# The Distribution of $v_2(3n+1)$ and Toward a State-Augmented Potential Function for the Collatz Conjecture

AUTHOR
7B7545EB2B5B22A28204066BD292A036
5D4989260318CDF4A7A0407C272E9AFB

April 24, 2025

**Abstract**

The Collatz conjecture posits that iterating the map $T(n) = n/2$ if $n$ is even and $T(n) = 3n + 1$ if $n$ is odd eventually leads to 1 for all positive integers $n$. Heuristic arguments supporting the conjecture often rely on the assumption that the map behaves pseudo-randomly, leading to an expected decrease in magnitude. A key component is the distribution of the 2-adic valuation $v_2(3n + 1)$ for odd $n$. This paper rigorously computes this distribution using natural density, proving the density of odd $n$ with $v_2(3n + 1) = k$ is $2^{-k}$ for each $k \geq 1$. This confirms $E[v_2(3n + 1)] = 2$, providing a formal basis for the heuristic downward drift argument ($2 > \log_2 3$). We discuss implications and limitations of density results. Motivated by this and the shortcomings of static potential functions, we propose a novel state-augmented potential function framework, incorporating trajectory history via the previous step's valuation ($k_{prev}$), to better model the effects of transient carry-bit dynamics and overcome limitations of static functions.

## 1 Introduction

The Collatz conjecture, also known as the $3n + 1$ problem, concerns the iteration of the function $T : \mathbb{N} \to \mathbb{N}$ defined by

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

The conjecture asserts that for every starting integer $n \geq 1$, the sequence of iterates $n, T(n), T(T(n)), \ldots$ eventually reaches the cycle $4 \to 2 \to 1$. Despite its simple statement, the conjecture remains unproven and is famously difficult [3].

Much intuition stems from heuristic arguments suggesting iterates decrease "on average". A crucial element is the map $T_{\text{odd}} : \mathcal{O} \to \mathcal{O}$ for odd $n$:

$$T_{\text{odd}}(n) = \frac{3n+1}{2^{v_2(3n+1)}}, \tag{1}$$

where $v_2(m)$ is the 2-adic valuation. The magnitude change is roughly $\log_2(3) - v_2(3n+1)$. Heuristics often assume $v_2(3n+1)$ follows $P(k) \approx 2^{-k}$, leading to $E[v_2(3n+1)] \approx 2$. Since $2 > \log_2(3) \approx 1.58$, this suggests an average decrease, supported by computational evidence and results on density [5] and boundedness [4].

The first goal here is to provide a rigorous basis for this heuristic by computing the exact natural density distribution of $v_2(3n+1)$.

Our main result is:

**Theorem 1.1.** *Let $\mathcal{O} = \{1, 3, 5, \dots\}$. For each integer $k \geq 1$, the natural density of $S_k = \{n \in \mathcal{O} \mid v_2(3n+1) = k\}$ within $\mathcal{O}$ is $2^{-k}$.*

This confirms $E[v_2(3n+1)] = 2$ under this measure. Yet, density results fall short of universal proof, and static potential functions struggle with the discontinuous dynamics [Lagarias2010]. Thus, recognizing these limitations, we propose a novel state-augmented framework for constructing a potential function, aiming to capture the influence of transient carry-bit dynamics more effectively.

## 2 Preliminaries

**Definition 2.1** (2-adic Valuation)**.** *For $m \in \mathbb{Z} \setminus \{0\}$, $v_2(m)$ is the exponent of the highest power of 2 dividing $m$. $v_2(0) = \infty$.*

**Definition 2.2** (Natural Density on Odd Integers)**.** *For $S \subseteq \mathcal{O}$, $\delta(S) = \lim_{m\to\infty} |\{n \in S \mid n \leq 2^m - 1\}|/2^{m-1}$, if it exists.*

## 3 Distribution of $v_2(3n+1)$: Proof of Theorem 1.1

**Lemma 3.1.** *For odd $n$, $v_2(3n+1) = k \iff 3n \equiv -1 \pmod{2^k}$ and $3n \not\equiv -1 \pmod{2^{k+1}}$.*

*Proof.* Equivalent to $3n + 1 \equiv 2^k \pmod{2^{k+1}}$. □

**Lemma 3.2.** *For $k \geq 1$, $a_k = -3^{-1} \pmod{2^k}$ exists, is unique, and is odd.*

*Proof.* Inverse exists as $\gcd(3, 2^k) = 1$. If $3n \equiv -1 \pmod{2^k}$, then $\gcd(n, 2^k) = 1$, so $n$ is odd. □

### 3.1 Examples for k=1, k=2, and k=3

We summarize the density calculation for small $k$:

*Example* 3.3 (Case k=1). Condition $n \equiv 3 \pmod 4$. Proportion = 1/2.

*Example* 3.4 (Case k=2). Condition $n \equiv 1 \pmod 8$. Proportion = 1/4.

*Example* 3.5 (Case k=3). Condition $n \equiv 13 \pmod{16}$. Proportion = 1/8.

### 3.2 General Proof

*Proof of Theorem 1.1.* Let $a_k = -3^{-1} \pmod{2^k}$ and $a_{k+1} = -3^{-1} \pmod{2^{k+1}}$. The condition $v_2(3n+1) = k$ is $n \equiv a_k \pmod{2^k}$ and $n \not\equiv a_{k+1} \pmod{2^{k+1}}$. The condition $n \equiv a_k \pmod{2^k}$ defines two odd classes mod $2^{k+1}$: $a_{k+1}$ and $a'_k = a_{k+1} + 2^k \pmod{2^{k+1}}$. The class $a_{k+1}$ yields $v_2 \geq k+1$; the class $a'_k$ yields $v_2 = k$. Thus, $v_2(3n+1) = k$ is equivalent to $n \equiv a'_k \pmod{2^{k+1}}$, a single odd class. For $m \geq k+1$, this condition defines $2^{m-(k+1)}$ classes among the $2^{m-1}$ odd classes modulo $2^m$. The proportion is $2^{m-k-1}/2^{m-1} = 2^{-k}$. This is independent of $m$ for $m \geq k+1$. The natural density $\delta(S_k) = 2^{-k}$. □

*Remark* 3.6 (Finite Moduli Behavior). The proportion is exact for moduli $2^m$ with $m \geq k+1$.

*Remark* 3.7 (Density vs. Finite Intervals). Density over $[1, X]$ might show deviations.

## 4 Discussion and Implications

### 4.1 Expected Value of $v_2(3n+1)$

**Corollary 4.1.** $E[v_2(3n+1)] = 2$ *under the natural density measure.*

*Proof.* $\sum_{k=1}^{\infty} k \cdot \delta(S_k) = \sum_{k=1}^{\infty} k \cdot 2^{-k} = 2$. □

Supports heuristic average decrease $\approx \log_2 3 - 2 \approx -0.415$ bits/odd step.

### 4.2 Implications for Collatz Approaches

- Supports pseudo-random / ergodic models [5, 4].
- Highlights importance of carry dynamics [6].
- Provides statistical constraints against cycles.

### 4.3 2-adic Perspective

Arises naturally from Haar measure on $\mathbb{Z}_2^*$ [2]. $T_{\text{odd}}$ is measure-preserving. Non-zero integral of $v_2(3x+1) - \log_2 3$ shows it is not a coboundary, supporting ergodic negative drift $\mu$-a.e. Bridging to $\mathbb{N}^+$ remains open.

### 4.4 Limitations

Density results don't prove universal convergence [1]. Static potential functions fail, often due to insufficient state information [3].

## 5 Toward a State-Augmented Potential Function

The limitations motivate exploring potential functions incorporating more state, aiming to model carry effects. We propose a **state-augmented (or history-augmented) potential function** framework.

**Definition 5.1** (Augmented State)**.** *For $n \in \mathcal{O}, n > 1$, if $n \neq n_0$, let $n = T_{odd}(n_{prev})$. The state includes $n$ and $k_{prev} = v_2(3n_{prev} + 1)$. For the initial $n_0$, a convention is needed; for simplicity, we suggest setting $k_{prev} = 2$ (the expected value) as a default, although optimal initialization might depend on $n_0 \pmod{2^J}$ and requires further analysis.*

**Definition 5.2** (Hypothetical State-Augmented Potential Function Form)**.** *We hypothesize $f$ depending on this state, possibly:*

$$f(n, k_{prev}) = \log_2 n + \Psi(n \pmod{2^J}, k_{prev})$$

*for large $J$, with $f(1, \cdot) = 0$.*

The Lyapunov requirement is: for $n \to n' = (3n + 1)/2^k$ with $k = v_2(3n + 1)$,

$$f(n', k) \leq f(n, k_{prev}) \tag{2}$$

implying the condition on $\Psi$:

$$\Psi(n \pmod{2^J}, k_{prev}) - \Psi(n' \pmod{2^J}, k) \geq \log_2(3 + 1/n) - k$$

**Illustrative Example of $\Psi$:** The simple form $\Psi(a, k_{prev}) = C - c \cdot k_{prev}$ fails. A viable $\Psi$ likely requires more complex dependence on its arguments. Future work might explore $\Psi$ with terms like $c \cdot k_{prev} \cdot \phi(n \pmod{2^J})$, where $\phi$ might weight residue classes based on their influence on subsequent carry dynamics or 2-adic properties.

**Potential Interpretation and Advantages:** Allows potential $\Psi$ to depend explicitly on the previous step's outcome ($k_{prev}$), potentially tracking "energy" storage/release better than static functions.

**Challenges and Novelty:** Defining and analyzing $\Psi$ is the main challenge. Proving inequality (2) universally is the goal. Incorporating history via $k_{prev}$ is a **novel conceptual direction**, addressing limitations of static functions by accounting for transient carry outcomes.

## 6 Conclusion

We rigorously established the $2^{-k}$ natural density for $v_2(3n+1) = k$, confirming $E[v_2(3n+1)] = 2$. This grounds heuristic arguments but doesn't resolve the conjecture. The failure of standard potential functions underscores the need to model the discontinuous carry dynamics more accurately.

We propose a novel framework exploring state-augmented potential functions, $f(n, k_{prev})$, incorporating memory via $k_{prev}$. This approach aims to model the system's dynamics more faithfully than static functions. While constructing and validating such a function is challenging, this direction offers a potential path forward, meriting further investigation.

## References


[1] C. J. Everett, Iteration of the number-theoretic function f(2n) = n, f(2n+1) = 3n+2. *Advances in Mathematics*, **25**(1) (1977), pp. 42–45.

[2] F. Q. Gouvêa, *p-adic Numbers: An Introduction*. Second Edition, Springer, 1997.

[3] J. C. Lagarias, ed., *The Ultimate Challenge: The 3x+1 Problem*. American Mathematical Society, 2010.

[4] T. Tao, Almost all orbits of the Collatz map attain almost bounded values. *Forum of Mathematics, Pi*, **10** (2022), e1.

[5] R. Terras, A stopping time problem on the positive integers. *Acta Arithmetica*, **30** (1976), pp. 241–252.

[6] G. J. Wirsching, *The Dynamical System Generated by the 3n+1 Function*. Lecture Notes in Mathematics 1681, Springer, 1998.