Collatz Conjecture

William F. Barnes

August 17, 2020

Introduction

In 1937, Lothar Collatz pointed out a pattern followed by (seemingly) all positive integers. Start with any integer $n > 1$. If n is even, change n to $n/2$. If n is odd, change n to $3n+1$. Repeat this until n changes to 1. The so-called *Collatz conjecture* states that any integer $n > 1$ will eventually reduce down to 1.

Any acceptable proof the Collatz conjecture has remained elusive to the most accomplished mathematicians. Paul Erdõs himself conceded that 'mathematics may not be ready for such problems', while others have speculated that a proof, if one exists, cannot be built from standard mathematical axioms. The problem has nonetheless been explored in several directions and has gained a slew of nicknames along the way, namely (but not limited to) the $3n+1$ problem, the hailstone sequence, the hailstone numbers, and the wondrous numbers.

Common literature is ubiquitous with comments and conclusions on the Collatz conjecture, thus none are purposely repeated here. Instead, following is the summary of back-of-the envelope notes that capture an exploration of the problem.

1 Data

Let us write each positive integer $\{n\} = 1, 2, 3, \ldots$, and apply the recursive rule

$$n \to \begin{cases} \dfrac{n}{2} & n \text{ even} \\ 3n+1 & n \text{ odd} \end{cases}$$

until $n \to 1$. Stopping at $n = 15$, a table

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
	1	10	2	16	3	22	4	28	5	34	6	40	7	46
		5	1	8	10	11	2	14	16	17	3	20	22	23
		16		4	5	34	1	7	8	52	10	10	11	70
		8		2	16	17		22	4	26	5	5	34	35
		4		1	8	52		11	2	13	16	16	17	106
		2			4	26		34	1	40	8	8	52	53
		1			2	13		17		20	4	4	26	160
					1	40		52		10	2	2	13	80
						20		26		5	1	1	40	40
						10		13		16			20	20
						5		40		8			10	10
						.		.		.			.	.

emerges, where columns that end with a dot (.) continue to a sub-sequence already written elsewhere in the table (merely for saving space).

Sequence Notation

In the table above, each column contains a sequence $\phi(n, k)$ that starts at number n and ends at number k. The simplest meaningful case starts with $n = 2$, meaning $\phi(2, 1) = (2, 1)$, whereas for $n = 3$ we have $\phi(3,1) = (3, 10, 5, 16, 8, 4, 2, 1)$. Let us denote the 'width' of a sequence, i.e. the number of elements, as $\tilde{n} + 1$, where $\tilde{n}$ is the number of jumps required to reach 1 from the base number n.

For certain base integers, we observe that the resulting sequence may start repeating that of a previous base integer. Taking $n = 6$ for example, we find

$$\begin{aligned}\phi(6,1) &= (6, 3, 10, 5, 16, 8, 4, 2, 1) \\ &= (6, 3, 10)\ (5, 16, 8, 4, 2, 1) \\ &= \phi(6,10)\ \phi(5,1)\ .\end{aligned}$$

Of course, the same exercise can be repeated on $\phi(5,1)$, all the way down to $\phi(2,1)$. In the general case, sequences that satisfy the Collatz conjecture obey

$$\phi(n,1) = \phi(n,x)\ \phi(j,k)\ \phi(y,1)\ ,$$

where y must follow from x across a known sequence $\phi(j,k)$.

2 Progress Operator

Let us construct the *progress* operator $f(n)$ that applies a single step of the transformation rule

$$f(n) = \begin{cases} \dfrac{n}{2} & n \text{ even} \\ 3n+1 & n \text{ odd} \end{cases}$$

(i.e. non-recursively) to any integer $n > 0$, generating a table:

$f(1) = 4$	$f(16) = 8$	$f(31) = 94$
$f(2) = 1$	$f(17) = 52$	$f(32) = 16$
$f(3) = 10$	$f(18) = 9$	$f(33) = 100$
$f(4) = 2$	$f(19) = 58$	$f(34) = 17$
$f(5) = 16$	$f(20) = 10$	$f(35) = 106$
$f(6) = 3$	$f(21) = 44$	$f(36) = 17$
$f(7) = 22$	$f(22) = 11$	$f(37) = 112$
$f(8) = 4$	$f(23) = 70$	$f(38) = 19$
$f(9) = 28$	$f(24) = 12$	$f(39) = 118$
$f(10) = 5$	$f(25) = 76$	$f(40) = 20$
$f(11) = 34$	$f(26) = 13$	$f(41) = 124$
$f(12) = 6$	$f(27) = 82$	$f(42) = 21$
$f(13) = 40$	$f(28) = 14$	$f(43) = 130$
$f(14) = 7$	$f(29) = 88$	$f(44) = 22$
$f(15) = 46$	$f(30) = 15$	$f(45) = 136$

Also denote the *compound progress* operator $f_k(n)$ as the progress operator applied k times to a base number n such that

$$f_2(n) = f(f(n)) \qquad\qquad f_k(n) = f(f(f(\cdots k \cdots (f(n)))))\ .$$

Collatz Condition

In the language of the progress operator, any integers that satisfy the Collatz conjecture must satisfy the condition

$$f_{\tilde{n}}(n) = f(f(f(\cdots \tilde{n} \cdots (f(n))))) = 1 \,,$$

where the number of iterations equals the number of forward jumps $\tilde{n}$ needed to traverse the sequence $\phi(n, 1)$.

3 Regress Operator

To accompany the progress operator $f(n)$, we have grounds to define its inverse called the *regress* operator $g(n)$ such that

$$g(f(n)) = \{n\} \,,$$

where the right side of $g(n)$ is multi-valued for certain n. That is, the operator $g(n)$ asks 'which number(s) could have brought us to n?' A partial answer is in the following table:

$g(1) = 2$	$g(16) = 32, 5$	$g(31) = 62$
$g(2) = 4$	$g(17) = 34$	$g(32) = 64$
$g(3) = 6$	$g(18) = 36$	$g(33) = 66$
$g(4) = 8, 1$	$g(19) = 38$	$g(34) = 68, 11$
$g(5) = 10$	$g(20) = 40$	$g(35) = 70$
$g(6) = 12$	$g(21) = 42$	$g(36) = 72$
$g(7) = 14$	$g(22) = 44, 7$	$g(37) = 74$
$g(8) = 16$	$g(23) = 46$	$g(38) = 76$
$g(9) = 18$	$g(24) = 48$	$g(39) = 78$
$g(10) = 20, 3$	$g(25) = 50$	$g(40) = 80, 13$
$g(11) = 22$	$g(26) = 52$	$g(41) = 82$
$g(12) = 24$	$g(27) = 54$	$g(42) = 84$
$g(13) = 26$	$g(28) = 56, 9$	$g(43) = 86$
$g(14) = 28$	$g(29) = 58$	$g(44) = 88$
$g(15) = 30$	$g(30) = 60$	$g(45) = 90$

There is clearly more 'order' in the $g(n)$-table as compared to the previous $f(n)$-table. Each $n > 0$ has at least one solution $2n$, accounting for all even-number results of $g(n)$. The odd-number results of $g(n)$ occur as second solutions to the cases 4, 10, 16, 22, 28, 34, and so on. Evidently, all operations $g(4 + 6j)$ for integers $j = 0, 1, 2, \ldots$ are multi-valued:

$$g(4 + 6j) = \begin{cases} 2(4 + 6j) \\ 1 + 2j \end{cases} \qquad j = 0, 1, 2, 3, \ldots$$

Let us finally denote the *compound regress* operator $g_k(n)$ as the regress operator applied k times to a base number n such that

$$g_2(n) = g(g(n)) \qquad\qquad g_k(n) = g(g(g(\cdots k \cdots (g(n))))) \,.$$

Note that the right hand side of $g(n)$ is generally multi-valued, giving rise to a tree-like right hand side of $g_k(n)$.

Regress Tree

Now, let us write the Collatz condition

$$f_{\tilde{n}}(n) = 1 \; ,$$

and apply the regress operator $g(\,)$ to each side. On the left, we have $g(f_{\tilde{n}}(n)) = f_{\tilde{n}-1}(n)$, and meanwhile $g(1)$ is produced on the right, giving us

$$f_{\tilde{n}-1}(n) = g(1) = 2 \; .$$

Apply $g(\,)$ again to get another statement

$$f_{\tilde{n}-2}(n) = g(g(1)) = g_2(1) = 4 \; ,$$

and again for yet another

$$f_{\tilde{n}-3}(n) = g(g(g(1))) = g_3(1) = 8 \; ,$$

and again for $\tilde{n}-4$:

$$f_{\tilde{n}-4}(n) = g_4(1) = 16$$

Proceeding carefully for the $\tilde{n}-5$ case, we get *two* possible results

$$f_{\tilde{n}-5}(n) = g_5(1) = g(16) = \begin{cases} 32 \\ 5 \end{cases} .$$

Applying $g(\,)$ to the left and both items on the right, we have

$$f_{\tilde{n}-6}(n) = g_6(1) = \begin{cases} g(32) \\ g(5) \end{cases} = \begin{cases} 64 \\ 10 \end{cases} ,$$

and such a pattern can continue. Doing so, we produce a *regress tree*. In the following equations, each elongated brace symbol represents a multi-valued result in $g_k(1)$:

$$f_{\tilde{n}-7}(n) = g_7(1) = \begin{cases} g(64) \\ g(10) \end{cases} = \begin{cases} \begin{cases} 128 \\ 21 \end{cases} \\ \begin{cases} 20 \\ 3 \end{cases} \end{cases}$$

$$f_{\tilde{n}-8}(n) = g_8(1) = \begin{cases} \begin{cases} g(128) \\ g(21) \end{cases} \\ \begin{cases} g(20) \\ g(3) \end{cases} \end{cases} = \begin{cases} \begin{cases} 256 \\ 42 \end{cases} \\ \begin{cases} 40 \\ 6 \end{cases} \end{cases}$$

$$f_{\tilde{n}-9}\left(n\right)=g_{9}\left(1\right)=\begin{cases}\begin{cases}g\left(256\right)\\ g\left(42\right)\end{cases}\\ \begin{cases}g\left(40\right)\\ g\left(6\right)\end{cases}\end{cases}=\begin{cases}\begin{cases}\begin{cases}512\\ 85\end{cases}\\ 84\end{cases}\\ \begin{cases}\begin{cases}80\\ 13\end{cases}\\ 12\end{cases}\end{cases}$$

$$f_{\tilde{n}-10}\left(n\right)=g_{10}\left(1\right)=\begin{cases}\begin{cases}\begin{cases}g\left(512\right)\\ g\left(85\right)\end{cases}\\ g\left(84\right)\end{cases}\\ \begin{cases}\begin{cases}g\left(80\right)\\ g\left(13\right)\end{cases}\\ g\left(12\right)\end{cases}\end{cases}=\begin{cases}\begin{cases}\begin{cases}1024\\ 170\end{cases}\\ 168\end{cases}\\ \begin{cases}\begin{cases}160\\ 26\end{cases}\\ 24\end{cases}\end{cases}$$

$$f_{\tilde{n}-11}\left(n\right)=g_{11}\left(1\right)=\begin{cases}\begin{cases}\begin{cases}g\left(1024\right)\\ g\left(170\right)\end{cases}\\ g\left(168\right)\end{cases}\\ \begin{cases}\begin{cases}g\left(160\right)\\ g\left(26\right)\end{cases}\\ g\left(24\right)\end{cases}\end{cases}=\begin{cases}\begin{cases}\begin{cases}\begin{cases}2048\\ 341\end{cases}\\ 340\end{cases}\\ 336\end{cases}\\ \begin{cases}\begin{cases}\begin{cases}320\\ 53\end{cases}\\ 52\end{cases}\\ 48\end{cases}\end{cases}$$

$$f_{\tilde{n}-12}\left(n\right)=g_{12}\left(1\right)=\begin{cases}\begin{cases}\begin{cases}\begin{cases}g\left(2048\right)\\ g\left(341\right)\end{cases}\\ g\left(340\right)\end{cases}\\ g\left(336\right)\end{cases}\\ \begin{cases}\begin{cases}\begin{cases}g\left(320\right)\\ g\left(53\right)\end{cases}\\ g\left(52\right)\end{cases}\\ g\left(48\right)\end{cases}\end{cases}=\begin{cases}\begin{cases}\begin{cases}\begin{cases}4096\\ 682\end{cases}\\ \begin{cases}680\\ 113\end{cases}\end{cases}\\ 672\end{cases}\\ \begin{cases}\begin{cases}\begin{cases}640\\ 106\end{cases}\\ \begin{cases}104\\ 17\end{cases}\end{cases}\\ 96\end{cases}\end{cases}$$

$$f_{\tilde{n}-13}(n)=g_{13}(1)=\begin{cases}\begin{cases}\begin{cases}\begin{cases}g(4096)\\ g(682)\end{cases}\\ \begin{cases}g(680)\\ g(113)\end{cases}\end{cases}\\ g(672)\end{cases}\\ \begin{cases}\begin{cases}\begin{cases}g(640)\\ g(106)\end{cases}\\ \begin{cases}g(104)\\ g(17)\end{cases}\end{cases}\\ g(96)\end{cases}\end{cases}=\begin{cases}\begin{cases}\begin{cases}\begin{cases}\begin{cases}8192\\ 1365\end{cases}\\ \begin{cases}1364\\ 227\end{cases}\end{cases}\\ \begin{cases}1360\\ 226\end{cases}\end{cases}\\ 1344\end{cases}\\ \begin{cases}\begin{cases}\begin{cases}\begin{cases}1280\\ 213\end{cases}\\ \begin{cases}212\\ 35\end{cases}\end{cases}\\ \begin{cases}208\\ 34\end{cases}\end{cases}\\ 192\end{cases}\end{cases}$$

$$f_{\tilde{n}-14}(n)=g_{14}(1)=\begin{cases}\begin{cases}\begin{cases}\begin{cases}\begin{cases}g(8192)\\ g(1365)\end{cases}\\ \begin{cases}g(1364)\\ g(227)\end{cases}\end{cases}\\ \begin{cases}g(1360)\\ g(226)\end{cases}\end{cases}\\ g(1344)\end{cases}\\ \begin{cases}\begin{cases}\begin{cases}\begin{cases}g(1280)\\ g(213)\end{cases}\\ \begin{cases}g(212)\\ g(35)\end{cases}\end{cases}\\ \begin{cases}g(208)\\ g(34)\end{cases}\end{cases}\\ g(192)\end{cases}\end{cases}=\begin{cases}\begin{cases}\begin{cases}\begin{cases}\begin{cases}16384\\ 2730\end{cases}\\ \begin{cases}2728\\ 454\end{cases}\end{cases}\\ \begin{cases}\begin{cases}2720\\ 453\end{cases}\\ \begin{cases}452\\ 75\end{cases}\end{cases}\end{cases}\\ 2688\end{cases}\\ \begin{cases}\begin{cases}\begin{cases}\begin{cases}2560\\ 426\end{cases}\\ \begin{cases}424\\ 70\end{cases}\end{cases}\\ \begin{cases}\begin{cases}416\\ 69\end{cases}\\ \begin{cases}68\\ 11\end{cases}\end{cases}\end{cases}\\ 384\end{cases}\end{cases}$$

Condensing notation, let us flatten each tree $g_{k=\tilde{n}}(1)$ into a sequence $\psi(\tilde{n})$:

$$
\begin{aligned}
\psi(0) &= (1) \\
\psi(1) &= (2) \\
\psi(2) &= (4) \\
\psi(3) &= (8) \\
\psi(4) &= (16) \\
\psi(5) &= (32, 5) \\
\psi(6) &= (64, 10) \\
\psi(7) &= (128, 21, 20, 3) \\
\psi(8) &= (256, 42, 40, 6) \\
\psi(9) &= (512, 85, 84, 80, 13, 12) \\
\psi(10) &= (1024, 170, 168, 160, 26, 24) \\
\psi(11) &= (2048, 341, 340, 336, 320, 54, 52, 48) \\
\psi(12) &= (4096, 682, 680, 113, 672, 640, 106, 104, 17, 96) \\
\psi(13) &= (8192, 1365, 1364, 227, 1360, 226, 1344, 1280, 213, 212, 35, 208, 34, 192) \\
&\cdots
\end{aligned}
$$

As constructed, any given sequence $\psi(\tilde{n})$ contains all of the integers that are $\tilde{n}$ jumps from one. Since no integer n can have two different 'jump numbers' $\tilde{n}$ and $\tilde{n}' \neq \tilde{n}$, no integer occurs more than once throughout the tree.

Tree Analysis

As we've seen, the operation $g(\psi(\tilde{n}))$ is used to calculate $\psi(\tilde{n}+1)$, and the number of elements per sequence never decreases as $\tilde{n}$ increases. (There are thus no empty branches.)As an informal exercise, we note that, if handed a random result of $g(n)$, the limit probability of the result being even versus odd is

$$\lim_{N \gg 0} \frac{N_{\text{odd}}}{N} \approx \frac{1}{7} \approx 0.143 \qquad \lim_{N \gg 0} \frac{N_{\text{even}}}{N} \approx \frac{6}{7} \approx 0.857\,.$$

Testing this, the count of even and odd occurrences included within all sequences $\psi(0)$, $\cdots$, $\psi(14)$ results in 79 total elements, comprising of 17 odd integers and 62 even integers. Calculating the ratio of each to the total, we find

$$N_{\text{odd}} = \frac{17}{79} = 0.218 \qquad N_{\text{even}} = \frac{62}{79} = 0.795\,,$$

in rough agreement with the above. Without requiring said ratios to represent all sequences $\psi(\tilde{n})$, it should generally follow that sequences with $\tilde{n} \gg 1$ contain roughly six even integers for every odd integer. Any odd integer n contained in $\psi(\tilde{n})$ becomes an even integer in $\psi(\tilde{n}+1)$, which is ultimately balanced by new odd integers emerging elsewhere in the tree.

Treating each element n in $\psi(\tilde{n})$ as a pseudo-random integer, it follows that each updated element $g(n)$ in $\psi(\tilde{n}+1)$ has a rough $1/6$ probability being multi-valued. Denoting $N(\tilde{n})$ as the number of elements in the sequence $\psi(\tilde{n})$, we should have, in the limit of large $\tilde{n}$,

$$N(\tilde{n}+1) \approx N(\tilde{n})\left(1+\frac{1}{6}\right) \qquad \rightarrow \qquad \frac{N(\tilde{n}+1)-N(\tilde{n})}{N(\tilde{n})} \approx \frac{1}{6} \,,$$

implying exponential growth in $N(\tilde{n})$,

$$N(\tilde{n}) \approx \exp(\tilde{n}/6) \,,$$

where the initial value $N(\tilde{n}=0)$ corresponds to one.

To proceed, condense all $\psi(\tilde{n})$ into the grand sequence

$$\Psi(\tilde{n}) = \psi(0) \cup \psi(1) \cdots \psi(\tilde{n}-1) \cup \psi(\tilde{n}) \,.$$

That is, $\Psi(\tilde{n})$ contains one instance of every integer that progresses to one in $\tilde{n}$ jumps or less. Also, let $T(\tilde{n})$ equal the total number of elements in $\Psi(\tilde{n})$. It quickly follows that

$$T(\tilde{n}) = \sum_{j=0}^{\tilde{n}} N(j) \approx \sum_{j=0}^{\tilde{n}} \exp(j/6) \,.$$

In the large-$\tilde{n}$ regime, we may further approximate the sum of exponential terms as a continuous integral in j having step size dj:

$$T(\tilde{n}) \approx \int_0^{\tilde{n}} \exp(j/6)\, dj = 6\, \exp(j/6)\Big|_0^{\tilde{n}} \approx 6\, \exp(\tilde{n}/6)$$

Evidently, an integer that satisfies the Collatz conjecture after $\tilde{n}$ jumps is a single element in a sequence $\psi(\tilde{n})$ containing approximately $\exp(\tilde{n}/6)$ elements. For a given number of jumps $\tilde{n}$, the total count of integers that satisfy the Collatz conjecture is approximately $6\exp(\tilde{n}/6)$.

4 Base-Odd Lattice

It's easy to see that odd integers are the star players in the problem, as all even integers eventually collapse down to an odd integer $\{\alpha_j\}$. Rewriting the whole set of positive integers in this light, we can write out a *base-odd lattice*, growing downward in multiples of $2^\lambda \alpha_j$, where $\lambda = 1, 2, 3, \ldots$, as follows (stopping at 10 rows, 9 columns):

1	*3*	*5*	*7*	*9*	*11*	*13*	15	*17*
2	6	**10**	14	18	**22**	26	30	**34**
4	12	20	**28**	36	44	**52**	60	68
8	24	**40**	56	72	88	104	120	136
16	48	80	112	144	176	208	240	272
32	96	160	224	288	352	416	480	544
64	192	320	448	576	704	832	960	1088
128	384	640	896	1152	1408	1664	1920	2176
256	768	1280	1792	2304	2816	3328	3840	4352
512	1536	2560	3584	4608	5632	6656	7680	8704

Using such a lattice, Collatz sequences can be easily visualized: any process $n \to 3n+1$ will land n in the field of even integers, which then 'boils upward' to an odd integer in the top row. In the above, italicized odd integers lead to boldface evens in nearby columns.

As it turns out, a complete base-odd lattice has (nearly) every second bold entry corresponding to *some* odd number. For instance, jumping down the leftmost column to 16, 64, 256, 1024, 4096, etc., it follows that each original $n \to 3n+1$ had to be 5, 31, 85, 341, 1365, etc. Continuing in the space allotted, we find a pattern in the boldface even integers that is conspicuously absent in any column whose elements are divisible by three:

1	3	5	7	9	11	13	15	17
2	6	**10**	14	18	**22**	26	30	**34**
4	12	20	**28**	36	44	**52**	60	68
8	24	**40**	56	72	**88**	104	120	**136**
16	48	80	**112**	144	176	**208**	240	272
32	96	**160**	224	288	**352**	416	480	**544**
64	192	320	**448**	576	704	**832**	960	1088
128	384	**640**	896	1152	**1408**	1664	1920	**2176**
256	768	1280	**1792**	2304	2816	**3328**	3840	4352
512	1536	**2560**	3584	4608	**5632**	6656	7680	**8704**

An odd integer α_j in the top row sits on an infinite stack of even numbers $\left\{\alpha_j 2^\lambda\right\}$. For the columns not divisible by three, the boldface evens could come from some *different* odd number $\alpha_{k \neq j}$ for each power of λ.

We deduce that any integer in the top row that satisfies the Collatz conjecture 'brings along' a column of subsequent odd integers that also satisfy the Collatz conjecture (excluding columns divisible by three). Reading down the first column, the odd integers $\{\beta_j\}$ implied by the bold evens are given by

$$\beta_{1,k} = \frac{1 \cdot 2^{2k} - 1}{3} \qquad\qquad k = 1, 2, 3, \ldots\ ,$$

where the next nontrivial column is begins with 5 and follows with

$$\beta_{5,k} = \frac{5 \cdot 2^{2k-1} - 1}{3} \qquad\qquad k = 1, 2, 3, \ldots\ .$$

In the same notation, note that $\beta_{3,k}$ contains nothing. These patterns repeat every three columns:

$$\begin{aligned}
\beta_{1+6j,k} &= \frac{(1+6j)\cdot 2^{2k}-1}{3} \qquad j=0,1,2,\ldots \qquad k=1,2,3,\ldots \\
\beta_{3+6j,k} &= \emptyset \\
\beta_{5+6j,k} &= \frac{(5+6j)\cdot 2^{2k-1}-1}{3} \qquad j=0,1,2,\ldots \qquad k=1,2,3,\ldots
\end{aligned}$$

Explicitly, this means if $n=1$ satisfies the Collatz conjecture, then so does $(5, 21, 85, \ldots)$. Similarly, if $n=5$ satisfies the Collatz conjecture, then so does $(3, 13, 53, \ldots)$, and so on:

$$\begin{aligned}
\{\beta_1\} &= (\mathit{1}, 5, \mathbf{21}, 85, 341, \mathbf{1365}, 5461, 21845, \mathbf{87381} \ldots) \\
\{\beta_5\} &= (\mathbf{3}, 13, 53, \mathbf{213}, 853, 3413, \mathbf{13653}, 54613, \ldots) \\
\{\beta_7\} &= (\mathbf{9}, 37, 149, \mathbf{597}, 2389, 9557, \mathbf{38229}, 152917, \ldots) \\
\{\beta_{11}\} &= (7, 29, \mathbf{117}, 469, 1877, \mathbf{7509}, 30037, 120149, \ldots) \\
\{\beta_{13}\} &= (17, \mathbf{69}, 277, 1109, \mathbf{4437}, 11749, 70997, \mathbf{283989}, \ldots) \\
\{\beta_{17}\} &= (11, \mathbf{45}, 181, 725, \mathbf{2901}, 11605, 46421, \mathbf{185685}, \ldots) \\
\{\beta_{19}\} &= (25, 101, \mathbf{405}, 1621, 6485, \mathbf{25941}, 103765, 415061, \ldots) \\
\{\beta_{23}\} &= (\mathbf{15}, 61, 245, \mathbf{981}, 3925, 15701, \mathbf{62805}, 251221, \ldots) \\
\{\beta_{25}\} &= (\mathbf{33}, 133, 533, \mathbf{2133}, 8533, 34133, \mathbf{136533}, 546133, \ldots) \\
\{\beta_{29}\} &= (19, 77, \mathbf{309}, 1237, 4949, \mathbf{19797}, 79189, 316757, \ldots) \\
\{\beta_{31}\} &= (41, \mathbf{165}, 661, 2645, \mathbf{10581}, 42325, 169301, \mathbf{677205}, \ldots)
\end{aligned}$$

Odd integers divisible by three are denoted in boldface, corresponding to the 'missing' sequences $\{\beta_3\}$, $\{\beta_9\}$, etc.

Odd Number Generator

While each sequence $\{\beta_j\}$ contains an infinite count of odd integers, naturally one wonders if *every* odd integer is contained somewhere in a β-sequence. Indeed, each odd number can be systematically generated by

$$\beta_{1+6j,k} = \frac{(1+6j)\cdot 2^{2k}-1}{3} \qquad\qquad \beta_{5+6j,k} = \frac{(5+6j)\cdot 2^{2k-1}-1}{3}\,,$$

where $j = 0, 1, 2, \ldots$ and $k = 1, 2, 3, \ldots$ in each:

j	k	$\beta_{1+6j,k}$	j	k	$\beta_{5+6j,k}$
0	1	1	0	1	3
0	2	5	1	1	7
1	1	9	2	1	11
2	1	17	0	2	13
0	3	21	3	1	15
3	2	25	4	1	19
4	1	33	5	1	23
1	2	37	6	1	27
5	1	41	1	2	29
6	1	49	7	1	31
			8	1	35
			9	1	39
			10	1	43
			2	2	45
			11	1	47
			12	1	51
			0	3	53
			13	1	55

The β-equations above can be generalized into a single equation

$$\beta_{x,y} = \frac{x \cdot 2^y - 1}{3} \qquad y = 1, 2, 3, \ldots \qquad \frac{x}{3} \neq \mathcal{Z},$$

where y is a positive integer, and x is an integer not divisible by three. Generating the same table of odd numbers, we have:

x	y	$\beta_{x,y}$	x	y	$\beta_{x,y}$
0	2	1	11	3	29
5	1	3	47	1	31
1	4	5	25	2	33
11	1	7	29	2	35
7	2	9	7	4	37
17	1	11	59	1	39
5	3	13	31	2	41
23	1	15	65	1	43
13	2	17	17	3	45
29	1	19	71	1	47
1	6	21	37	2	49
35	1	23	77	1	51
19	2	25	5	5	53
41	1	27	83	1	55

Since $\beta_{x,y}$ is always an odd integer, we may apply the progress operator $f\left(\beta_{x,y}\right)$ as a sanity check to write

$$f\left(\beta_{x,y}\right) = 3\left(\frac{x \cdot 2^y - 1}{3}\right) + 1 = x \cdot 2^y \ ,$$

which is surely an even integer. Applying the compound progress operator $f_{y+1}\left(\beta_{x,y}\right)$, the result reduces down to x via

$$f_{y+1}\left(\beta_{x,y}\right) = x \cdot \frac{2^y}{2^y} = x \ ,$$

reaffirming that an arbitrary odd integer $\beta_{x,y}$ eventually links forward to another odd integer x not divisible by three. Meanwhile, odd integers that *are* divisible by three, namely 3, 9, 15, etc., can only occur as the first odd member of a sequence $\phi\left(n, k\right)$.

Cutoff Analysis

The sequences $\{\beta_j\}$ avail a method for testing a range of odd integers starting from one and ending at a cutoff $\tilde{n}$. Suppose we were tasked with verifying the Collatz conjecture for odd integers $1 \leq n \leq 29$, arranged on the grid that follows. Treating $n = 1$ as the only 'tested' case, cross out from the grid any integers less than $\tilde{n}$ that occur in the sequence $\{\beta_1\}$:

$$\begin{array}{ccccc} \cancel{1} & 3 & \cancel{5} & 7 & 9 \\ 11 & 13 & 15 & 17 & 19 \\ \cancel{21} & 23 & 25 & 27 & 29 \end{array}$$

Now, since $n = 5$ has been crossed out, we may read across $\{\beta_5\}$ and cross out from the grid any integers less than $\tilde{n}$:

$$\begin{array}{ccccc} \cancel{1} & \cancel{3} & \cancel{5} & 7 & 9 \\ 11 & \cancel{13} & 15 & 17 & 19 \\ \cancel{21} & 23 & 25 & 27 & 29 \end{array}$$

This time $n = 13$ has been crossed out, which means we may read across $\{\beta_{13}\}$ and cross out $n = 17$, which then gives us $n = 11$, and then immediately 7, 29 from $\{\beta_{11}\}$:

$$\begin{array}{ccccc} \cancel{1} & \cancel{3} & \cancel{5} & \cancel{7} & 9 \\ \cancel{11} & \cancel{13} & 15 & \cancel{17} & 19 \\ \cancel{21} & 23 & 25 & 27 & \cancel{29} \end{array}$$

Continuing this pattern, we find that 9, 19, 25 can also be crossed out:

$$\begin{array}{ccccc} \cancel{1} & \cancel{3} & \cancel{5} & \cancel{7} & \cancel{9} \\ \cancel{11} & \cancel{13} & 15 & \cancel{17} & \cancel{19} \\ \cancel{21} & [23] & \cancel{25} & [27] & \cancel{29} \end{array}$$

So far then, *only* the integers 15, 23, 27 less than $\tilde{n} = 29$ don't follow automatically from $\{\beta_1\}$ and its branches. Of course, the first element in $\{\beta_{23}\}$ is 15, meaning there are only two unique calculations left, 23 and 27, denoted in brackets [ ] above. To see why, we solve

$$\frac{(5 + 6j) \cdot 2^{2k-1} - 1}{3} = 23 \ ,$$

to find $j = 5$, $k = 1$, meaning 23 is the first element in $\{\beta_{35}\}$, and is only a guaranteed solution if $n = 35$ is a solution, which is outside of the $\tilde{n}$-domain. Thus, we're stuck testing $n = 23$ for $\tilde{n}$ fixed at 29. Similar reasoning applies to $n = 27$.

Extending $\tilde{n}$ from 29 to 39, we need the elements of $\{\beta_{35}\}$, $\{\beta_{37}\}$ that are less than $\tilde{n}$. Using the β-equations above, we find

$$\{\beta_{35}\} = (23, 93, 373, \dots)$$
$$\{\beta_{37}\} = (49, 187, 789, \dots) \ ,$$

extending our table of odd numbers by one row, crossing out members that trace back to $n = 1$:

~~1~~	~~3~~	~~5~~	~~7~~	~~9~~
~~11~~	~~13~~	15	~~17~~	~~19~~
~~21~~	23	~~25~~	[27]	~~29~~
[31]	~~33~~	[35]	~~37~~	[39]

In contrast to the $\tilde{n} = 29$ case, there are four unique integers to verify, however note that $n = 23$ need not be uniquely verified.

Extending $\tilde{n}$ further by one member, making $\tilde{n} = 41$, first note that

$$\{\beta_{41}\} = (27, 109, 437, \dots) \ ,$$

reducing the burden of verifying $n = 27$ to verifying $n = 41$, which is already a member of $\{\beta_{31}\}$, leaving us with:

~~1~~	~~3~~	~~5~~	~~7~~	~~9~~
~~11~~	~~13~~	15	~~17~~	~~19~~
~~21~~	23	~~25~~	27	~~29~~
[31]	~~33~~	[35]	~~37~~	[39]
41				

So far, only the integers 31, 35, 39 need to be directly verified in the domain $1 \leq n \leq \tilde{n} = 41$.

Without extending $\tilde{n}$, we use $n \to f(n)$ to calculate the next odd integers that follow 31, 35, 39 respectively:

$$31 \to 94 \to 47$$
$$35 \to 106 \to 53$$
$$39 \to 118 \to 59$$

While the 31- and 39-cases don't help for our choice of $\tilde{n}$, the 35-case lands on 53, which is a member of $\{\beta_5\}$, already crossed out. Thus, 35 may be crossed out as well (along with 53 when we get there):

~~1~~	~~3~~	~~5~~	~~7~~	~~9~~
~~11~~	~~13~~	15	~~17~~	~~19~~
~~21~~	23	~~25~~	27	~~29~~
[31]	~~33~~	~~35~~	~~37~~	[39]
41				

Extending $\tilde{n}$ higher, the 'frontier' of non-guaranteed integers shifts upward, leaving crossed-out or otherwise guaranteed integers behind. We learn that the burden of testing a range of integers up to $\tilde{n}$ reduces to the smaller task of verifying a handful of bracketed integers near $\tilde{n}$. Cutoff analysis transforms the 'hard work' of carrying out repetitions of $n \to f(n)$ to table lookup.

Very Odd Integers

Looking once more at the generator for all odd integers, namely

$$\beta_{x,y} = \frac{x \cdot 2^y - 1}{3} \qquad y = 1, 2, 3, \ldots \qquad \frac{x}{3} \neq \mathcal{Z} \; ,$$

we assign a special name to any case when $y = 1$, giving rise to *very odd integers* $\beta_{x,1} < x$. That is, a very odd integer $\beta_{x,1}$ is greater than the odd integer x (not divisible by three) that generates it.

With respect to cutoff analysis, very odd integers whose generating x-values are greater than a given cutoff $\tilde{n}$ are precisely those that must be checked explicitly (in square brackets). Explicitly, the members 31 and 39 in the table above each (i) qualify as very odd integers, and (ii) point to odd integers greater than $\tilde{n}$.

As $\tilde{n}$ increases arbitrarily, it follows that most odd integers in the associated table are either crossed directly, or depend on other odd numbers that are eventually crossed out. The adjustment in $\tilde{n}$ required to cover such a very odd integer $\beta_{x,1}$ is

$$\tilde{n} \approx \frac{3\beta_{x,1} + 1}{2} \; .$$

Thus, no odd integer remains permanently out of reach with increasing $\tilde{n}$. In the worst-case scenario, the cutoff value would itself qualify as a very odd integer, giving $\tilde{n} = \beta_{x,1}$.

5 Cycles

Let us denote a *cycle* $\theta_k(\{n\})$ as an occurrence where an integer n lands back at its original value after k jumps such that

$$f_k(n) = n \; .$$

It follows that any integers that progress from n via $f(n)$ also repeat their values, and are thus members of the same cycle. Applying successive instances of $f(\,)$ to both sides of the above, we eventually land at

$$f_{2k}(n) = f_k(n) = n \; ,$$

affirming the k-periodicity of the cycle.

One-Odd Cycle

It is clear that a cycle must contain at least one odd integer, as jumping to strictly even integers will always have $n \to n/2$, and the cycle would never be established. Consider the general cycle

$$\theta_k = (a, \ldots, r, s, t, \alpha, v, w, x, \ldots, a) \ ,$$

where α is a forced odd integer in the qth position, with all other elements being unknown integers. Working outward from α, its immediate neighbors can only be even integers:

$$\theta_k = (a, \ldots, r, s, 2\alpha, \alpha, 3\alpha + 1, w, x, \ldots, a)$$

Supposing we seek a cycle with precisely one odd integer in the qth position, the general cycle becomes

$$\theta_k = \left(2^{q-1}\alpha, \ldots, 8\alpha, 4\alpha, 2\alpha, \alpha, 3\alpha + 1, \frac{3\alpha+1}{2}, \frac{3\alpha+1}{4}, \ldots, \frac{3\alpha+1}{2^{k-q}}\right) \ .$$

Comparing the first and last terms, see that

$$a = 2^{q-1}\alpha = \frac{3\alpha+1}{2^{k-q}} \qquad \to \qquad \alpha = \frac{3\alpha+1}{2^{k-1}} \ ,$$

which loses all q-dependence, as the cycle is invariant with respect to which element is listed first. Isolating α, we find

$$\alpha = \frac{1}{2^{k-1} - 3} \ ,$$

which is only solved by the pair $\alpha = 1$, $k = 3$. Evidently then, we find

$$\theta_3 = (4, 2, 1, 4) = (1, 4, 2, 1) = (2, 1, 4, 2)$$

to be the *only* cycle allowed to contain one odd integer.

Two-Odd Cycle

Consider a cycle that contains exactly two odd integers α and β

$$\theta_k = (a, \ldots, 2\alpha, \alpha, 3\alpha + 1, \ldots, 2\beta, \beta, 3\beta + 1, \ldots, a) \ ,$$

with all other elements being even integers. Following the even integers from α, we advance λ positions to the right until encountering β. Similarly, starting from β, we advance λ' positions to the right until encountering α, giving simultaneous equations

$$\frac{3\alpha+1}{2^{\lambda}} = \beta \qquad\qquad \frac{3\beta+1}{2^{\lambda'}} = \alpha \ ,$$

where the exponents λ, λ' determine the total length of the cycle. Isolating α, β, we have

$$\alpha = \frac{2^{\lambda}+3}{2^{\lambda+\lambda'} - 9} \qquad\qquad \beta = \frac{2^{\lambda'}+3}{2^{\lambda+\lambda'} - 9} \ .$$

For two-odd cycles with λ and λ' very large, we approximately have

$$\alpha \approx \frac{1}{2^{\lambda'}} \qquad\qquad \beta \approx \frac{1}{2^{\lambda}}\ ,$$

which has no valid solution. We thus deduce that, if the two-odd cycle is to contain very many elements, then the two odd numbers α, β cannot be separated by very many jumps. To capture this, we write

$$\epsilon = |\lambda - \lambda'| \ll \lambda + \lambda'\ ,$$

where ϵ is a positive integer far smaller than the number of elements in the cycle. Next, the ratio β/α tells us

$$\frac{(3\alpha+1)\,\alpha}{(3\beta+1)\,\beta} = 2^{\lambda-\lambda'} \qquad \rightarrow \qquad \ln\left(\left|\frac{(3\alpha+1)\,\alpha}{(3\beta+1)\,\beta}\right|\right)\frac{1}{\ln 2} = |\lambda-\lambda'| = \epsilon\ ,$$

reaffirming α, β cannot vastly differ in value. To proceed, choose the case that α follows from a long descent of even numbers, and then β occurs $\approx \epsilon$ jumps after. As β is reached, the next jump is $3\beta+1$, which is by construction far smaller than the even numbers that lead to α. The descent from β cannot link back to α, telling us that a two-odd cycle with many elements cannot exist.

For two-odd cycles with α, β very large, we first notice from the first set of equations that

$$\alpha \gg 2^{\lambda} \qquad\qquad \beta \gg 2^{\lambda'}\ .$$

If so, the second set of equations demands that the denominator $2^{\lambda+\lambda'}-9$ be a small number, maximizing α, β at $2^{\lambda+\lambda'} \approx 9$, severely restricting the sum $\lambda+\lambda'$. The smallest denominator we legally make corresponds to $\lambda+\lambda'=4$, meaning the cycle must have very few elements. Moreover, the numerators $2^{\lambda}+3$, $2^{\lambda'}+3$ already correspond to numbers smaller than α, β, respectively. Therefore, a cycle containing exactly two very large odd integers implies contractions, and cannot exist.

Clearly, the two-odd cycle cannot contain one large element and one non-large element, leaving the last non-trivial case, in where the cycle has two non-large odd integers. At this point we note the chore of testing the Collatz conjecture can be automated, where the domain of successfully tested integers spans from one to $\approx 2^{60}$ (citation needed). Thus, all small-enough integers are already 'used up' in valid Collatz sequences, and cannot participate in two-odd cycles. In summary, no two-odd cycle can occur at all.

Multi-Odd Cycle

Generalizing the two-odd cycle analysis entails noticing that the structure $2\alpha_j, \alpha_j, 3\alpha_j+1$ occurs once per odd integer, with index $j = 1, 2, 3, \ldots, N$ tracking each:

$$\theta_k = (a, \ldots, 2\alpha_1, \alpha_1, 3\alpha_1+1, \ldots, 2\alpha_2, \alpha_2, 3\alpha_2+1, \ldots, 2\alpha_3, \alpha_3, 3\alpha_3+1, \ldots, a)$$

This generates a list of j equations relating to the number of jumps between the odd numbers α_j:

$$\frac{3\alpha_j+1}{2^{\lambda_j}} = \alpha_{j+1} \qquad\qquad \frac{3\alpha_N+1}{2^{\lambda_N}} = \alpha_1$$

Looking closely, we notice that any odd integer α_j 'ratchets upward' by a factor of $3\alpha_j+1$, and then downward by a factor of 2^{λ_j} to produce the next odd member α_{j+1} in the cycle. By similar arguments that apply to the two-odd cycle, the exponents $\{\lambda_j\}$ cannot be very large, as a high density of $n \to n/2$ operations induces a downward trend in n, preventing the end of the cycle from linking to the beginning. Thus if multi-odd cycles exist, the population of odd integers is 'dense', i.e. not separated by many jumps.

Maximal-Odd Cycle

A special case of the multi-odd cycle occurs if all $\lambda_j = 1$, meaning every second element is an odd integer. Such a cycle cannot exist, as the odd base numbers produced by $n \to (3n+1)/2$ can only increase, and the cycle is forever open. This in fact corresponds to a divergent sequence with $n \to \infty$.

6 Summary Analysis

The Role of Three

Odd integers divisible by three play a special role in the problem. It was shown that such integers cannot occur as intermediate elements in any given sequence $\phi(n,k)$, but only occur as the first odd element. As a corollary, we find that no integer divisible by three can occur in any cycle θ_k.

On a separate note, the base number $n = 3^3 = 27$ leads to an unexpectedly long Collatz sequence, finally reaching one after 111 jumps. Supposing we undertake the computational burden of verifying $f_{111}(27) = 1$, it follows that $n = 27$ lives in the regress tree $\psi(\tilde{n} = 111)$, which has approximately

$$N(111) \approx \exp(111/6) \approx 10^8$$

members.

Regress Trees and Cycles

With exception of the trivial cycle $\theta_3 = (4, 2, 1, 4)$, it follows that the existence of any cycle θ_k that arises via $f_k(n) = n$ would give rise to recursive branches in the grand regress tree $\Psi(\tilde{n})$. Of course, regress trees are built from the $n = 1$ case, which is never reached by the cycle θ_k. Thus, cycles are not represented in regress trees.

For an odd integer α_j in a cycle θ_k, there eventually exists another odd integer γ, namely a member of $\left\{\beta_{\alpha_j}\right\}$, that leads to α_j from outside the cycle. We similarly conclude that no γ-like integer can be a member of a regress tree.

Stepping to Infinity

Consider the set of regress trees

$$\Psi(\tilde{n}) = \{\psi(\tilde{n})\} = \{n\} \ ,$$

where any given $\psi(\tilde{n})$ contains all of the integers $\{n\}$ that are $\tilde{n}$ jumps from one. In condensed notation, we found:

$$
\begin{aligned}
0 &\,|\, 1 \\
1 &\,|\, 2 \\
2 &\,|\, 4 \\
3 &\,|\, 8 \\
4 &\,|\, 16 \\
5 &\,|\, 32, 5 \\
6 &\,|\, 64, 10 \\
7 &\,|\, 128, 21, 20, 3 \\
8 &\,|\, 256, 42, 40, 6 \\
9 &\,|\, 512, 85, 84, 80, 13, 12 \\
10 &\,|\, 1024, 170, 168, 160, 26, 24 \\
11 &\,|\, 2048, 341, 340, 336, 320, 54, 52, 48 \\
12 &\,|\, 4096, 682, 680, 113, 672, 640, 106, 104, 17, 96 \\
13 &\,|\, 8192, 1365, 1364, 227, 1360, 226, 1344, 1280, 213, 212, 35, 208, 34, 192 \\
\dots &\,|\, \dots
\end{aligned}
$$

The left column lists the set of all jump numbers $\tilde{n} = (1, 2, 3, \dots)$, whereas the field of integers $\{n\}$ on the right are those that satisfy the Collatz condition for a given $\tilde{n}$. While we guarantee no cycle exists in $\Psi(\tilde{n})$, there is no air-tight assurance that $\{n\}$ does not 'skip' any integers.

Tree vs. Lattice

Base-odd lattice analysis allows all odd integers to be written in a Collatz-ready format, namely

$$\beta_{x,y} = \frac{x \cdot 2^y - 1}{3} \qquad y = 1, 2, 3, \dots \qquad \frac{x}{3} \neq \mathcal{Z} \,.$$

In contrast to the regress tree, the base-odd lattice has an automatic plan for all integers. That is, no integer is excluded from the lattice, however the existence of a mysterious multi-odd cycle is not ruled out from existing when the progress operator is to applied some undiscovered special set of integers.

As our final move, let us combine the cycle-free advantage of the regress tree with the all-integers-included advantage of the base-odd lattice. Perhaps not astonishingly, the regress tree and the base-odd lattice seem to contain the *same* data. The most obvious pattern in each structure is $(1, 2, 4, 8, 16, \dots)$, however looking closely, the pattern $(5, 10, 20, 40, 80, \dots)$ also occurs in each, and so on. That is, the content of $\Psi(\tilde{n})$, reading 'downward' across $\tilde{n}$, embeds the columns $\{\alpha_j 2^\lambda\}$. By the same token, a reciprocal reading of the base-odd lattice can recover the regress tree. In the equivalence

$$\lim_{\tilde{n} \to \infty} \Psi(\tilde{n}) \leftrightarrow \{\alpha_j 2^\lambda\} \ \forall\, j, \lambda \in \mathcal{Z} > 0 \,,$$

we must have that (i) there exist no multi-odd cycles in the sequences $\{\beta_j\}$, and (ii) the regress tree $\Psi(\tilde{n})$ eventually contains every integer. Since $\Psi(\tilde{n})$ is the list of all integers that satisfy the Collatz conjecture, and all integers are seemingly on the list, we can finally stop.