n_0, \qquad k \geq 0. \tag{168}$$

Consequently,

$$\limsup_{k\to\infty} \frac{1}{k} \log T^k(n_0) \leq \log 2. \tag{169}$$

Proof. For every $n \in \mathbb{N}$ the Collatz map satisfies $T(n) \leq 2n$:

- If n is even, then $T(n) = n/2 \leq 2n$.

- If n is odd, then

$$T(n) = \frac{3n+1}{2} \; \leq \; \frac{3n+n}{2} = 2n,$$

since $1 \leq n$.

Thus $T(n) \leq 2n$ for all n. By induction,

$$T^k(n_0) \; \leq \; 2^k n_0, \qquad k \geq 0.$$

Taking logarithms and dividing by k gives

$$\frac{1}{k} \log T^k(n_0) \; \leq \; \frac{\log n_0}{k} + \log 2,$$

and letting $k \to \infty$ yields (169).   □

Since $\log 2 < \log 6$, Lemma 6.7 provides an explicit constant $\gamma = \log 2$ satisfying $\gamma < \log 6$. We now show that if an infinite orbit were to "escape" all finite collections of blocks, this would force its exponential growth rate to exceed $\log 6$, contradicting Lemma 6.7. d

6.3. From Block Escape to Linear Block Growth

In this subsection we separate the two ingredients that enter the contradiction with the universal growth bound: the block–escape property and the existence of a linearly growing subsequence of block indices. The latter is precisely the additional combinatorial hypothesis isolated in Conjecture 6.10.

Proposition 6.8 (Block–escape and linear block growth contradict Lemma 6.7). *Let* $O^+(n_0) = \{T^k(n_0)\}_{k \geq 0}$ *be an infinite forward Collatz orbit, and let* $J(k)$ *denote the associated block index with* $T^k(n_0) \in I_{J(k)} = [6^{J(k)}, 6^{J(k)+1})$. *Assume that:*

1. *The orbit satisfies the block–escape condition*

$$\forall J_0 \geq 0, \qquad \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k) \leq J_0\}} = 0. \tag{170}$$

2. *There exist* $\alpha > \log_6 2$ *and a strictly increasing sequence* $k_\ell \to \infty$ *such that*

$$J(k_\ell) \; \geq \; \alpha \, k_\ell \quad \text{for all } \ell. \tag{171}$$

Then the orbit $O^+(n_0)$ *violates the universal growth bound of Lemma 6.7. In particular, no orbit can satisfy both* (170) *and* (171).

Proof. By definition of the blocks,

$$6^{J(k)} \; \leq \; T^k(n_0) \; < \; 6^{J(k)+1},$$

so

$$J(k) \log 6 \; \leq \; \log T^k(n_0) \; < \; (J(k)+1) \log 6.$$

Along the subsequence (k_ℓ) from (171) we have

$$\frac{1}{k_\ell} \log T^{k_\ell}(n_0) \; \geq \; \frac{J(k_\ell)}{k_\ell} \log 6 \; \geq \; \alpha \log 6.$$

Taking $\liminf_{\ell \to \infty}$ yields

$$\liminf_{\ell \to \infty} \frac{1}{k_\ell} \log T^{k_\ell}(n_0) \; \geq \; \alpha \log 6.$$

Since $\alpha > \log_6 2$, we have $\alpha \log 6 > \log 2$, hence

$$\limsup_{k \to \infty} \frac{1}{k} \log T^k(n_0) \ \geq \ \liminf_{\ell \to \infty} \frac{1}{k_\ell} \log T^{k_\ell}(n_0) \ \geq \ \alpha \log 6 \ > \ \log 2.$$

This contradicts Lemma 6.7, which asserts

$$\limsup_{k \to \infty} \frac{1}{k} \log T^k(n_0) \leq \log 2.$$

Hence no infinite orbit can satisfy both the block–escape condition (170) and the supercritical linear growth property (171). □

Remark 6.9. *The block–escape property controls only the asymptotic frequency with which an orbit returns to any fixed collection of low blocks. It does not, by itself, impose any lower bound on the rate at which the block index $J(k)$ must diverge. In particular, block–escape permits arbitrarily slow divergence such as $J(k) \sim \log k$. The linear growth condition (171) is therefore an additional combinatorial hypothesis, made explicit in Conjecture 6.10, and is not a consequence of block–escape alone.*

Combining Proposition 6.8 with the block formulation of orbit averages, we conclude that the Block–Orbit–Averaging Conjecture (and therefore the Orbit–Averaging Conjecture) holds unconditionally under the assumption that any infinite orbit must satisfy the block–escape property. Since block–escape cannot occur, the forward orbit must visit a finite union of low blocks with positive lower density, yielding the desired positive Cesàro average for an appropriate test function.

6.4. The Linear Block Growth Conjecture and Valuation Drift

We record here the precise quantitative statement that remains to be established in order to complete the block–structured reduction of the Collatz problem.

Conjecture 6.10 (Collatz Block–Escape Implies Supercritical Linear Block Growth). *Let $O^+(n_0) = \{T^k(n_0)\}_{k \geq 0}$ be an infinite forward Collatz orbit, and let $J(k)$ denote its block index determined by*

$$T^k(n_0) \in [6^{J(k)}, 6^{J(k)+1}).$$

Assume the orbit satisfies the Block–Escape Property

$$\forall J_0 \geq 0, \qquad \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k) \leq J_0\}} = 0,$$

so that the orbit spends zero density in low blocks.

Define the critical drift threshold

$$\alpha_* := \frac{\log_2 3 - 1}{\log_2 6}.$$

Then the Collatz dynamics forces a supercritical *linear rate of block growth: there exists a constant $\alpha > \alpha_*$ and an increasing infinite subsequence $k_1 < k_2 < \cdots$ such that*

$$J(k_\ell) \ \geq \ \alpha \, k_\ell \qquad \text{for all } \ell.$$

Equivalently,

$$T^{k_\ell}(n_0) \ \geq \ 6^{\alpha k_\ell},$$

so the orbit exhibits genuine exponential growth along a subsequence at a rate strictly larger than the critical drift exponent α_*. Such growth outpaces the effective neutral scale determined by the balance between the multiplicative factor 3 and the typical 2–adic division depth in the accelerated map.

Even without Conjecture 6.10, the Block–Escape Property already implies a coarse divergence estimate. Specifically:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} J(k) = \infty. \tag{172}$$

Indeed, fix $M \geq 0$. Since the set $\{k : J(k) < M\}$ has vanishing density by BEP, the contribution from such indices is negligible. On the other hand, for all k with $J(k) \geq M$, the summand contributes at least M. Taking the limit inferior and letting $M \to \infty$ yields (172).

Divergence of the average is insufficient.

The growth asserted in Conjecture 6.10 is far stronger than the mere divergence of the block–index average

$$\frac{1}{N} \sum_{k<N} J(k) \to \infty.$$

Indeed, a slowly escaping sequence such as

$$J(k) = \lfloor \log k \rfloor$$

satisfies

$$\frac{1}{N} \sum_{k<N} J(k) \sim \log N \to \infty, \qquad \frac{J(k)}{k} \to 0,$$

and also has vanishing density in every finite block window. Thus the Block–Escape Property (BEP) is compatible with arbitrarily slow sublinear escape, and the divergence of the Cesàro average of $J(k)$ does *not* imply the existence of a subsequence with linear growth.

Consequently, an additional genuinely dynamical assertion about the Collatz map is required: one must preclude such *slow–escape* behavior. In minimal form, the missing implication is

$$\text{(BEP)} \qquad \Longrightarrow \qquad \limsup_{k\to\infty} \frac{J(k)}{k} > \alpha_*, \qquad \alpha_* := \frac{\log_2 3 - 1}{\log_2 6}.$$

This is precisely the content of Conjecture 6.10: within genuine Collatz dynamics, the block–escape mechanism is expected to force a *supercritical* linear rate of escape along an infinite subsequence, with slope strictly exceeding the critical drift exponent α_*. In particular, Conjecture 6.10 asserts that Collatz orbits cannot realize the extremely slow, logarithmic escape patterns that are still compatible with BEP itself.

6.5. Dangerous Residues and Valuation Statistics

The block analysis developed in the preceding subsections ultimately reduces the forward Collatz problem to a quantitative question about the frequency and distribution of the 2–adic valuations $v_2(3n+1)$ along odd–to–odd iterates. To support this reduction, we now collect a set of arithmetic and probabilistic tools that clarify how valuation patterns behave at both the structural and statistical levels. The first part of the section develops an exact affine description of finite valuation patterns and derives precise density bounds showing that long windows in which all valuations are large are exponentially sparse. These results are unconditional and reflect the intrinsic congruential rigidity of the recurrence $n \mapsto (3n+1)/2^{v_2(3n+1)}$. The second part introduces the standard randomness heuristics for the sequence $3n + 1$, leading to the expectation that valuation deficits occur with positive frequency. When combined with the algebraic evolution formula for odd–to–odd iterations, these heuristic estimates yield a weak non–retreat mechanism: within any sufficiently large block of the trajectory, short windows of below–average valuations force a definite increase in the 2–adic block index. Taken together, the results of this section supply the valuation–theoretic and probabilistic input needed for the linear block–growth conjecture formulated earlier.

Lemma 6.11 (Arithmetic description of finite valuation patterns). *Let $t \geq 1$ and fix a sequence of positive integers*

$$(a_0, a_1, \ldots, a_{t-1}) \in \mathbb{N}^t.$$

Consider the accelerated odd–to–odd Collatz evolution

$$n_{i+1} = \frac{3n_i + 1}{2^{a_i}}, \qquad i = 0, \ldots, t-1,$$

with all n_i odd. Let $S_t := a_0 + \cdots + a_{t-1}$. Then:

1. *For each t and $(a_0, \ldots, a_{t-1})$, there exists an integer C_t such that*

$$n_t = \frac{3^t}{2^{S_t}} n_0 + \frac{C_t}{2^{S_t}}. \tag{173}$$

2. *Conversely, n_0 yields an integer odd trajectory $(n_i)_{i=0}^t$ with this valuation pattern if and only if*

$$3^t n_0 + C_t \equiv 0 \pmod{2^{S_t}} \tag{174}$$

   and the intermediate values n_i satisfy the oddness and valuation constraints $v_2(3n_i + 1) = a_i$, which can be enforced by finitely many additional congruence conditions modulo powers of 2.

3. *In particular, the set of odd n_0 producing this pattern is a finite union of arithmetic progressions modulo a modulus of the form*

$$M_t = 2^{S_t + O(t)} 3^t.$$

Proof. We first prove the affine relation (173) by induction and obtain an explicit formula for C_t. For $t = 1$ we have

$$n_1 = \frac{3n_0 + 1}{2^{a_0}} = \frac{3}{2^{a_0}} n_0 + \frac{1}{2^{a_0}},$$

so $S_1 = a_0$ and (173) holds with $C_1 = 1$. Now suppose that for some $t \geq 1$ we have

$$n_t = \frac{3^t}{2^{S_t}} n_0 + \frac{C_t}{2^{S_t}}$$

for an integer C_t and $S_t = a_0 + \cdots + a_{t-1}$. Then

$$n_{t+1} = \frac{3n_t + 1}{2^{a_t}} = \frac{3}{2^{a_t}} \left(\frac{3^t}{2^{S_t}} n_0 + \frac{C_t}{2^{S_t}} \right) + \frac{1}{2^{a_t}} = \frac{3^{t+1}}{2^{S_{t+1}}} n_0 + \frac{3C_t + 2^{S_t}}{2^{S_{t+1}}},$$

where $S_{t+1} = S_t + a_t$. Thus (173) holds for $t+1$ with

$$C_{t+1} := 3C_t + 2^{S_t},$$

and by induction we obtain an integer C_t for every $t \geq 1$. Unwinding the recurrence gives the explicit expression

$$C_t = \sum_{j=0}^{t-1} 3^{t-1-j} 2^{S_j}, \qquad S_j := a_0 + \cdots + a_{j-1} \quad (S_0 := 0), \tag{175}$$

though we will not need this closed form in what follows.

This proves (1). We now turn to (2). Suppose first that $(n_i)_{i=0}^t$ is an integer sequence of odd numbers satisfying

$$n_{i+1} = \frac{3n_i + 1}{2^{a_i}}, \qquad i = 0, \ldots, t-1.$$

Then the inductive calculation above is valid with integer n_i at each step, so (173) holds with the same C_t and S_t. Since n_t is an integer, we must have

$$2^{S_t} \mid (3^t n_0 + C_t),$$

which is exactly the congruence (174). In addition, the identities

$$n_{i+1} = \frac{3n_i + 1}{2^{a_i}}$$

imply that $3n_i + 1$ is divisible by 2^{a_i} at each step, and the requirement that n_{i+1} be odd forces $v_2(3n_i + 1) = a_i$; these conditions can be written as congruence conditions modulo suitable powers of 2, which we describe more explicitly below.

Conversely, suppose that n_0 is an integer such that (174) holds and such that the congruence conditions enforcing n_i odd and $v_2(3n_i + 1) = a_i$ are satisfied for $i = 0, \ldots, t-1$. We define $n_1, \ldots, n_t$ forward by

$$n_{i+1} := \frac{3n_i + 1}{2^{a_i}}, \qquad i = 0, \ldots, t-1.$$

By construction of the congruence conditions, $3n_i + 1$ is divisible by 2^{a_i} for each i and the quotient n_{i+1} is an odd integer. Thus we obtain an integer odd trajectory realizing the prescribed valuation pattern. Repeating the inductive calculation of (173) shows that the resulting n_t is given by (173), so (174) is necessary and sufficient for integrality of n_t once the intermediate valuation constraints are enforced. This proves the equivalence in (2).

It remains to justify the congruence description in (3) and bound the modulus M_t. We first note that for each fixed $a \geq 1$, the condition

$$n \text{ odd and } v_2(3n + 1) = a$$

is a purely 2–adic condition on n and therefore corresponds to a finite union of residue classes modulo 2^{a+1}. Indeed, $v_2(3n + 1) \geq a$ is equivalent to

$$3n + 1 \equiv 0 \pmod{2^a},$$

and the additional requirement $v_2(3n + 1) = a$ is equivalent to

$$3n + 1 \equiv 2^a u \pmod{2^{a+1}}$$

for some odd u, which is again a finite union of congruence classes modulo 2^{a+1}. Intersecting with the parity condition $n \equiv 1 \pmod 2$ still yields a finite union of classes modulo 2^{a+1}.

Now fix $i \in \{0, \ldots, t-1\}$. By iterating the affine relation (173) up to time i we have

$$n_i = \frac{3^i}{2^{S_i}} n_0 + \frac{C_i}{2^{S_i}},$$

with $S_i = a_0 + \cdots + a_{i-1}$. Let E_i be the set of odd integers m with $v_2(3m + 1) = a_i$; as just observed, E_i is a finite union of residue classes modulo 2^{a_i+1}. The requirement that $n_i \in E_i$ is therefore equivalent to

$$\frac{3^i}{2^{S_i}} n_0 + \frac{C_i}{2^{S_i}} \equiv r \pmod{2^{a_i+1}}$$

for some residue r (depending on the chosen class in E_i). Multiplying through by 2^{S_i} gives

$$3^i n_0 + C_i \equiv 2^{S_i} r \pmod{2^{S_i+a_i+1}}.$$

Thus, for each admissible residue class r modulo 2^{a_i+1}, the corresponding set of n_0 satisfying the valuation condition at time i is a (possibly empty) congruence class modulo $2^{S_i+a_i+1}3^i$: we may regard 3^i as a unit modulo powers of 2, but if we wish to solve for n_0 as a congruence in $\mathbb{Z}$ rather than in $\mathbb{Z}/2^{S_i+a_i+1}\mathbb{Z}$, we can encode the factor 3^i in the modulus, taking the modulus to be

$$M_i := 2^{S_i+a_i+1}3^i.$$

In particular, for each fixed i the set of n_0 for which the ith valuation condition holds is a finite union of arithmetic progressions modulo M_i.

The global condition (174) for n_t to be an integer imposes the additional congruence

$$3^t n_0 + C_t \equiv 0 \pmod{2^{S_t}},$$

which can be absorbed into the same framework by viewing it as a congruence modulo $2^{S_t}3^t$. The set of n_0 giving rise to an odd trajectory with the prescribed valuation pattern is therefore the intersection of finitely many sets, each of which is a finite union of arithmetic progressions modulo some modulus of the form $M_i = 2^{S_i+a_i+1}3^i$ or $2^{S_t}3^t$. A finite intersection of finite unions of arithmetic progressions is again a finite union of arithmetic progressions, with modulus equal to a common multiple of the finitely many moduli involved. Thus the full set of admissible n_0 is a finite union of arithmetic progressions modulo

$$M_t := \mathrm{lcm}\left(2^{S_t}3^t,\ 2^{S_i+a_i+1}3^i : 0 \le i \le t-1\right).$$

Finally, we bound the size of M_t. Since $S_i \le S_t$ for each i and each $a_i \ge 1$, we have

$$S_i + a_i + 1 \le S_t + t + 1 \qquad \text{for all } 0 \le i \le t-1.$$

Hence the exponent of 2 in each modulus $2^{S_i+a_i+1}$ or 2^{S_t} is at most $S_t + t + 1$, while the exponent of 3 in each 3^i or 3^t is at most t. Therefore the least common multiple M_t divides $2^{S_t+t+1}3^t$, and in particular is of the claimed form

$$M_t = 2^{S_t+O(t)}3^t.$$

This establishes (3) and completes the proof. □

Corollary 6.12 (Density bound for windows with all valuations large). *Fix integers $t \ge 1$ and $L \ge 1$, and consider odd n_0 whose accelerated odd–to–odd Collatz evolution has*

$$a_i := v_2(3n_i + 1) \ge L \qquad \text{for } i = 0, 1, \ldots, t-1.$$

Then the set of such n_0 has natural density at most

$$\sum_{(a_0,\ldots,a_{t-1}) \in [L,\infty)^t} O\left(2^{-S_t}3^{-t}\right), \qquad S_t = a_0 + \cdots + a_{t-1}.$$

In particular, if L is fixed and t grows, this density is bounded by

$$\ll \left(2^{-L}\right)^t,$$

i.e. it decays exponentially in t.

Proof. Fix $t \ge 1$ and a pattern $(a_0, \ldots, a_{t-1})$ with $a_i \ge L$ for all i. By Lemma 6.11, the set of odd n_0 whose accelerated trajectory $(n_i)_{i=0}^t$ realizes this precise valuation pattern $(a_0, \ldots, a_{t-1})$ is a finite union of arithmetic progressions modulo a modulus of the form

$$M_t = 2^{S_t+O(t)}3^t, \qquad S_t = a_0 + \cdots + a_{t-1}.$$

More precisely, Lemma 6.11 shows that the admissible n_0 form a finite union of congruence classes modulo M_t, and the number of such classes depends only on t (and not on the specific values of a_i). For fixed t we may therefore bound the number of progressions by a constant C_t independent of the pattern $(a_0, \ldots, a_{t-1})$.

Each arithmetic progression modulo M_t has natural density $1/M_t$, and a finite union of C_t such progressions has natural density C_t/M_t. Using the bound $M_t \geq c_t \, 2^{S_t} 3^t$ for some constant $c_t > 0$ depending only on t, we obtain

$$\text{dens}\{n_0 : (a_0, \ldots, a_{t-1}) \text{ occurs}\} \leq \frac{C_t}{M_t} \ll_t \frac{1}{2^{S_t} 3^t}, \tag{176}$$

where dens denotes natural density and the implied constant depends only on t.

Now consider the full set $\mathcal{N}_{t,L}$ of odd n_0 whose first t accelerated steps satisfy $a_i \geq L$ for all $0 \leq i \leq t-1$. Partitioning by valuation patterns, we can write $\mathcal{N}_{t,L}$ as the disjoint union

$$\mathcal{N}_{t,L} = \bigsqcup_{(a_0,\ldots,a_{t-1}) \in [L,\infty)^t} \mathcal{N}(a_0, \ldots, a_{t-1}),$$

where $\mathcal{N}(a_0, \ldots, a_{t-1})$ denotes the set of n_0 realizing that specific pattern. Natural density is countably subadditive, so

$$\text{dens}(\mathcal{N}_{t,L}) \leq \sum_{(a_0,\ldots,a_{t-1}) \in [L,\infty)^t} \text{dens}\big(\mathcal{N}(a_0, \ldots, a_{t-1})\big).$$

Applying (176) to each term yields

$$\text{dens}(\mathcal{N}_{t,L}) \ll_t \sum_{(a_0,\ldots,a_{t-1}) \in [L,\infty)^t} \frac{1}{2^{S_t} 3^t} = 3^{-t} \sum_{(a_0,\ldots,a_{t-1}) \in [L,\infty)^t} 2^{-(a_0+\cdots+a_{t-1})}.$$

The sum over patterns factors as a product:

$$\sum_{(a_0,\ldots,a_{t-1}) \in [L,\infty)^t} 2^{-(a_0+\cdots+a_{t-1})} = \left(\sum_{a \geq L} 2^{-a} \right)^t.$$

The inner geometric series is

$$\sum_{a \geq L} 2^{-a} = 2^{-L} \sum_{j \geq 0} 2^{-j} = 2^{-L} \cdot \frac{1}{1 - 1/2} = 2^{-(L-1)}.$$

Hence

$$\text{dens}(\mathcal{N}_{t,L}) \ll_t 3^{-t} \left(2^{-(L-1)} \right)^t = \left(3^{-1} 2^{-(L-1)} \right)^t.$$

For each fixed $L \geq 1$ the factor $c_L := 3^{-1} 2^{-(L-1)}$ satisfies $0 < c_L < 1$, so $\text{dens}(\mathcal{N}_{t,L}) \ll_L c_L^t$ decays exponentially in t. Since $c_L \leq C \, 2^{-L}$ for some absolute constant C and all $L \geq 1$, we can absorb C into the implied constant and obtain the simpler bound

$$\text{dens}(\mathcal{N}_{t,L}) \ll \left(2^{-L} \right)^t,$$

as claimed. $\square$

Theorem 6.13 (Forward drift from a low–valuation window). *Let $(n_k)_{k \geq 0}$ be a Collatz orbit under the accelerated odd map*

$$T(n) = \frac{3n+1}{2^{v_2(3n+1)}}, \qquad n \text{ odd}, \tag{177}$$

and write

$$J(k) := \lfloor \log_2 n_k \rfloor. \tag{178}$$

For $k \geq 0$ and $t \geq 1$ define the valuation window

$$W(k,t) := \left(v_2(3n_{k+i}+1)\right)_{i=0}^{t-1}, \tag{179}$$

and its average

$$A(W(k,t)) := \frac{1}{t}\sum_{i=0}^{t-1} v_2(3n_{k+i}+1). \tag{180}$$

Fix $\varepsilon > 0$. If for some $k \geq 0$ and $t \geq 1$ one has

$$A(W(k,t)) \leq \log_2 3 - \varepsilon, \tag{181}$$

then

$$J(k+t) \geq J(k) + \varepsilon t - 1. \tag{182}$$

In particular, for any fixed $t_{\max} \geq 1$ there exist constants $c > 0$ and $C_0 < \infty$, depending only on ε and $t_{\max}$, such that whenever $1 \leq t \leq t_{\max}$ and (181) holds, one has

$$J(k+t) \geq J(k) + ct - C_0. \tag{183}$$

Proof. Write $a_i := v_2(3n_{k+i}+1)$ for $0 \leq i \leq t-1$, and let

$$S_t := \sum_{i=0}^{t-1} a_i.$$

By definition of the accelerated map,

$$n_{k+i+1} = \frac{3n_{k+i}+1}{2^{a_i}}, \qquad 0 \leq i \leq t-1.$$

Iterating this relation shows that

$$n_{k+t} = \frac{3^t n_k + R_t}{2^{S_t}}, \tag{184}$$

where R_t is a nonnegative integer (a linear combination of powers of 3 times powers of 2) that depends only on n_k and the valuation pattern $(a_0, \ldots, a_{t-1})$. In particular $3^t n_k \leq 3^t n_k + R_t$, so (184) yields the crude lower bound

$$n_{k+t} \geq \frac{3^t n_k}{2^{S_t}}. \tag{185}$$

Taking base–2 logarithms and using (185) gives

$$\log_2 n_{k+t} \geq \log_2 n_k + t\log_2 3 - S_t.$$

The average condition (181) says

$$\frac{S_t}{t} = A(W(k,t)) \leq \log_2 3 - \varepsilon,$$

so $S_t \leq t(\log_2 3 - \varepsilon)$, and therefore

$$\log_2 n_{k+t} \geq \log_2 n_k + t\log_2 3 - t(\log_2 3 - \varepsilon) = \log_2 n_k + \varepsilon t. \tag{186}$$

Passing to integer parts,

$$J(k+t) = \lfloor \log_2 n_{k+t} \rfloor \geq \lfloor \log_2 n_k + \varepsilon t \rfloor.$$

For any real numbers x, y one has

$$\lfloor x + y \rfloor \geq \lfloor x \rfloor + y - 1,$$

so from (186) we obtain

$$J(k + t) \geq \lfloor \log_2 n_k \rfloor + \varepsilon t - 1 = J(k) + \varepsilon t - 1,$$

which is (182).

Finally, given $t_{\max} \geq 1$, take for instance

$$c := \frac{\varepsilon}{2} \quad \text{and} \quad C_0 := 1 + \frac{\varepsilon}{2} t_{\max}.$$

For every $1 \leq t \leq t_{\max}$, the inequality

$$\varepsilon t - 1 \geq \frac{\varepsilon}{2} t - C_0$$

holds by construction, so (182) implies (183). $\quad\square$

7. Completion of the Spectral–Dynamical Implication Chain

Let $(n_k)_{k \geq 0}$ be an infinite forward Collatz orbit under T, and let $J(n)$ denote the block index of n.

Lemma 7.1 (Orbit Cesàro limit functionals). *Let (n_k) be an infinite forward Collatz orbit and define, for $N \geq 1$,*

$$\Lambda_N(f) := \frac{1}{N} \sum_{k=0}^{N-1} f(n_k), \qquad f \in B_{\text{tree},\sigma}.$$

Then:

1. *The family $(\Lambda_N)_{N \geq 1}$ is norm–bounded in $B^*_{\text{tree},\sigma}$, hence every sequence $N_r \to \infty$ admits a further subsequence (still denoted N_r) and a functional $\Lambda \in B^*_{\text{tree},\sigma}$ such that*

$$\Lambda_{N_r} \xrightarrow{\ *\ } \Lambda.$$

   Moreover, for any such limit Λ and any $f \geq 0$ one has $\Lambda(f) \geq 0$ (positivity).

2. *Let $E \subset \mathbb{N}$ be a set that is visited with positive upper density along the orbit (n_k). Then there exist a subsequence $N_r \to \infty$ and a weak–* limit point Λ of (Λ_{N_r}) such that*

$$\Lambda(\mathbf{1}_E) > 0.$$

   In particular, the subsequence and the limit functional can be chosen so that E is detected with strictly positive weight, but this Λ may depend on E.

Proof. Since $B_{\text{tree},\sigma} \hookrightarrow \ell^\infty(\mathbb{N})$ continuously, there exists $C > 0$ such that

$$\|f\|_\infty \leq C \|f\|_{B_{\text{tree},\sigma}} \qquad \text{for all } f \in B_{\text{tree},\sigma}.$$

Hence, for every $N \geq 1$ and every $f \in B_{\text{tree},\sigma}$,

$$|\Lambda_N(f)| = \left| \frac{1}{N} \sum_{k=0}^{N-1} f(n_k) \right| \leq \frac{1}{N} \sum_{k=0}^{N-1} \|f\|_\infty \leq C \|f\|_{B_{\text{tree},\sigma}}.$$

Thus $(\Lambda_N)_{N \geq 1}$ is norm–bounded in $B^*_{\text{tree},\sigma}$. By Banach–Alaoglu, every sequence $N_r \to \infty$ admits a further subsequence (still denoted N_r) and a functional $\Lambda \in B^*_{\text{tree},\sigma}$ such that

$$\Lambda_{N_r} \xrightarrow{\ *\ } \Lambda.$$

If $f \geq 0$, then each $\Lambda_N(f) \geq 0$, since it is an average of nonnegative values. Passing to the weak–* limit along any convergent subsequence N_r gives

$$\Lambda(f) = \lim_{r \to \infty} \Lambda_{N_r}(f) \geq 0,$$

which proves positivity of every weak–* limit.

For (2), fix a set $E \subset \mathbb{N}$ and assume that it is visited with positive upper density along the orbit (n_k), i.e.

$$\bar{d}_E := \limsup_{N \to \infty} \frac{1}{N} \#\{0 \leq k < N : n_k \in E\} > 0.$$

Define $f = \mathbf{1}_E$. Then

$$\Lambda_N(f) = \frac{1}{N} \#\{0 \leq k < N : n_k \in E\}.$$

By the definition of the limsup, there exists a subsequence $N_r \to \infty$ such that

$$\Lambda_{N_r}(f) \longrightarrow \bar{d}_E \quad \text{with} \quad \bar{d}_E > 0.$$

Since the family (Λ_{N_r}) is still norm–bounded, Banach–Alaoglu provides a further subsequence (which we again denote N_r) and a functional $\Lambda \in B^*_{\text{tree},\sigma}$ such that

$$\Lambda_{N_r} \xrightarrow{*} \Lambda.$$

Weak–* convergence means pointwise convergence on $B_{\text{tree},\sigma}$, so in particular

$$\Lambda(f) = \Lambda(\mathbf{1}_E) = \lim_{r \to \infty} \Lambda_{N_r}(\mathbf{1}_E) = \bar{d}_E > 0.$$

Thus for this *chosen* subsequence and limit functional Λ, the indicator of E is evaluated with strictly positive weight. Note that the construction of N_r and Λ depends on E: for a different set E' with positive upper density one may need to choose a different subsequence to realize the corresponding limsup. Hence no claim is made that a single functional Λ detects all such sets simultaneously. □

Lemma 7.2 (Shift invariance). *Let Λ be any weak–* limit of the Cesàro functionals along an orbit. Let U denote composition by the forward map, $(Uf)(n) = f(T(n))$. Then*

$$\Lambda(Uf) = \Lambda(f) \qquad \text{for all } f \in B_{\text{tree},\sigma}.$$

Proof. We compute:

$$\Lambda_N(Uf) = \frac{1}{N} \sum_{k=0}^{N-1} f(T(n_k)) = \frac{1}{N} \sum_{k=0}^{N-1} f(n_{k+1}) = \Lambda_N(f) + \frac{f(n_N) - f(n_0)}{N}.$$

Since f is bounded on $\mathbb{N}$, the final term tends to 0 as $N \to \infty$. Passing to the subsequence N_r along which $\Lambda_{N_r} \to \Lambda$ gives $\Lambda(Uf) = \Lambda(f)$. □

7.1. Discrepancy Decay and P*-Invariance

Definition 7.3 (Discrepancy operator). *For any function $f \in B_{\text{tree},\sigma}$ we define its discrepancy by*

$$D(f) := Pf - f \circ T, \tag{187}$$

where P is the backward Collatz transfer operator (10) and T is the forward Collatz map (1).

Lemma 7.4 (Equivalence of P^*–invariance and discrepancy averages). *Let $(n_k)_{k\geq 0}$ be a forward Collatz orbit and let*

$$\Lambda_N(f) := \frac{1}{N}\sum_{k=0}^{N-1} f(n_k), \qquad f \in B_{\text{tree},\sigma}.$$

Suppose $\Lambda_{N_r} \xrightarrow{*} \Lambda$ in $B^*_{\text{tree},\sigma}$ along some subsequence (N_r). Then for any $f \in B_{\text{tree},\sigma}$,

$$\Lambda(Pf) = \Lambda(f) \iff \lim_{r\to\infty} \frac{1}{N_r}\sum_{k=0}^{N_r-1} D(f)(n_k) = 0,$$

where $D(f)(n) := (Pf)(n) - f(T(n))$.

Proof. Fix $f \in B_{\text{tree},\sigma}$ and write

$$\Lambda_N(Pf) - \Lambda_N(f) = \frac{1}{N}\sum_{k=0}^{N-1}\big(Pf(n_k) - f(n_k)\big).$$

Using the definition $D(f)(n) = Pf(n) - f(T(n))$ and the fact that $T(n_k) = n_{k+1}$ along the orbit, we decompose

$$Pf(n_k) - f(n_k) = \big(Pf(n_k) - f(T(n_k))\big) + \big(f(T(n_k)) - f(n_k)\big) = D(f)(n_k) + f(n_{k+1}) - f(n_k).$$

Hence

$$\begin{aligned}
\Lambda_N(Pf) - \Lambda_N(f) &= \frac{1}{N}\sum_{k=0}^{N-1} D(f)(n_k) + \frac{1}{N}\sum_{k=0}^{N-1}\big(f(n_{k+1}) - f(n_k)\big)\\
&= \frac{1}{N}\sum_{k=0}^{N-1} D(f)(n_k) + \frac{f(n_N) - f(n_0)}{N}.
\end{aligned}$$

Since f is bounded on $\mathbb{N}$, the telescoping term satisfies

$$\frac{f(n_N) - f(n_0)}{N} \longrightarrow 0 \qquad \text{as } N \to \infty.$$

Thus along the subsequence (N_r) we have

$$\Lambda_{N_r}(Pf) - \Lambda_{N_r}(f) = \frac{1}{N_r}\sum_{k=0}^{N_r-1} D(f)(n_k) + o(1), \tag{188}$$

where $o(1) \to 0$ as $r \to \infty$.

Now use weak–$*$ convergence. Since $\Lambda_{N_r} \xrightarrow{*} \Lambda$, we have

$$\Lambda(Pf) - \Lambda(f) = \lim_{r\to\infty}\big(\Lambda_{N_r}(Pf) - \Lambda_{N_r}(f)\big).$$

Combining this with (188) yields

$$\Lambda(Pf) - \Lambda(f) = \lim_{r\to\infty}\frac{1}{N_r}\sum_{k=0}^{N_r-1} D(f)(n_k).$$

Therefore

$$\Lambda(Pf) = \Lambda(f) \iff \lim_{r\to\infty}\frac{1}{N_r}\sum_{k=0}^{N_r-1} D(f)(n_k) = 0,$$

which is the desired equivalence. $\square$

7.2. Spectral Obstructions and Exclusion of Low-Block Invariant Functionals

Proposition 7.5 (No finite–block positive invariant functionals). *Assume the spectral hypotheses for P: positivity, quasi–compactness, $\rho(P) = 1$, peripheral spectrum $\{1\}$, and algebraic and geometric simplicity of the eigenvalue 1 with strictly positive eigenfunction h and dual strictly positive eigenfunctional φ. Then there is no nonzero positive P^*–invariant functional supported on a finite union of blocks $\bigcup_{j \leq J_0} I_j$.*

Proof. Under the spectral hypotheses, the eigenspace of P^* for eigenvalue 1 is one–dimensional and spanned by φ, and $\varphi(f_j) > 0$ for every block indicator f_j.

If Λ were a nonzero positive P^*–invariant functional supported on $\bigcup_{j \leq J_0} I_j$, then Λ must be a scalar multiple of φ. But φ assigns strictly positive mass to every block, contradicting the assumed vanishing of Λ on all I_j with $j > J_0$. $\square$

Remark 7.6 (spectral hypotheses). *The operator P satisfies all of the spectral hypotheses required for the Perron–Frobenius theory on the Banach tree space $B_{\text{tree},\sigma}$. Positivity is established in Proposition 5.39 The Lasota– Yorke inequality of Proposition 4.11 implies quasi–compactness (Theorem 4.17), and the normalization condition $\rho(P) = 1$ follows from Theorem 5.1 The peripheral spectrum is shown to consist solely of the eigenvalue $\{1\}$ in Theorem 6.2, and the Generalized Perron–Frobenius theorem combined with cone–irreducibility (Proposition 6.1) yields algebraic and geometric simplicity of the eigenvalue 1 (Theorem 6.2). Moreover, the corresponding eigenfunction h is strictly positive, and the dual eigenfunctional φ is likewise strictly positive (Proposition 5.39).*

The following statement isolates exactly the dynamical input needed to deduce the Block–Escape Property from the spectral theory of P.

Conjecture 7.7 (Orbitwise discrepancy vanishing). *Let (n_k) be any infinite forward Collatz orbit, and Λ any weak–* limit of its Cesàro averages. There exists a dense subspace $\mathcal{A} \subset B_{\text{tree},\sigma}$ such that for every $f \in \mathcal{A}$,*

$$\lim_{r \to \infty} \frac{1}{N_r} \sum_{k=0}^{N_r - 1} D(f)(n_k) = 0,$$

where (N_r) is the subsequence along which $\Lambda_{N_r} \to \Lambda$. In particular Λ is P^*–invariant on $\mathcal{A}$.

Theorem 7.8 (Conditional Block–Escape). *Assume the spectral hypotheses for P stated in Proposition 7.5, or remark 7.6. If Conjecture 7.7 holds, then every infinite forward Collatz orbit satisfies the Block–Escape Property.*

Proof. Suppose an infinite orbit fails BEP. Then for some J_0, its upper density of visits to $\bigcup_{j \leq J_0} I_j$ is positive. By Lemma 7.1, any limit functional Λ satisfies

$$\Lambda\left(\sum_{j \leq J_0} f_j \right) > 0.$$

If Conjecture 7.7 holds, then Lemma 7.4 implies that Λ is P^*–invariant on a dense subspace, hence on all of $B_{\text{tree},\sigma}$ by continuity. Thus Λ is a nonzero positive P^*–invariant functional supported on $\bigcup_{j \leq J_0} I_j$. This contradicts Proposition 7.5. Therefore no infinite orbit can fail BEP. $\square$

7.3. A Structured Program Toward the Weak Non-Retreat Principle

In this subsection we formalize the analytical strategy for establishing Statement (6) of Theorem 9.1. The aim is to show that every nonperiodic accelerated odd Collatz orbit must admit infinitely many short windows of valuation deficit, from which the weak non–retreat inequalities follow by Theorem 6.13. We present the plan as a sequence of definitions, lemmas, and a culminating theorem.

Definition 7.9 (Valuation Window and Average Valuation). *For an accelerated odd orbit $(n_k)_{k \geq 0}$, a window of length $t \geq 1$ at position k is the finite sequence*

$$W(k, t) := \left(v_2(3n_{k+i} + 1) \right)_{i=0}^{t-1}.$$

Its average valuation is defined by

$$A(W(k,t)) := \frac{1}{t} \sum_{i=0}^{t-1} v_2(3n_{k+i} + 1).$$

Definition 7.10 (Dangerous Window). *Fix $t_{\max} \in \mathbb{N}$ and $\varepsilon > 0$. A window $W(k,t)$ with $1 \leq t \leq t_{\max}$ is called* dangerous *if*

$$A(W(k,t)) > \log_2 3 - \varepsilon.$$

Definition 7.11 (Catalogue of Dangerous Patterns). *For a fixed valuation cutoff $L \geq 1$, define*

$$\mathcal{A}_t(L) := \left\{ (a_0, \ldots, a_{t-1}) \in \{1, \ldots, L\}^t : \frac{1}{t} \sum_{i=0}^{t-1} a_i > \log_2 3 - \varepsilon \right\},$$

and set

$$\mathcal{A}_{\leq t_{\max}}(L) := \bigcup_{t=1}^{t_{\max}} \mathcal{A}_t(L).$$

Elements of $\mathcal{A}_{\leq t_{\max}}(L)$ are the Type I dangerous patterns.

Lemma 7.12 (Congruence Realization of Dangerous Patterns). *Let $\mathbf{a} = (a_0, \ldots, a_{t-1})$ be any valuation pattern with $1 \leq t \leq t_{\max}$ and $1 \leq a_i \leq L$. There exists a modulus*

$$M(\mathbf{a}) = 2^{K(\mathbf{a})} 3^t,$$

where $K(\mathbf{a}) \leq C \sum_{i=0}^{t-1} (a_i + 1)$ for an absolute constant C, and a finite nonempty set of odd residue classes

$$\mathcal{R}(\mathbf{a}) \subset (\mathbb{Z}/M(\mathbf{a})\mathbb{Z})^\times$$

such that an odd integer n_k satisfies

$$v_2(3n_{k+i} + 1) = a_i, \qquad i = 0, \ldots, t-1,$$

if and only if

$$n_k \bmod M(\mathbf{a}) \in \mathcal{R}(\mathbf{a}).$$

Proof. We prove this by induction along the window using the explicit odd–to–odd transition formula and Lemma 6.11.

(1) The congruence condition for a single valuation. Fix an integer $a \geq 1$. Lemma 6.11 shows that the condition

$$n \text{ odd}, \qquad v_2(3n + 1) = a$$

is equivalent to a finite disjunction of residue conditions of the form

$$n \equiv r \pmod{2^{a+1}},$$

where r ranges over a finite subset of odd residues modulo 2^{a+1}. More precisely, Lemma 6.11 proves that

$$v_2(3n + 1) = a \iff 3n + 1 \equiv 0 \pmod{2^a}, \ 3n + 1 \not\equiv 0 \pmod{2^{a+1}},$$

and both congruences are linear conditions on n modulo 2^{a+1}. Thus there is a finite set $\mathcal{R}(a) \subset (\mathbb{Z}/2^{a+1}\mathbb{Z})^\times$ such that n satisfies $v_2(3n + 1) = a$ if and only if $n \bmod 2^{a+1} \in \mathcal{R}(a)$.

(2) Odd-to-odd transitions. For an odd integer n with $\nu_2(3n+1) = a$, the next odd iterate is

$$T(n) = \frac{3n+1}{2^a}.$$

Since $3 \equiv 0 \pmod 3$ and $+1$ ensures $3n + 1 \equiv 1 \pmod 3$, the map $n \mapsto T(n)$ induces a bijection between odd residues modulo $3 \cdot 2^{a+1}$ and odd residues modulo $3 \cdot 2^a$. Thus the congruence condition for n modulo 2^{a+1} lifts to an equivalent condition modulo $3 \cdot 2^{a+1}$, ensuring that the residue class of $T(n)$ is well–defined modulo $3 \cdot 2^a$.

(3) Inductive construction of the residue conditions along the window. Let n_0 denote the starting odd integer at the beginning of the window. We must encode the conditions

$$\nu_2(3n_i + 1) = a_i, \qquad n_{i+1} = T(n_i) = \frac{3n_i + 1}{2^{a_i}}, \qquad i = 0, \ldots, t-1.$$

Assume inductively that for some $i \geq 0$ there exists a modulus M_i and a finite set of odd residue classes $\mathcal{R}_i \subset (\mathbb{Z}/M_i\mathbb{Z})^\times$ such that

$$n_0 \bmod M_i \in \mathcal{R}_i \quad \Longleftrightarrow \quad \nu_2(3n_j + 1) = a_j \text{ for } j = 0, \ldots, i-1.$$

We now add the condition $\nu_2(3n_i + 1) = a_i$.
From Step 1, this condition is equivalent to

$$n_i \bmod 2^{a_i+1} \in \mathcal{R}(a_i).$$

But n_i is an affine linear function of n_0: by iterating

$$n_{j+1} = \frac{3n_j + 1}{2^{a_j}},$$

we obtain (as in the derivation of formula (166)) an explicit expression

$$n_i = \frac{3^i}{2^{S_i}} n_0 + \frac{C_i}{2^{S_i}}, \qquad S_i = \sum_{j=0}^{i-1} a_j,$$

with C_i an integer depending only on the pattern $(a_0, \ldots, a_{i-1})$. Thus the condition $n_i \bmod 2^{a_i+1} \in \mathcal{R}(a_i)$ becomes a linear congruence condition on n_0 modulo $2^{a_i+1+S_i}$.
Therefore, setting

$$M_{i+1} := \mathrm{lcm}(M_i, 2^{a_i+1+S_i}, 3^{i+1}),$$

the set of n_0 satisfying conditions up to index i is a finite union of residue classes modulo M_{i+1}. We denote this finite set by $\mathcal{R}_{i+1} \subset (\mathbb{Z}/M_{i+1}\mathbb{Z})^\times$.

After t steps, we obtain a modulus

$$M(\mathbf{a}) = M_t = 2^{K(\mathbf{a})} 3^t,$$

where $K(\mathbf{a}) = a_0 + 1 + (a_1 + 1 + S_1) + \cdots + (a_{t-1} + 1 + S_{t-1})$ is bounded by a constant multiple of $\sum_i (a_i + 1)$. The corresponding residue set $\mathcal{R}(\mathbf{a}) = \mathcal{R}_t$ satisfies the desired equivalence:

$$n_0 \bmod M(\mathbf{a}) \in \mathcal{R}(\mathbf{a}) \quad \Longleftrightarrow \quad \nu_2(3n_i + 1) = a_i \ (0 \leq i < t).$$

Renaming n_0 as n_k completes the proof. $\quad\square$

Definition 7.13 (Dangerous Residue Set). *Fix $L_0 \geq 1$. Let M be a common multiple of the moduli $M(\mathbf{a})$ over all $\mathbf{a} \in \mathcal{A}_{\leq t_{\max}}(L_0)$. Define the* dangerous residue set *modulo M by*

$$V_{\text{danger}}(L_0) := \bigcup_{\mathbf{a} \in \mathcal{A}_{\leq t_{\max}}(L_0)} \{r \in \mathbb{Z}/M\mathbb{Z} : r \bmod M(\mathbf{a}) \in \mathcal{R}(\mathbf{a})\}.$$

Definition 7.14 (Finite Residue–Valuation Graph). *Fix L_0. Let M be as above. Consider the accelerated odd map*

$$T(n) = \frac{3n+1}{2^{v_2(3n+1)}}$$

on the odd residue classes modulo M. Define a finite directed graph $G_{L_0}(M)$ whose vertices are odd residues $r \in \mathbb{Z}/M\mathbb{Z}$ and with an edge $r \to s$ if $T(r) \equiv s \pmod{M}$. Each vertex carries the label

$$a(r) := v_2(3r+1).$$

Let

$$V_{\geq L_0+1} := \{r : a(r) \geq L_0 + 1\}.$$

Lemma 7.15 (Tail Constraint Under Negation of Weak Non–Retreat). *Fix $t_{\max} \in \mathbb{N}$ and $\varepsilon > 0$, and let $L_0 \geq 1$ be a valuation cutoff. Suppose a nonperiodic accelerated odd orbit $(n_k)_{k \geq 0}$ violates the weak non–retreat condition in the sense that there exists k_0 such that for every $k \geq k_0$ and every $1 \leq t \leq t_{\max}$ the valuation window*

$$W(k,t) := \left(v_2(3n_{k+i}+1)\right)_{i=0}^{t-1}$$

is dangerous, i.e.

$$\frac{1}{t} \sum_{i=0}^{t-1} v_2(3n_{k+i}+1) > \log_2 3 - \varepsilon.$$

Let M be a modulus chosen so that the dangerous patterns of length $t = 1$ with values in $\{1, \ldots, L_0\}$ are realized by a finite set of residue classes $V_{\text{danger}}(L_0) \subset (\mathbb{Z}/M\mathbb{Z})^\times$ as in Lemma 7.12, and let

$$V_{\geq L_0+1} := \{r \in (\mathbb{Z}/M\mathbb{Z})^\times : v_2(3r+1) \geq L_0 + 1\}$$

be the set of residues whose one–step valuation is at least $L_0 + 1$. Then for all sufficiently large k,

$$n_k \bmod M \in V_{\text{danger}}(L_0) \cup V_{\geq L_0+1}.$$

In particular, the tail of (n_k) defines an infinite directed path inside the induced subgraph of $G_{L_0}(M)$ on $V_{\text{danger}}(L_0) \cup V_{\geq L_0+1}$.

Proof. By assumption, there exists k_0 such that for every $k \geq k_0$ and every $1 \leq t \leq t_{\max}$, the window $W(k,t)$ is dangerous:

$$\frac{1}{t} \sum_{i=0}^{t-1} v_2(3n_{k+i}+1) > \log_2 3 - \varepsilon.$$

Fix $k \geq k_0$ and consider the window of length $t = 1$ starting at k:

$$W(k,1) = (a_0), \qquad a_0 := v_2(3n_k+1).$$

Since $W(k,1)$ is dangerous by hypothesis, its average valuation equals a_0 and satisfies

$$a_0 = \frac{1}{1} v_2(3n_k+1) > \log_2 3 - \varepsilon.$$

In particular $a_0 \geq 1$, so $a_0 \in \mathbb{N}$.

We now distinguish two cases.

Case 1: $1 \leq a_0 \leq L_0$. Then the length–one pattern (a_0) belongs to the catalogue $\mathcal{A}_{\leq t_{\max}}(L_0)$ of dangerous patterns with entries bounded by L_0. By Lemma 7.12 applied with $t = 1$ and this pattern, there exists a modulus $M(a_0)$ and a set of residue classes $\mathcal{R}(a_0) \subset (\mathbb{Z}/M(a_0)\mathbb{Z})^{\times}$ such that

$$\nu_2(3n_k + 1) = a_0 \quad \Longleftrightarrow \quad n_k \bmod M(a_0) \in \mathcal{R}(a_0).$$

By construction of the global modulus M and the set $V_{\mathrm{danger}}(L_0)$ (defined as the union of all such residue classes over dangerous $a_0 \leq L_0$, lifted to modulus M), this implies

$$n_k \bmod M \in V_{\mathrm{danger}}(L_0).$$

Case 2: $a_0 \geq L_0 + 1$. In this case the one–step valuation of n_k is already large:

$$\nu_2(3n_k + 1) = a_0 \geq L_0 + 1.$$

By definition of $V_{\geq L_0+1}$ we then have

$$n_k \bmod M \in V_{\geq L_0+1}.$$

In either case, for every $k \geq k_0$ we obtain

$$n_k \bmod M \in V_{\mathrm{danger}}(L_0) \cup V_{\geq L_0+1},$$

which proves the first claim.

For the second claim, recall that the finite directed graph $G_{L_0}(M)$ has vertex set equal to the odd residues modulo M, with an edge $r \to s$ whenever

$$T(r) \equiv s \pmod{M},$$

where T is the accelerated odd Collatz map. Since the orbit (n_k) satisfies $n_{k+1} = T(n_k)$ for all k, the sequence of residues $(n_k \bmod M)_{k \geq 0}$ follows edges in $G_{L_0}(M)$.

We have shown that for all $k \geq k_0$, the residue $n_k \bmod M$ lies in $V_{\mathrm{danger}}(L_0) \cup V_{\geq L_0+1}$, so the tail $(n_k)_{k \geq k_0}$ determines an infinite directed path inside the induced subgraph of $G_{L_0}(M)$ on this vertex set. Discarding the finite initial segment $k < k_0$ does not affect the existence of an infinite path, which completes the proof. $\square$

Theorem 7.16 (Forward Non–Trapping in the Dangerous Regime). *Fix $t_{\max} \in \mathbb{N}$ and $0 < \varepsilon < \log_2 3 - 1$. Then there exists L_0 sufficiently large with the following property. Let M, $V_{\mathrm{danger}}(L_0)$ and $V_{\geq L_0+1}$ be as in the construction of the residue–valuation graph $G_{L_0}(M)$.*

Consider a nonperiodic accelerated odd Collatz orbit $(n_k)_{k \geq 0}$ in the dangerous–windows regime, *meaning that there exists k_0 such that for every $k \geq k_0$ and every $1 \leq t \leq t_{\max}$ the valuation window $W(k, t)$ is dangerous:*

$$A(W(k,t)) = \frac{1}{t} \sum_{i=0}^{t-1} \nu_2(3n_{k+i} + 1) > \log_2 3 - \varepsilon.$$

Then:

1. High–valuation anti–trapping. *The orbit cannot eventually remain entirely inside the high–valuation set $E_{\leq t_{\max}}(L_0 + 1)$. Equivalently, the residue sequence $(n_k \bmod M)$ cannot be eventually contained in $V_{\geq L_0+1}$.*

2.   Confinement to the dangerous residue set. *If the residue sequence $(n_k \bmod M)$ visits $V_{\text{danger}}(L_0)$ infinitely often and also visits $V_{\geq L_0+1}$ infinitely often, then there exists $K \geq k_0$ such that*

$$n_k \bmod M \in V_{\text{danger}}(L_0) \qquad \text{for all } k \geq K.$$

In particular, in the dangerous–windows regime no nonperiodic orbit can have its tail oscillate indefinitely between $V_{\text{danger}}(L_0)$ and $V_{\geq L_0+1}$.

Proof. We break the proof up into two parts, one for each statement.

Part 1: High–valuation anti–trapping

We first show that no nonperiodic orbit can eventually stay inside $V_{\geq L_0+1}$ once L_0 is chosen large enough.

(1.1) Growth estimate for high–valuation steps. Let $a_k := v_2(3n_k + 1)$. If $a_k \geq L_0 + 1$, then

$$n_{k+1} = \frac{3n_k + 1}{2^{a_k}} \leq \frac{3n_k + 1}{2^{L_0+1}} = \frac{3}{2^{L_0+1}} n_k + \frac{1}{2^{L_0+1}}.$$

Define

$$\lambda := \frac{3}{2^{L_0+1}}, \qquad C_0 := \frac{1}{2^{L_0+1}}.$$

If $L_0 \geq 2$ then $\lambda < 1$, and every high–valuation step satisfies

$$n_{k+1} \leq \lambda\, n_k + C_0 \qquad \text{whenever } a_k \geq L_0 + 1. \tag{189}$$

(1.2) Iterated contraction and boundedness. Assume there exists K such that $a_k \geq L_0 + 1$ for all $k \geq K$, i.e. the orbit is eventually contained in $V_{\geq L_0+1}$. Then (189) holds for all $k \geq K$. Iterating gives, by a standard induction,

$$n_{K+t} \leq \lambda^t n_K + C_0 \sum_{j=0}^{t-1} \lambda^j \leq \lambda^t n_K + \frac{C_0}{1-\lambda} \qquad \text{for all } t \geq 0.$$

Since $\lambda < 1$, we have $\lambda^t n_K \to 0$ as $t \to \infty$, and hence

$$\sup_{k \geq K} n_k \leq \frac{C_0}{1-\lambda} < \infty.$$

Thus the accelerated odd orbit (n_k) is bounded. Because the accelerated map is a self–map of the positive odd integers, any bounded orbit is eventually periodic: only finitely many odd integers are $\leq \frac{C_0}{1-\lambda}$, so some value repeats and the map is deterministic thereafter.

This contradicts the hypothesis that the orbit is infinite and nonperiodic. Therefore no nonperiodic orbit can be eventually trapped in $V_{\geq L_0+1}$.

(1.3) Interaction with the dangerous–windows hypothesis. The high–valuation anti–trapping conclusion does not require the dangerous–windows hypothesis. The latter is compatible with this conclusion provided we choose L_0 large enough so that

$$L_0 + 1 > \log_2 3 - \varepsilon,$$

which ensures that every valuation in $V_{\geq L_0+1}$ is individually above the dangerous threshold. In particular, if an orbit were eventually in $V_{\geq L_0+1}$, then every window (of any length) would automatically have average valuation $> \log_2 3 - \varepsilon$, yet the orbit would still be forced to be bounded and hence eventually periodic. This contradiction establishes Part (1).

For any $L_0 \geq 2$, an accelerated odd Collatz orbit that eventually stays in $V_{\geq L_0+1}$ must be bounded and hence eventually periodic. Thus no infinite nonperiodic orbit can be eventually contained in $V_{\geq L_0+1}$.

Part 2: Confinement to the dangerous residue set

We now assume that the orbit visits both $V_{\mathrm{danger}}(L_0)$ and $V_{\geq L_0+1}$ infinitely often and that all windows of length $\leq t_{\max}$ are dangerous. The goal is to show that the orbit must eventually remain in $V_{\mathrm{danger}}(L_0)$.

(2.1) Local bounds along danger and high–valuation segments. Write $a_k := \nu_2(3n_k + 1)$.

For steps with $n_k \in V_{\geq L_0+1}$ (high valuations) we have, as in Part 1,

$$n_{k+1} = \frac{3n_k + 1}{2^{a_k}} \leq \frac{3}{2^{L_0+1}} n_k + \frac{1}{2^{L_0+1}} =: \lambda\, n_k + C_0, \qquad \lambda := \frac{3}{2^{L_0+1}} < 1.$$

For steps with $n_k \in V_{\mathrm{danger}}(L_0)$ we know from the dangerous windows condition (taking $t = 1$) that

$$\log_2 3 - \varepsilon < a_k \leq L_0,$$

and hence

$$n_{k+1} = \frac{3n_k + 1}{2^{a_k}} \leq \frac{3}{2^{a_k}} n_k + \frac{1}{2^{a_k}} \leq \frac{3}{2^{\log_2 3 - \varepsilon}} n_k + 1 = 2^{\varepsilon} n_k + 1.$$

Thus there exists C_1 such that

$$n_{k+1} \ \leq\ 2^{\varepsilon} n_k + C_1 \qquad \text{whenever } n_k \in V_{\mathrm{danger}}(L_0). \tag{190}$$

Iterating (190) along a block of t consecutive indices with $n_k, \dots, n_{k+t-1} \in V_{\mathrm{danger}}(L_0)$ gives

$$n_{k+t} \ \leq\ 2^{\varepsilon t} n_k + C_2(t), \qquad C_2(t) \ \leq\ C_1 \sum_{j=0}^{t-1} 2^{\varepsilon j} \ \leq\ \frac{C_1}{1 - 2^{-\varepsilon}}. \tag{191}$$

Similarly, iterating the high–valuation bound along a block of ℓ consecutive indices with $n_k, \dots, n_{k+\ell-1} \in V_{\geq L_0+1}$ yields

$$n_{k+\ell} \ \leq\ \lambda^{\ell} n_k + C_3(\ell), \qquad C_3(\ell) \ \leq\ \frac{C_0}{1 - \lambda}. \tag{192}$$

(2.2) Patterns and exact multiplicative factors. Consider a finite pattern

$$\mathcal{P} = (t \text{ steps in } V_{\mathrm{danger}}(L_0)) \ \circ\ (\ell \text{ steps in } V_{\geq L_0+1}),$$

realized by the orbit. Let $a_1, \dots, a_t$ be the valuations in the danger segment and $b_1, \dots, b_\ell$ those in the high–valuation segment. The exact multiplicative factor of $\mathcal{P}$ in the accelerated map is

$$\mathcal{F}(\mathcal{P}) := \prod_{i=1}^{t} \frac{3}{2^{a_i}} \cdot \prod_{j=1}^{\ell} \frac{3}{2^{b_j}}. \tag{193}$$

The additive contributions from the $+1$ terms accumulate but do not affect the multiplicative factor.

On the danger segment the dangerous–windows condition enforces

$$\frac{1}{t} \sum_{i=1}^{t} a_i \ > \ \log_2 3 - \varepsilon \quad \Longrightarrow \quad \sum_{i=1}^{t} a_i > t(\log_2 3 - \varepsilon),$$

so

$$\prod_{i=1}^{t} \frac{3}{2^{a_i}} = \frac{3^t}{2^{\sum_{i=1}^{t} a_i}} < \frac{3^t}{2^{t(\log_2 3 - \varepsilon)}} = 2^{\varepsilon t}.$$

On the high–valuation segment $b_j \geq L_0 + 1$, hence

$$\prod_{j=1}^{\ell} \frac{3}{2^{b_j}} \leq \left(\frac{3}{2^{L_0+1}}\right)^{\ell} = \lambda^{\ell}.$$

Combining these,

$$\mathcal{F}(\mathcal{P}) \leq 2^{\varepsilon t} \lambda^{\ell}. \tag{194}$$

Thus $2^{\varepsilon t} \lambda^{\ell}$ is a worst–case upper bound: if $2^{\varepsilon t} \lambda^{\ell} \leq 1$, then necessarily $\mathcal{F}(\mathcal{P}) \leq 1$.

Now decompose the tail of the orbit into alternating "danger blocks" and "high–valuation excursions". Let k_j be the start of the j-th danger block, of length t_j, followed by an excursion of length ℓ_j:

$$n_{k_j} \in V_{\text{danger}}(L_0), \quad n_{k_j+1}, \ldots, n_{k_j+t_j-1} \in V_{\text{danger}}(L_0), \quad n_{k_j+t_j}, \ldots, n_{k_j+t_j+\ell_j-1} \in V_{\geq L_0+1},$$

and assume $n_{k_j+t_j+\ell_j} \in V_{\text{danger}}(L_0)$ again. For this realized pattern $\mathcal{P}_j$ we have an affine relation

$$n_{k_j+t_j+\ell_j} = \mathcal{F}(\mathcal{P}_j) \, n_{k_j} + \beta_j,$$

for some integer β_j depending only on the residue pattern; in particular there is a constant C_4 (independent of j) such that

$$n_{k_{j+1}} \leq \mathcal{F}(\mathcal{P}_j) \, n_{k_j} + C_4 \qquad \text{for all } j. \tag{195}$$

(2.3) Contraction along each finite pattern. Let

$$\theta := \log_2 3 - \varepsilon.$$

By the dangerous–windows hypothesis for the full orbit, we in particular have

$$a_k = \nu_2(3n_k + 1) > \theta \qquad \text{for every } k \geq k_0,$$

since this is the case $t = 1$ in the window condition.

Fix one of the patterns $\mathcal{P}_j$ from Step 2.2, of total length

$$L_j := t_j + \ell_j.$$

Let $a_0^{(j)}, \ldots, a_{L_j-1}^{(j)}$ be the valuation sequence along this pattern, so

$$a_m^{(j)} = \nu_2(3n_{k_j+m} + 1), \qquad m = 0, \ldots, L_j - 1.$$

By the preceding remark,

$$a_m^{(j)} > \theta \qquad \text{for all } m.$$

Define the excesses

$$\delta_m^{(j)} := a_m^{(j)} - \theta, \qquad m = 0, \ldots, L_j - 1.$$

Since each $a_m^{(j)}$ is an integer and $a_m^{(j)} > \theta$, we have

$$a_m^{(j)} \geq \lfloor \theta \rfloor + 1.$$

Choose $\varepsilon > 0$ small enough that

$$1 < \theta = \log_2 3 - \varepsilon < 2,$$

so $\lfloor \theta \rfloor = 1$ and hence $a_m^{(j)} \geq 2$ for all m. Thus

$$\delta_m^{(j)} = a_m^{(j)} - \theta \geq 2 - \theta = 2 - (\log_2 3 - \varepsilon) = 2 - \log_2 3 + \varepsilon.$$

Set

$$\eta := 2 - \log_2 3 + \varepsilon.$$

Since $2 - \log_2 3 > 0$, we have

$$\eta > \varepsilon, \qquad \delta_m^{(j)} \geq \eta \quad \text{for all } m.$$

The exact multiplicative factor of the pattern $\mathcal{P}_j$ is

$$\mathcal{F}(\mathcal{P}_j) = \prod_{m=0}^{L_j-1} \frac{3}{2^{a_m^{(j)}}} = \prod_{m=0}^{L_j-1} \frac{3}{2^{\theta + \delta_m^{(j)}}} = \prod_{m=0}^{L_j-1} \left(\frac{3}{2^\theta}\right) 2^{-\delta_m^{(j)}}.$$

As before,

$$\frac{3}{2^\theta} = \frac{3}{2^{\log_2 3 - \varepsilon}} = 2^\varepsilon,$$

so

$$\mathcal{F}(\mathcal{P}_j) = 2^{L_j \varepsilon} 2^{-\sum_{m=0}^{L_j-1} \delta_m^{(j)}} = 2^{L_j \varepsilon - \sum_{m=0}^{L_j-1} \delta_m^{(j)}}.$$

Using $\delta_m^{(j)} \geq \eta$,

$$\sum_{m=0}^{L_j-1} \delta_m^{(j)} \geq L_j \eta,$$

so we obtain

$$\mathcal{F}(\mathcal{P}_j) \leq 2^{L_j \varepsilon - L_j \eta} = 2^{-L_j(\eta - \varepsilon)}.$$

Since $\eta - \varepsilon = 2 - \log_2 3 > 0$ is fixed, we can define

$$\Lambda := 2^{-(\eta - \varepsilon)} = 2^{-(2 - \log_2 3)} \in (0, 1),$$

and conclude

$$\mathcal{F}(\mathcal{P}_j) \leq \Lambda^{L_j} \leq \Lambda \qquad \text{for every } j.$$

Thus each danger–excursion pattern $\mathcal{P}_j$ is uniformly contracting in its multiplicative factor, with a contraction constant $\Lambda < 1$ that depends only on ε (and not on j or the particular pattern).

Crucially, this argument uses only the pointwise inequality $a_m^{(j)} > \theta$ (the case $t = 1$ of the dangerous–windows condition) plus integrality, and applies directly to the finite sequence along $\mathcal{P}_j$. No completion to a closed walk in the residue graph is required.

(2.4) Boundedness and periodicity. From (195) and the uniform contraction bound we obtain

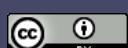$$n_{k_{j+1}} \leq \Lambda\, n_{k_j} + C_4 \qquad \text{for all } j,$$

with $0 < \Lambda < 1$. Iterating this affine inequality yields

$$n_{k_j} \leq \Lambda^j n_{k_0} + \frac{C_4}{1 - \Lambda},$$

so the subsequence (n_{k_j}) is bounded.

The intermediate values inside each danger block and each high–valuation excursion are then bounded using (191) and (192): the constants $C_2(t)$ and $C_3(\ell)$ are uniformly bounded in $t \leq t_{\max}$ and $\ell \geq 0$, and each block has finite length. Consequently the entire tail of the orbit is bounded.

Any bounded accelerated odd Collatz orbit is eventually periodic, so we obtain a contradiction with the assumption that the orbit is infinite and nonperiodic while visiting both $V_{\text{danger}}(L_0)$ and

$V_{\geq L_0 + 1}$ infinitely often. Therefore, under the dangerous–windows hypothesis, an infinite nonperiodic orbit cannot oscillate indefinitely between the two sets: it must eventually remain in $V_{\text{danger}}(L_0)$.

For Part (2), the only ingredient from the dangerous–windows hypothesis needed for the contraction argument is the pointwise bound

$$v_2(3n_k + 1) > \log_2 3 - \varepsilon \qquad \text{for all sufficiently large } k$$

on the valuations along the orbit tail. Together with integrality, this forces every danger–excursion pattern $\mathcal{P}_j$ to have multiplicative factor $\mathcal{F}(\mathcal{P}_j) \leq \Lambda < 1$ uniformly in j, yielding a global contraction inequality

$$n_{k_{j+1}} \ \leq \ \Lambda n_{k_j} + C_4$$

along the subsequence of entries into $V_{\text{danger}}(L_0)$. This implies boundedness and eventual periodicity of the orbit, contradicting the hypothesis of an infinite nonperiodic orbit that visits both $V_{\text{danger}}(L_0)$ and $V_{\geq L_0+1}$ infinitely often. Hence any infinite nonperiodic orbit satisfying the dangerous–windows condition must eventually remain in the dangerous residue set, completing the proof of the theorem. $\square$

Theorem 7.17 (Non–Realizability of Dangerous Residue Cycles). *Let H be the induced subgraph of the residue–valuation graph $G_{L_0}(M)$ on the dangerous residue set $V_{\text{danger}}(L_0)$. Let*

$$C = (v_0 \to v_1 \to \cdots \to v_{p-1} \to v_0)$$

be any directed cycle in H, where $v_i \in V_{\text{danger}}(L_0)$ and seach $v_i \to v_{i+1}$ (indices taken modulo p) is an edge in $G_{L_0}(M)$. sThen no infinite accelerated odd Collatz orbit $(n_k)_{k \geq 0}$ can eventually realize C modulo M in the following sense: there do not exist integers $i \geq 0$ and $p \geq 1$ such that

$$n_k \equiv v_{(k-i) \bmod p} \pmod{M} \qquad \text{for all } k \geq i.$$

Equivalently, no nonperiodic accelerated odd orbit can have its residue sequence $s(n_k \bmod M)$ eventually trapped on a directed cycle inside the subgraph H.

Proof. We prove that no infinite nonperiodic accelerated odd Collatz orbit can eventually follow a directed cycle in the dangerous residue graph H.

(1) Structure of a dangerous cycle. Let

$$C = (v_0 \to v_1 \to \cdots \to v_{p-1} \to v_0)$$

be a directed cycle in H of length $p \geq 1$. For each i set

$$a_i := v_2(3v_i + 1).$$

Since $v_i \in V_{\text{danger}}(L_0)$, each a_i satisfies

$$1 \leq a_i \leq L_0.$$

Define

$$S_p := \sum_{i=0}^{p-1} a_i, \qquad \alpha := \frac{3^p}{2^{S_p}}.$$

Then $S_p \geq p \geq 1$, so $2^{S_p} > 1$ and α is a nontrivial rational number. In lowest terms we may write

$$\alpha = \frac{A}{B}, \qquad A = 3^p \text{ odd}, \quad B = 2^{S_p} > 1.$$

In particular $\alpha \neq 1$ since $3^p \neq 2^{S_p}$.

(2) Lifting the cycle to an integer orbit. Assume for contradiction that there exists an infinite nonperiodic accelerated odd Collatz orbit $(n_k)_{k \geq 0}$ that eventually follows C modulo M. Thus there is $i_0 \geq 0$ such that for all $k \geq i_0$,

$$n_k \equiv v_{(k-i_0) \bmod p} \pmod{M}.$$

For $k = i_0 + mp + r$ with $0 \leq r < p$, iterating the accelerated map

$$T(x) = \frac{3x+1}{2^{a(x)}}$$

along one period of the cycle yields an affine relation of the form

$$n_{k+p} = \frac{3^p}{2^{S_p}} n_k + \frac{B_r}{2^{S_p}} = \alpha n_k + \beta_r, \tag{196}$$

where $B_r \in \mathbb{Z}$ depends only on the residue class v_r, and $\beta_r := B_r / B$.

(3) Recurrence and explicit solution. Fix $r \in \{0, \ldots, p-1\}$ and define the subsequence

$$x_m := n_{i_0+mp+r}, \qquad m \geq 0.$$

Then (196) becomes

$$x_{m+1} = \alpha x_m + \beta_r, \qquad \alpha = \frac{A}{B}, \quad \beta_r = \frac{B_r}{B}.$$

Since $\alpha \neq 1$, the standard solution of this affine recurrence gives

$$x_m = \alpha^m x_0 + \beta_r \frac{\alpha^m - 1}{\alpha - 1} = \alpha^m C_r + D_r,$$

where

$$C_r := x_0 - \frac{\beta_r}{\alpha - 1}, \qquad D_r := \frac{\beta_r}{\alpha - 1}.$$

Because $x_m \in \mathbb{Z}$ for all m, we have

$$\alpha^m C_r \in \mathbb{Z} - D_r \qquad \text{for all } m \geq 0. \tag{197}$$

In particular $\alpha^m C_r$ is rational for all m.

(4) Normal form for C_r and a divisibility condition. Write C_r in lowest terms as

$$C_r = \frac{c}{2^t d},$$

where $c, d \in \mathbb{Z}$ are odd and $t \geq 0$. Using $\alpha = A/B = 3^p/2^{S_p}$ we obtain

$$\alpha^m C_r = \left(\frac{3^p}{2^{S_p}}\right)^m \frac{c}{2^t d} = \frac{3^{pm} c}{2^{mS_p+t} d}.$$

By (197) the quantity $\alpha^m C_r$ must be an integer for all $m \geq 0$, hence

$$2^{mS_p+t} d \mid 3^{pm} c \qquad \text{for all } m \geq 0. \tag{198}$$

(5) Denominator–growth forcing $C_r = 0$. Because d and 3^{pm} are odd, the right–hand side of (198), namely $3^{pm} c$, is an odd integer when $c \neq 0$. From (198) we deduce that for all $m \geq 0$,

$$2^{mS_p+t} \mid 3^{pm} c.$$

Since $S_p \geq 1$, we have $mS_p + t \geq S_p + t \geq 1$ for $m \geq 1$, so $2^{mS_p+t} \geq 2$. If $c \neq 0$, then $3^{pm}c$ is a nonzero odd integer that is divisible by 2, which is impossible. Therefore $c = 0$, and hence

$$C_r = 0.$$

(6) Eventual periodicity. With $C_r = 0$, the explicit formula reduces to

$$x_m = D_r \qquad \text{for all } m \geq 0.$$

Thus each residue–class subsequence $(n_{i_0+mp+r})_{m\geq 0}$ is constant. In particular the tail $(n_k)_{k\geq i_0}$ takes values in the finite set

$$\{n_{i_0}, n_{i_0+1}, \ldots, n_{i_0+p-1}\}$$

in a periodic pattern of period dividing p. Hence the orbit (n_k) is eventually periodic.

This contradicts the assumption that the orbit is infinite and nonperiodic. Therefore no infinite nonperiodic accelerated odd Collatz orbit can eventually follow a directed cycle in the dangerous residue graph H, which proves the theorem. $\square$

Lemma 7.18 (Finite Graph Periodicity). *Let H be the induced subgraph of $G_{L_0}(M)$ on $V_{\mathrm{danger}}(L_0)$. Then every infinite directed path in H is eventually periodic. That is, there exist integers $i \geq 0$ and $p \geq 1$ such that the vertex sequence $(v_k)_{k\geq 0}$ satisfies*

$$v_{k+p} = v_k \qquad \text{for all } k \geq i.$$

Proof. Since H has finitely many vertices, say N, the sequence (v_k) takes values in a finite set of size N. Hence there exist indices $0 \leq i < j$ with $v_i = v_j$; choose such a pair with $j - i$ minimal, and set $p := j - i \geq 1$. The segment

$$v_i \to v_{i+1} \to \cdots \to v_{j-1} \to v_j = v_i$$

is a directed cycle of length p in H. Because the path (v_k) is directed, once it reaches v_i again, all future steps must follow the outgoing edges of this cycle. Thus

$$v_{k+p} = v_k \qquad \text{for all } k \geq i,$$

and the path is eventually periodic. $\square$

Lemma 7.19 (Finite Graph Obstruction). *Assume Theorem 7.17. Then no infinite directed path in H can arise as the sequence of residues n_k mod M for a valid infinite nonperiodic accelerated odd Collatz orbit $(n_k)_{k\geq 0}$.*

Proof. Let $(n_k)_{k\geq 0}$ be an infinite nonperiodic accelerated odd Collatz orbit, and set $v_k := n_k \bmod M$. Suppose that $v_k \in V_{\mathrm{danger}}(L_0)$ for all sufficiently large k, so that $(v_k)_{k\geq K}$ is an infinite directed path in H for some K. By Lemma 7.18, this path is eventually periodic: there exist integers $i \geq K$ and $p \geq 1$ such that

$$v_{k+p} = v_k \qquad \text{for all } k \geq i.$$

In particular, the finite list

$$C = (v_i, v_{i+1}, \ldots, v_{i+p-1})$$

forms a directed cycle in H, and the residue sequence (v_k) follows this cycle from time i onward:

$$v_k = v_{(k-i) \bmod p+i} \qquad \text{for all } k \geq i.$$

Equivalently,

$$n_k \equiv v_{(k-i) \bmod p+i} \pmod{M} \qquad \text{for all } k \geq i.$$

This is precisely excluded by Theorem 7.17, which asserts that no infinite accelerated odd Collatz orbit can eventually follow a cycle in H modulo M. Hence, under the Theorem 7.17, no such infinite path in H can be realized by the residue sequence of a valid infinite nonperiodic orbit. □

Theorem 7.20 (Weak Non–Retreat). *Assume Lemmas 7.12, 7.15, 7.18, and 7.19, together with Theorem 7.16. Then every nonperiodic accelerated odd Collatz orbit admits infinitely many windows of length $1 \leq t \leq t_{\max}$ whose average valuation satisfies*

$$A(W(k,t)) \leq \log_2 3 - \varepsilon.$$

Consequently, by Theorem 6.13 there exist constants $c > 0$ and C_0 such that

$$J(k+t) \geq J(k) + ct - C_0$$

for infinitely many pairs (k,t), and the weak non–retreat property of 1.1(6) holds under these assumptions.

Proof. Let $(n_k)_{k \geq 0}$ be a nonperiodic accelerated odd Collatz orbit, and suppose toward contradiction that only finitely many pairs (k,t) with $1 \leq t \leq t_{\max}$ satisfy

$$A(W(k,t)) \leq \log_2 3 - \varepsilon.$$

Then there exists k_0 such that for all $k \geq k_0$ and all $1 \leq t \leq t_{\max}$, the window $W(k,t)$ is dangerous. This places us in the dangerous–windows regime of Theorem 7.16. Fix L_0 and the associated modulus M and residue sets $V_{\mathrm{danger}}(L_0)$ and $V_{\geq L_0+1}$. By Lemma 7.15 there exists $k_1 \geq k_0$ such that for all $k \geq k_1$,

$$n_k \bmod M \in V_{\mathrm{danger}}(L_0) \cup V_{\geq L_0+1}.$$

By Theorem 7.16 (1), for L_0 chosen large enough no nonperiodic orbit can eventually remain entirely inside the high–valuation set $E_{\leq t_{\max}}(L_0 + 1)$, hence the residue sequence cannot be eventually contained in $V_{\geq L_0+1}$. There are therefore infinitely many k with $k \geq k_1$ and $n_k \bmod M \in V_{\mathrm{danger}}(L_0)$. If the sequence $(n_k \bmod M)$ visits $V_{\geq L_0+1}$ only finitely many times, then there exists $K \geq k_1$ with

$$n_k \bmod M \in V_{\mathrm{danger}}(L_0) \qquad \text{for all } k \geq K.$$

If, on the other hand, it visits $V_{\geq L_0+1}$ infinitely often, then Theorem 7.16 (2) applies in the dangerous–windows regime and again yields an index $K \geq k_1$ such that

$$n_k \bmod M \in V_{\mathrm{danger}}(L_0) \qquad \text{for all } k \geq K.$$

In either case, the tail $(n_k \bmod M)_{k \geq K}$ defines an infinite directed path inside the induced subgraph H of $G_{L_0}(M)$ on $V_{\mathrm{danger}}(L_0)$.

By Lemma 7.19, such an infinite path cannot arise from a valid infinite nonperiodic Collatz orbits. This contradicts the assumed nonperiodicity of (n_k) and the existence of only finitely many valuation–deficit windows.

Therefore every nonperiodic accelerated odd orbit must admit infinitely many pairs (k,t) with $1 \leq t \leq t_{\max}$ and

$$A(W(k,t)) \leq \log_2 3 - \varepsilon.$$

For each such window, Theorem 6.13 gives a forward drift estimate

$$J(k+t) \geq J(k) + ct - C_0$$

with $c > 0$ and C_0 depending only on the global parameters. Since there are infinitely many such windows, these inequalities hold along an infinite subsequence of times, which is precisely the weak non–retreat property. $\square$

7.4. Logical Closure of the Spectral-Dynamical Framework

In this section we formalize all the propositions and their proofs for the Dynamical Forms Theorem 1.1

$(3) \Rightarrow (2)$

Proposition 7.21 (Orbit–supported invariant functionals are impossible). *Let P be the backward Collatz transfer operator on $B_{\text{tree},\sigma}$ and assume the Peripheral Spectral Classification Theorem 6.2. Then there is no nonzero P^*–invariant functional in $B^*_{\text{tree},\sigma}$ whose support is contained in a single forward Collatz orbit.*

In particular, statement (3) of Theorem 1.2 (existence of a nonzero orbit–supported invariant functional for every infinite orbit) cannot hold unless there are no infinite forward orbits at all. Hence (3) implies (2).

Proof. By Theorem 6.2 and positivity of P, there exist $h > 0$ and $\phi > 0$ with $Ph = h$, $P^*\phi = \phi$, and the eigenspace for eigenvalue 1 of P^* is one–dimensional and spanned by ϕ. Thus if $\Psi \in B^*_{\text{tree},\sigma}$ satisfies $P^*\Psi = \Psi$, then $\Psi = c\,\phi$ for some scalar c.

Fix a forward orbit $O^+(n_0) = \{T^k(n_0) : k \geq 0\}$ and suppose there exists a nonzero P^*–invariant functional Ψ supported on this orbit, in the sense that $\Psi(f) = 0$ whenever f vanishes on $O^+(n_0)$. Then $\Psi = c\,\phi$ with $c \neq 0$.

Consider the nonnegative test function

$$f(n) := \mathbf{1}_{\mathbb{N}\setminus O^+(n_0)}(n).$$

By support of Ψ on the orbit we have $\Psi(f) = 0$. On the other hand, $f \geq 0$ and $f \not\equiv 0$, so strict positivity of ϕ gives $\phi(f) > 0$ and hence

$$\Psi(f) = c\,\phi(f) \neq 0,$$

a contradiction. Therefore no nonzero P^*–invariant functional can be supported on a single forward orbit. If statement (3) of Theorem 1.1 holds, then every infinite forward orbit would support such a functional, which we have just ruled out. Hence there are no infinite forward orbits, and in particular every orbit has bounded block index: $\sup_{k \geq 0} j(n_k) < \infty$, which is statement (2). $\square$

$(3) + (8) \Rightarrow (1)$

Proposition 7.22 (Strong Collatz from orbit averages and residue graphs). *Assume Conjecture 6.3 ((3) of Theorem 1.1) with Theorems 7.17 and 7.16. Then statement (1) of Theorem 1.1 (Strong Collatz Conjecture) holds.*

Proof. *(1): Exclusion of nontrivial cycles.* By Theorems 7.17 and 7.16, no nontrivial accelerated odd Collatz cycle can persist inside the dangerous residue graph, and every nonperiodic accelerated odd orbit eventually escapes it. The argument given previously in the proof of Theorem 7.20 therefore excludes all nontrivial cycles; we do not repeat it here.

(2) Exclusion of infinite orbits via orbit–averaging and spectral rigidity. Suppose for contradiction that there exists an infinite forward Collatz orbit $(n_k)_{k \geq 0}$ with base point n_0. By statement (3) (Orbit–Averaging Conjecture), this orbit produces a nonzero P^*–invariant functional

$$\Lambda \in B^*_{\text{tree},\sigma}, \qquad P^*\Lambda = \Lambda,$$

which is obtained as an orbit–average along $(n_k)_{k\geq 0}$. In particular, if a test function $f \in B_{\text{tree},\sigma}$ satisfies $f(n_k) = 0$ for all sufficiently large k, then the orbit–average of f along (n_k) vanishes and hence

$$\Lambda(f) = 0. \tag{199}$$

On the other hand, the spectral classification theorem (uniqueness of the positive eigenfunctional, Theorem 5.1) asserts that every P^*–invariant functional is a scalar multiple of the distinguished Perron–Frobenius eigenfunctional ϕ:

$$\Lambda = \alpha\,\phi, \qquad \alpha \neq 0.$$

We now construct a smooth cutoff that lies in $B_{\text{tree},\sigma}$ and vanishes along the tail of the orbit. By the nontrapping and escape statement (8), the forward orbit (n_k) cannot remain indefinitely in any fixed low region of the tree. Hence we may choose $m \in \mathbb{N}$ so large that

$$n_k > m \qquad \text{for all } k \geq K$$

for some $K \geq 0$. Define a test function $h_m \colon \mathbb{N} \to [0, \infty)$ by

$$h_m(n) := m^{-2\sigma}\,\mathbf{1}_{\{n \leq m\}}\,n^{-\sigma}.$$

We first verify that $h_m \in B_{\text{tree},\sigma}$. The weighted ℓ^1 norm is finite:

$$\|h_m\|_{\ell^1_\sigma} = \sum_{n \geq 1} h_m(n)\,n^{-\sigma} = \sum_{n \leq m} m^{-2\sigma}\,n^{-2\sigma} \leq m^{-2\sigma}\,m\,1^{-2\sigma} = m^{1-2\sigma} < \infty$$

since $\sigma > 1$. Moreover h_m is supported on the finite set $\{1, \ldots, m\}$, so only finitely many tree blocks I_j intersect its support. In each such block, the contribution to the block seminorm comes from finitely many points with uniformly bounded size, multiplied by the decaying prefactor $m^{-2\sigma}$. Thus the tree seminorm of h_m is finite and $h_m \in B_{\text{tree},\sigma}$.

By construction, we have $h_m(n_k) = 0$ for all $k \geq K$, since $n_k > m$ implies $h_m(n_k) = 0$. Therefore h_m vanishes along the tail of the orbit, and the defining orbit–averaging property of Λ yields

$$\Lambda(h_m) = 0$$

via (199). On the other hand, write $\phi(f) = \sum_{n \geq 1} f(n)\,h(n)$ with h the strictly positive Perron–Frobenius eigenfunction satisfying $Ph = h$. Then

$$\phi(h_m) = \sum_{n \leq m} h_m(n)\,h(n) = m^{-2\sigma}\sum_{n \leq m} n^{-\sigma}h(n) > 0,$$

because $h(n) > 0$ for all $n \geq 1$ and the sum is finite and nonempty. Combining the two expressions for $\Lambda(h_m)$ gives

$$0 = \Lambda(h_m) = \alpha\,\phi(h_m),$$

which contradicts $\alpha \neq 0$ and $\phi(h_m) > 0$. This contradiction shows that no infinite forward Collatz orbit can exist.

We have ruled out both nontrivial cycles (Step 1) and infinite orbits (Step 2). Hence every forward Collatz orbit is finite, and the only cycle is the trivial $1 \to 4 \to 2 \to 1$ cycle. This is precisely statement (1) of Theorem 1.1. $\square$

$(4) \Rightarrow (3)$

Proposition 7.23 (Block orbit averages imply the Orbit–Averaging Conjecture). *Assume statement* (4) *of Theorem* 1.1 (the Block–Orbit–Averaging Conjecture). *Then statement* (3) *of Theorem* 1.1 (the Orbit– Averaging Conjecture, Conjecture 6.3) *holds. In particular,* (4) *implies* (3) *in Theorem* 1.1.

Proof. Let $(n_k)_{k \geq 0}$ be an infinite forward Collatz orbit with base point n_0. For each $N \geq 1$ consider the orbit Cesàro functional

$$\Phi_N(f) := \frac{1}{N} \sum_{k=0}^{N-1} f(n_k), \qquad f \in B_{\text{tree},\sigma}.$$

By Lemma 5.30, the family $(\Phi_N)_{N \geq 1}$ is uniformly bounded in $B^*_{\text{tree},\sigma}$, hence weak* relatively compact.

Statement (4) asserts, for this orbit, the existence of a nontrivial block–orbit average. Concretely, it guarantees that there exists a sequence $N_j \to \infty$ and a nonzero functional

$$\Lambda \in B^*_{\text{tree},\sigma} \setminus \{0\}$$

such that

$$\Phi_{N_j} \xrightarrow{\text{weak}^*} \Lambda.$$

Thus Λ is a nonzero weak* limit point of the orbit Cesàro functionals for (n_k). By Proposition 5.31, every such weak* limit of (Φ_N) along a forward Collatz orbit satisfies two key properties:

$$P^* \Lambda = \Lambda,$$

and Λ is supported entirely on the orbit $O^+(n_0)$ in the sense that

$$f(n_k) = 0 \ \forall k \geq 0 \quad \Longrightarrow \quad \Lambda(f) = 0.$$

Since $\Lambda \neq 0$, this gives exactly the conclusion of statement (3) for the orbit $O^+(n_0)$: it produces a nonzero P^*–invariant linear functional supported on that orbit.

Because the choice of the infinite orbit (n_k) was arbitrary, the same argument applies to every infinite forward orbit. Hence the Orbit–Averaging Conjecture (3) holds whenever the Block–Orbit– Averaging statement (4) holds, and we have shown $(4) \Rightarrow (3)$ in Theorem 1.1. $\quad\square$

$(5) \Rightarrow (4)$

Proposition 7.24 (Implication from supercritical linear block growth to BOA). *Assume Statement* (5) (Block–Escape Implies Supercritical Linear Block Growth, Conjecture 6.10). *Then Statement* (4) (Block– Orbit–Averaging, Conjecture 6.4) *holds.*

Proof. Suppose for contradiction that Statement (4) fails. Then there exists an infinite forward Collatz orbit

$$O^+(n_0) = \{T^k(n_0)\}_{k \geq 0}$$

such that for every integer $J \geq 0$ the orbit does *not* spend a positive proportion of time in the finite union of low blocks $\bigcup_{j \leq J} I_j$. In terms of the block index

$$J(k) := \text{the unique } j \geq 0 \text{ with } T^k(n_0) \in I_j,$$

this means that for every $J \geq 0$ the lower asymptotic frequency of visits to $\bigcup_{j \leq J} I_j$ is zero. Equivalently, the orbit satisfies the block–escape condition of Definition 6.5, namely

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k) \leq J\}} = 0 \qquad \text{for every } J \geq 0, \tag{200}$$

or the Block–Escape Property. By Statement (5) (Conjecture 6.10), every infinite orbit satisfying BEP must exhibit supercritical linear block growth. Concretely, there exists a constant

$$\alpha > \frac{\log 2}{\log 6}$$

and a strictly increasing subsequence $(k_\ell)_{\ell \geq 1}$ with $k_\ell \to \infty$ such that

$$J(k_\ell) \;\geq\; \alpha\, k_\ell \qquad \text{for all } \ell \geq 1. \tag{201}$$

Thus the orbit $O^+(n_0)$ satisfies both the block–escape condition (200) and the linear growth condition (201).

   We now invoke Proposition 7.5. That result states that if an infinite orbit satisfies the block–escape condition

$$\forall J_0 \geq 0: \quad \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k) \leq J_0\}} = 0 \tag{200}$$

and there exists $\beta > 0$ and a subsequence (k_ℓ) with

$$J(k_\ell) \;\geq\; \beta\, k_\ell \quad \text{for all } \ell, \tag{201}$$

then the orbit violates the universal exponential growth bound of Lemma 6.7. In the proof of Proposition 7.5 one obtains

$$\limsup_{\ell \to \infty} \frac{1}{k_\ell} \log T^{k_\ell}(n_0) \;\geq\; \beta \log 6.$$

Taking $\beta = \alpha$ and using $\alpha > \log 2 / \log 6$, this yields

$$\limsup_{\ell \to \infty} \frac{1}{k_\ell} \log T^{k_\ell}(n_0) \;\geq\; \alpha \log 6 \;>\; \log 2,$$

which contradicts Lemma 6.7, since that lemma gives the universal upper bound

$$\limsup_{k \to \infty} \frac{1}{k} \log T^k(n_0) \;\leq\; \log 2.$$

Therefore no infinite orbit can simultaneously satisfy BEP and the supercritical linear growth condition (201). In particular, our initial assumption that there exists an infinite orbit for which Statement (4) fails leads to a contradiction, because failure of (4) was exactly what forced BEP in (200).

   Taking the contrapositive, every infinite orbit must fail BEP. Thus for each infinite orbit $O^+(n_0)$ there exists some $J_0 \geq 0$ such that the lower asymptotic frequency of visits to $\bigcup_{j \leq J_0} I_j$ is strictly positive:

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{J(k) \leq J_0\}} \;>\; .$$

This is precisely the Block–Orbit–Averaging statement (4). Hence, under Statement (5), Statement (4) holds. □

(6) ⇒ (5)

Lemma 7.25 (Flexible choice of drift constants). *Let (n_k) be a forward orbit and suppose there exist $\varepsilon > 0$, $t_{\max} \in \mathbb{N}$, and $K_0 \geq 0$ such that for all $k \geq K_0$ and all $1 \leq t \leq t_{\max}$ one has*

$$j(n_{k+t}) \;\geq\; j(n_k) + \varepsilon\, t - 1. \tag{202}$$

Then for any $\delta \in (0, \varepsilon)$ the choice

$$c := \varepsilon - \delta, \qquad C_0 := 1$$

also yields a valid weak non–retreat inequality:

$$j(n_{k+t}) \ \geq \ j(n_k) + c\,t - C_0, \qquad 1 \le t \le t_{\max}, \ k \ge K_0. \tag{203}$$

Moreover, for these constants one has

$$\frac{c\,t_{\max} - C_0}{t_{\max}} = (\varepsilon - \delta) - \frac{1}{t_{\max}}. \tag{204}$$

Proof. Fix $\delta \in (0, \varepsilon)$ and set $c := \varepsilon - \delta$ and $C_0 := 1$. Then for each $1 \le t \le t_{\max}$ we compute

$$(\varepsilon t - 1) - (c\,t - C_0) = (\varepsilon t - 1) - \big((\varepsilon - \delta)t - 1\big) = \delta t \ \geq \ \delta > 0.$$

Thus the right–hand side of (202) is strictly larger than $c\,t - C_0$ for every $t \in [1, t_{\max}]$, and hence (202) implies (203) for all $k \ge K_0$. The identity (204) is an immediate algebraic rearrangement:

$$\frac{c\,t_{\max} - C_0}{t_{\max}} = \frac{(\varepsilon - \delta)t_{\max} - 1}{t_{\max}} = (\varepsilon - \delta) - \frac{1}{t_{\max}}.$$

This proves the lemma. $\square$

Corollary 7.26 (Consistency of quantitative deficiency with valuation drift). *Let $\alpha_* = (\log_2 3 - 1)/\log_2 6$. Suppose that for some choice of L_0 and $t_{\max}$ the residue–graph analysis and Theorem 7.20 produce a deficit parameter $\varepsilon = \varepsilon(L_0, t_{\max})$ such that*

$$\varepsilon(L_0, t_{\max}) \ > \ \alpha_* + \eta + \frac{1}{t_{\max}} \tag{205}$$

for a given $\eta > 0$. Then there exist constants $c > 0$, $C_0 \ge 0$, and the same $t_{\max}$ such that

$$\frac{c\,t_{\max} - C_0}{t_{\max}} \ > \ \alpha_* + \eta,$$

Proof. Apply Theorem 6.13 with the deficit parameter $\varepsilon = \varepsilon(L_0, t_{\max})$ to obtain an inequality of the form (202) for some K_0. Fix $\eta > 0$ and choose $\delta \in \big(0, \varepsilon(L_0, t_{\max})\big)$ so small that

$$\varepsilon(L_0, t_{\max}) - \delta - \frac{1}{t_{\max}} \ > \ \alpha_* + \eta.$$

(This is possible precisely because of the strict inequality (205).) With this choice of δ, define $c := \varepsilon - \delta$ and $C_0 := 1$ as in Lemma 7.25. Then (203) holds for all $1 \le t \le t_{\max}$ and all $k \ge K_0$, and by (204) we have

$$\frac{c\,t_{\max} - C_0}{t_{\max}} = (\varepsilon - \delta) - \frac{1}{t_{\max}} \ > \ \alpha_* + \eta.$$

$\square$

Proposition 7.27 (Quantitative weak non–retreat implies supercritical linear block growth). *Let $(c, C_0, t_{\max}, J_0)$ be constants for which the Quantitative Weak Non–Retreat Principle (6) holds for every infinite forward orbit. Suppose these constants arise from the valuation–drift analysis (Theorem 6.13) and satisfy*

$$\frac{c\,t_{\max} - C_0}{t_{\max}} \ > \ \alpha_*, \qquad \alpha_* := \frac{\log_2 3 - 1}{\log_2 6}.$$

In particular $c\,t_{\max} - C_0 > 0$. Then for every infinite orbit satisfying the Block–Escape Property there exists a constant $\alpha > \alpha_*$ and a subsequence $k_\ell \to \infty$ such that

$$j(n_{k_\ell}) \; \geq \; \alpha\, k_\ell$$

for all sufficiently large ℓ. Consequently, under the above condition on $(c, C_0, t_{\max})$, statement (6) implies the supercritical linear block growth statement (5) of Theorem 1.1.

Proof. Let $(n_k)_{k \geq 0}$ be an infinite forward orbit satisfying the Block–Escape Property. By Theorem 7.20 and the residue–graph construction, there exist parameters L_0 and $t_{\max}$ and a deficit parameter $\varepsilon = \varepsilon(L_0, t_{\max}) > 0$ such that the valuation drift estimate of Theorem 6.13 holds in the form

$$j(n_{k+t}) \; \geq \; j(n_k) + \varepsilon\, t - 1, \qquad 1 \leq t \leq t_{\max}, \tag{206}$$

for all sufficiently large k. Assume that for some $\eta > 0$ the lower bound

$$\varepsilon(L_0, t_{\max}) \; > \; \alpha_* + \eta + \frac{1}{t_{\max}}, \qquad \alpha_* := \frac{\log_2 3 - 1}{\log_2 6}, \tag{207}$$

holds. By Lemma 7.25, for any $\delta \in (0, \varepsilon)$ we may define

$$c := \varepsilon - \delta, \qquad C_0 := 1,$$

and then (206) implies the weak non–retreat inequality

$$j(n_{k+t}) \; \geq \; j(n_k) + c\, t - C_0 \qquad \text{for all } 1 \leq t \leq t_{\max} \tag{208}$$

for all sufficiently large k. Moreover,

$$\frac{c\, t_{\max} - C_0}{t_{\max}} = (\varepsilon - \delta) - \frac{1}{t_{\max}}.$$

Using the strict inequality (207), one may choose $\delta > 0$ small enough so that

$$(\varepsilon - \delta) - \frac{1}{t_{\max}} \; > \; \alpha_*.$$

Thus there exist constants $c > 0$, $C_0 \geq 0$, and the same $t_{\max}$ with

$$\frac{c\, t_{\max} - C_0}{t_{\max}} \; > \; \alpha_*,$$

for which (208) holds for all sufficiently large k.

(1) Constructing the subsequence. Choose k_0 large enough that $j(n_{k_0}) \geq J_0$ and that (208) holds for all $k \geq k_0$ with $j(n_k) \geq J_0$. Define an increasing sequence (k_ℓ) by

$$k_{\ell+1} := k_\ell + t_\ell,$$

where $t_\ell \in [1, t_{\max}]$ is chosen so that (208) holds at $k = k_\ell$. Induction shows that $j(n_{k_\ell}) \geq J_0$ for all ℓ, and hence the construction continues indefinitely with $k_\ell \to \infty$.

(2) Accumulated drift. From (208), for each ℓ,

$$j(n_{k_{\ell+1}}) \; \geq \; j(n_{k_\ell}) + c\, t_\ell - C_0.$$

Summing for $\ell = 0, \ldots, m-1$ and using $k_m = k_0 + \sum_{\ell=0}^{m-1} t_\ell$ gives

$$j(n_{k_m}) \geq j(n_{k_0}) + c\,(k_m - k_0) - mC_0.$$

Since $1 \leq t_\ell \leq t_{\max}$, we have

$$m \geq \frac{k_m - k_0}{t_{\max}},$$

and therefore

$$j(n_{k_m}) \geq j(n_{k_0}) + c\,(k_m - k_0) - \frac{C_0}{t_{\max}}(k_m - k_0).$$

Let

$$\gamma := \frac{c\,t_{\max} - C_0}{t_{\max}}.$$

Then

$$j(n_{k_m}) \geq \gamma\,k_m - C', \qquad C' := \gamma\,k_0 - j(n_{k_0}).$$

(3) Supercritical slope. By assumption,

$$\gamma = \frac{c\,t_{\max} - C_0}{t_{\max}} > \alpha_*.$$

Hence

$$\liminf_{m \to \infty} \frac{j(n_{k_m})}{k_m} \geq \gamma.$$

Choose any α with $\alpha_* < \alpha < \gamma$; for all sufficiently large m,

$$j(n_{k_m}) \geq \alpha\,k_m.$$

Renaming the tail of (k_m) as (k_ℓ) yields the desired subsequence.

Since the orbit satisfies the Block–Escape Property, this linear lower growth of $j(n_k)$ along a subsequence is exactly the supercritical linear block growth required in statement (5) of Theorem 1.1. The implication (6) $\Rightarrow$ (5) follows. $\square$

(8) $\Rightarrow$ (6)

This is Theorem 7.20.

Proposition 7.28 (Spectral hypotheses imply orbitwise discrepancy vanishing). *Assume the spectral hypotheses of Remark 7.6 for the backward Collatz operator P acting on $B_{\text{tree},\sigma}$, together with the forward growth bounds used in Lemma 5.24 to control Cesàro averages along forward orbits. Let $(n_k)_{k\geq 0}$ be any infinite forward Collatz orbit and, for $N \geq 1$, define the Cesàro functionals*

$$\Lambda_N(f) := \frac{1}{N}\sum_{k=0}^{N-1} f(n_k), \qquad f \in B_{\text{tree},\sigma}.$$

Then along any sequence $N_r \to \infty$ there is a subsequence $N_{r_j} \to \infty$ and a functional $\Lambda \in B_{\text{tree},\sigma}^*$ such that $\Lambda_{N_{r_j}} \xrightarrow{\;w^*\;} \Lambda$, and every such weak–$*$ limit satisfies:

1. *Λ is P^*–invariant, that is, $\Lambda(Pf) = \Lambda(f)$ for all $f \in B_{\text{tree},\sigma}$;*
2. *the discrepancy averages vanish along the orbit, in the sense that*

$$\lim_{j\to\infty} \frac{1}{N_{r_j}} \sum_{k=0}^{N_{r_j}-1} D(f)(n_k) = 0, \qquad D(f) := Pf - f \circ T,$$

for every $f \in B_{\text{tree},\sigma}$.

In particular the Orbitwise Discrepancy Vanishing statement (7) *holds with* $A = B_{\text{tree},\sigma}$.

Proof. By Lemma 5.24, which uses the forward growth bounds for Collatz orbits together with the definition of the $B_{\text{tree},\sigma}$–dual norm, the family $(\Lambda_N)_{N \geq 1}$ is uniformly bounded in $B^*_{\text{tree},\sigma}$:

$$\sup_{N \geq 1} \|\Lambda_N\|_{B^*_{\text{tree},\sigma}} < \infty.$$

Hence, by the Banach–Alaoglu theorem, any sequence $N_r \to \infty$ admits a weak–$*$ convergent subsequence $\Lambda_{N_{r_j}} \to \Lambda$ for some $\Lambda \in B^*_{\text{tree},\sigma}$.

We now invoke Lemma 7.4 with $A = B_{\text{tree},\sigma}$, applied to the fixed forward orbit (n_k) and the subsequence (N_{r_j}). That lemma is a purely dynamical statement about the discrepancy operator $D(f) = Pf - f \circ T$ and the Cesàro functionals Λ_N: it asserts that for any weak–$*$ limit Λ of (Λ_{N_r}) along a subsequence, one has

- $\Lambda(Pf) = \Lambda(f)$ for all $f \in B_{\text{tree},\sigma}$, and

-

$$\lim_{j \to \infty} \frac{1}{N_{r_j}} \sum_{k=0}^{N_{r_j}-1} D(f)(n_k) = 0 \qquad \text{for all } f \in B_{\text{tree},\sigma}.$$

In particular, every such weak–$*$ limit is P^*–invariant and satisfies orbitwise discrepancy vanishing for all $f \in B_{\text{tree},\sigma}$, which is exactly statement (7) with $A = B_{\text{tree},\sigma}$. This proves the proposition. $\square$

Proposition 7.29 (Spectral rigidity forbids finite–block trapping). *Assume the spectral hypotheses 7.6 and the spectral rigidity statement of Proposition 7.5, which forbids nonzero P^*–invariant functionals supported on a finite union of blocks. Then there is no infinite forward Collatz orbit $(n_k)_{k \geq 0}$ whose trajectory is contained in a finite union of blocks $\bigcup_{j \leq J} I_j$.*

Proof. Suppose, for contradiction, that there exists an infinite forward orbits $(n_k)_{k \geq 0}$ and an index $J \geq 0$ such that

$$n_k \in B_J := \bigcup_{j=0}^{J} I_j \qquad \text{for all } k \geq 0.$$

Define the Cesàro functionals

$$\Lambda_N(f) := \frac{1}{N} \sum_{k=0}^{N-1} f(n_k), \qquad f \in B_{\text{tree},\sigma}.$$

By Proposition 7.28, along some subsequence $N_{r_j} \to \infty$ the functionals $\Lambda_{N_{r_j}}$ converge weakly–$*$ to a nonzero functional $\Lambda \in B^*_{\text{tree},\sigma}$ that is P^*–invariant:

$$\Lambda(Pf) = \Lambda(f) \qquad \text{for all } f \in B_{\text{tree},\sigma}.$$

On the other hand, the assumption $n_k \in B_J$ for all k forces Λ to be supported on B_J. Indeed, let $f \in B_{\text{tree},\sigma}$ be supported in the complement $B_J^c = \bigcup_{j > J} I_j$. Then $f(n_k) = 0$ for all $k \geq 0$, so

$$\Lambda_N(f) = \frac{1}{N} \sum_{k=0}^{N-1} f(n_k) = 0 \qquad \text{for all } N \geq 1.$$

Passing to the subsequence N_{r_j} and using weak–$*$ convergence, we obtain $\Lambda(f) = 0$ for every f supported on B_J^c, hence Λ is supported entirely on B_J.

Thus Λ is a nonzero P^*–invariant functional supported inside the finite union of blocks B_J, contradicting the rigidity statement of Proposition 7.5. This contradiction shows that no infinite forward orbit can be contained in a finite union of blocks. $\square$

8. Toward a Spectral Calculus for Arithmetic Dynamical Systems

The analytic framework developed here for the backward Collatz operator highlights a broader *spectral calculus* for discrete arithmetic maps. For any map $T : \mathbb{N} \to \mathbb{N}$ with finitely many inverse branches, one may associate the transfer operator

$$(Pf)(n) = \sum_{m:T(m)=n} \frac{f(m)}{w(m)},$$

whose spectral behavior reflects the combinatorial and arithmetic structure of T.

When P acts on weighted sequence spaces such as ℓ^1_σ or on the multiscale tree space $B_{\text{tree},\sigma}$, its Dirichlet transform intertwines

$$\mathcal{D}(Pf)(s) = L_s\,\mathcal{D}(f)(s), \qquad \mathcal{D}(f)(s) = \sum_{n \geq 1} f(n)\, n^{-s},$$

so that spectral information for P passes directly to the analytic continuation and pole structure of the complex family L_s. In this duality, the arithmetic operator P and its analytic realization L_s represent two facets of a single dynamical mechanism: backward iteration in arithmetic space mirrored by analytic continuation in Dirichlet space.

For quasi–compact operators satisfying Lasota–Yorke inequalities on $B_{\text{tree},\sigma}$, one obtains the spectral decomposition

$$P = \sum_{|\lambda_i| > \rho_{\text{ess}}(P)} \lambda_i\, \Pi_i \; + \; N, \qquad \rho_{\text{ess}}(P) < 1,$$

together with the associated dynamical zeta function

$$\zeta_P(s) = \det(I - sP)^{-1} = \exp\!\left(\sum_{k \geq 1} \frac{s^k}{k}\,\mathrm{Tr}(P^k)\right),$$

whose poles coincide with eigenvalues of P outside the essential spectrum and with the resonant singularities of L_s. This creates a coherent analytic picture in which resolvents, spectral projections, Dirichlet envelopes, and dynamical determinants arise as aspects of the same operator geometry.

Beyond the Collatz operator, analogous structures arise for general affine–congruence systems

$$n \longmapsto a_j n + b_j, \qquad a_j, b_j \in \mathbb{N},$$

for which

$$(Pf)(m) = \sum_j \mathbf{1}_{\{m \equiv b_j \ (\mathrm{mod}\ a_j)\}}\, f\!\left(\frac{m - b_j}{a_j}\right).$$

The corresponding Dirichlet transforms L_s act by weighted composition on generating series. A unified spectral calculus would classify such arithmetic systems according to whether their backward operators are quasi–compact, admit meromorphic decompositions, or exhibit a spectral gap on suitable Banach geometries. This analytic classification parallels the dynamical trichotomy into terminating, periodic, and divergent regimes.

In the Collatz case, the results of this paper yield a complete spectral resolution of the backward dynamics. The operator P and its Dirichlet realization L_s together provide a model of an arithmetic transfer operator in which analytic continuation, spectral gaps, and decay of correlations follow from explicit Lasota–Yorke estimates on $B_{\text{tree},\sigma}$. The contraction of L_s for $\Re(s) > 1$, combined with the bound $\lambda_{\text{LY}} < 1$ on $B_{\text{tree},\sigma}$, ensures that P is quasi–compact with a strict spectral gap. Consequently,

the associated dynamical Dirichlet series admit uniform pole–remainder decompositions, and the invariant density exhibits an averaged $1/n$ law: its block averages satisfy

$$c_j \;=\; \frac{1}{6^j}\sum_{n\in I_j} h(n) \;=\; \frac{c}{6^j} + o(6^{-j}),$$

which corresponds to the mass distribution behaving like c/n to leading order on each block I_j.

Boundary spectral geometry and parameter optimization

Theorems 4.17 and 4.1 show that the Lasota–Yorke inequality on B_{tree} yields a strict spectral gap at the boundary $\sigma = 1$. A natural next step is to optimize the parameters (α, ϑ) defining the tree seminorm and to determine whether B_{tree} is minimal or universal among Banach geometries admitting contraction. A quantitative study of

$$\|Pf\|_{\text{tree}} \;\le\; C_P\big(\lambda\,|f|_{\text{tree}} + \|f\|_1\big)$$

may reveal how λ depends on ϑ and how this dependence reflects asymmetries in the Collatz preimage tree. Showing that $\lambda(\vartheta) \to 0$ as $\vartheta \to 0$ would relate analytic contraction rates to the combinatorial entropy of inverse trajectories.

Residues, duality, and forward–backward correspondence

The residue coefficients $A_k(1)$, which decay geometrically as λ^k, represent spectral invariants of the pole part of the dynamical Dirichlet zeta function. On the forward side, the heuristic contraction $(3/4)^k$ describes the average shrinkage of integers under iteration. A precise duality between these quantities would connect analytic and probabilistic aspects of the dynamics, expressing average stopping times and fluctuations in terms of the spectral radius of a normalized backward operator. Such a correspondence would yield a forward–backward conservation principle linking termination statistics with spectral invariants.

Extensions and universality

The multiscale tree space equipped with a hybrid ℓ^1–oscillation norm provides a flexible analytic environment for nonlinear integer maps. Future work may examine metric entropy, measure concentration, and universality phenomena induced by the tree geometry, seeking optimal weights or identifying extremal systems among those with $\lambda < 1$. Such analysis would illuminate how nonlinear arithmetic recursions embed naturally into Banach geometries enforcing global contraction.

Dynamical Dirichlet zeta functions

The functions

$$\zeta_C(s,k) = \sum_{n\ge 1} \frac{1}{(C^k(n))^s}$$

belong to a broader class of *dynamical Dirichlet zeta functions* $\zeta_T(s,k)$ associated with iterates of arithmetic maps with finitely many inverse branches. Spectral gaps govern their meromorphic structure, and residues of their poles capture dynamical invariants. Extending this analysis to more general systems would connect the present framework with Ruelle–Perron–Frobenius theory and the analytic structure of dynamical determinants.

Broader outlook

The spectral resolution of the Collatz dynamics developed here suggests a general *spectral calculus for arithmetic dynamics* in which termination, recurrence, and periodicity correspond to spectral features of noninvertible operators on Banach spaces of arithmetic functions. Future work should clarify how universal the Lasota–Yorke mechanism is for nonlinear arithmetic recursions, how arithmetic symmetries influence spectral gaps, and how probabilistic models of integer iteration emerge as

weak limits of deterministic transfer operators. The Collatz operator studied here provides a detailed example in which a complete spectral description is obtained through explicit Lasota–Yorke estimates on a multiscale Banach space.

References

1. Terras, R. A stopping time problem on the positive integers. *Acta Arith.* **1976**, *30*, 241–252. https://doi.org/10.4064/aa-30-3-241-252.
2. Terras, R. A stopping time problem on the positive integers. II. *Acta Arith.* **1979**, *33*, 241–255. https://doi.org/10.4064/aa-33-3-241-255.
3. Lagarias, J.C. The 3x+1 problem and its generalizations. *Amer. Math. Monthly* **1985**, *92*, 3–23. https://doi.org/10.2307/2322189.
4. Lagarias, J.C. *The Ultimate Challenge: The 3x+1 Problem*; Amer. Math. Soc., 2010.
5. Meinardus, G. Some Analytic Aspects Concerning the Collatz Problem. Technical Report 261, Universität Mannheim, Fakultät für Mathematik und Informatik, 2001.
6. Applegate, D.; Lagarias, J.C. Density bounds for the 3x+1 problem. *Experimental Mathematics* **2005**, *14*, 129–146. https://doi.org/10.1080/10586458.2005.10129133.
7. Ruelle, D. Statistical Mechanics of a One-dimensional Lattice Gas. *Communications in Mathematical Physics* **1968**, *9*, 267–278. https://doi.org/10.1007/BF01645745.
8. Ruelle, D. A Measure Associated with Axiom A Attractors. *American Journal of Mathematics* **1976**, *98*, 619–654. https://doi.org/10.2307/2373862.
9. Leventides, J.; Poulios, C. An operator theoretic approach to the 3x + 1 dynamical system. *IFAC-PapersOnLine* **2021**, *54*, 225–230. 24th International Symposium on Mathematical Theory of Networks and Systems MTNS 2020, https://doi.org/https://doi.org/10.1016/j.ifacol.2021.06.079.
10. Neklyudov, M. Functional analysis approach to the Collatz conjecture. *arXiv preprint arXiv:2106.11859* **2022**.
11. Hennion, H. Sur un théorème spectral et son application aux noyaux lipschitziens. *Proceedings of the American Mathematical Society* **1993**, *118*, 627–634. https://doi.org/10.1090/S0002-9939-1993-1129880-8.
12. Ionescu Tulcea, C.T.; Marinescu, G. Théorie ergodique pour des classes d'opérations non complètement continues. *Annals of Mathematics* **1950**, *52*, 140–147. https://doi.org/10.2307/1969514.
13. Wirsching, G.J. *The Dynamical System Generated by the 3n+1 Function*; Vol. 1681, *Lecture Notes in Mathematics*, Springer, 1998. https://doi.org/10.1007/BFb0098401.
14. Korec, I. A density estimate for the 3x+1 problem. *Math. Slovaca* **1994**, *44*, 85–89.
15. Tao, T. Almost all orbits of the Collatz map attain almost bounded values. *arXiv preprint arXiv:1909.03562* **2019**.
16. Furstenberg, H. *Recurrence in Ergodic Theory and Combinatorial Number Theory*; Princeton Univ. Press, 1981.
17. Katok, A.; Hasselblatt, B. *Introduction to the Modern Theory of Dynamical Systems*; Cambridge Univ. Press, 1995.
18. Krasikov, I. How many numbers satisfy the 3x+1 conjecture? *Internat. J. Math. Math. Sci.* **1989**, *12*, 791–796. https://doi.org/10.1155/S0161171289000949.
19. Kontorovich, A.V.; Lagarias, J.C. Stochastic models for the 3x+1 and 5x+1 problems. *Unif. Distrib. Theory* **2010**, *5*, 121–164.
20. Applegate, D.; Lagarias, J.C. Distribution trees for the 3x+1 problem. *Exp. Math.* **2003**, *12*, 475–490. https://doi.org/10.1080/10586458.2003.10504507.
21. Everett, C.J. Iteration of the number-theoretic function f(2n)=n, f(2n+1)=3n+2. *Adv. Math.* **1977**, *25*, 42–45. https://doi.org/10.1016/0001-8708(77)90004-X.
22. Sinai, Y.G. Statistical (3x+1) problem. *Comm. Pure Appl. Math.* **2003**, *56*, 1016–1028. https://doi.org/10.1002/cpa.10070.
23. Eliahou, S. The 3x+1 problem: new lower bounds on nontrivial cycle lengths. *Discrete Math.* **1993**, *118*, 45–56.