Collatz Conjecture Proof

Abstract. Collatz sequences are formed by applying the Collatz algorithm to any positive integer. If it is even repeatedly divide by two until it is odd, then multiply by three and add one to get an even number and vice versa. If the Collatz conjecture is true eventually you always get back to one. A connected Collatz Structure is created, which contains all positive integers exactly once. The terms of the Collatz Structure are joined together via the Collatz algorithm. Thus, every positive integer forms a Collatz sequence with unique terms terminating in the number one.

Introduction. The Collatz conjecture was made in 1937 by Lothar Collatz. Through 2017 the conjecture has been checked for all starting values up to $(87)(2^{60})$, but very little progress has been made toward proving the conjecture. Paul Erödos said about the Collatz conjecture: "Mathematics may not be ready for such problems." https://en.wikipedia.org/wiki/Collatz_conjecture

9→28→14→7→22→11→34→17→52→26→13→40 is a typical Collatz structure branch. Only terms of the form ***6n+3*** appear at the beginning of a branch. Since three divides *6n+3* evenly, **the only predecessors** of ***6n+3*** are of the form $(2^j)(6n+3)$, and they are not considered part of the branch. The Collatz algorithm is applied until the **last term** of the form ***24k+16*** appears. In appendix 1 we show there can be no more than two consecutive even terms in a branch. *24k+16* terms being divisible by eight must appear at the end of a branch. We will show there are no unending branches. By applying the Collatz algorithm to terms within a branch you will eventually come to a *24k+16* term. By applying the Collatz algorithm in reverse you will eventually come to a *6n+3* term.

In appendix 1 we show that the **only successor of odd terms** in a branch are terms of the form ***24m+4, 24m+10, 24m+16, or 24m+22.*** As shown immediately below, every ***24k+16*** term is of the form $4^j a$, ***a=24m+4, 24m+10, or 24m+22,*** $j \geq 1$, $m \geq 0$). As ***j*** increases, *24k+16* terms rise vertically in towers atop a *24m+4, 24m+10, or 24m+22* base within the Collatz structure.

Counting back towards its base term, $24k_n+16$ is the *n-th* term in a tower. Applying the Collatz algorithm

$24k_n+16 \rightarrow 12k_n+8 \rightarrow \underline{6k_n+4}$. $k_n = 4k_{n-1}+2$, $\underline{6k_n+4} = 24k_{n-1}+16 \rightarrow 12k_{n-1}+8 \rightarrow 6k_{n-1}+4$ ….

$24k_2+16 \rightarrow 12k_2+8 \rightarrow \underline{6k_2+4}$. $k_2 = 4k_1+2$, $\underline{6k_2+4} = 24k_1+16 \rightarrow 12k_1+8 \rightarrow \underline{6k_1+4}$.

If $k_1=4m$, $\underline{6k_1+4} = 24m+4$ or $k_1=4m+1$, $\underline{6k_1+4} = 24m+10$. If $k_1=4m+3$, $\underline{6k_1+4} = 24m+22$.

The Collatz Structure starts with the **Trunk Tower.** Each $(4^j)(4)$, $j=1,2,3$…Trunk Tower term is the last term in a branch: a ***24k+16*** term with ***a=4***. At every ***a=24m+4, 24m+10,*** and ***24m+22*** base term in the Trunk Tower branches is a $4^j a$, $j=1,2,3…$ secondary tower. Each of these $4^j a$ terms in the secondary towers is the last term in a branch. At every ***a=24m+4, 24m+10,*** and ***24m+22*** base term in these secondary branches is a $4^j a$, $j=1,2,3…$ secondary tower. Each $4^j a$ is the last term in a branch. This process is repeated indefinitely.

Collatz Structure Branches and Towers ↓indicates apply the Collatz algorithm to terms above

				↓				↓				↓				↓		↓							
↓				160	←	53	←	106	←	35	←	70	←	23	←	46	←	15							
128		↓		80					↓				↓				↓						↓		↓
64	←	21		40←	13	←	26	←	52	←	17	←	34	←	11	←	22	←	7	←	14	←	28	←	9
32				20	↓																				
16	←	5	←	10←	3																				
8																									
4																									
2																									
1																									

The ***6n+3*** branch first terms are sub-divided into four types: ***24h+3, 24h+9, 24h+15*** **and** ***24h+21,*** $h \geq 0$.
A branch **binary series** counts the number of divisions by two on its tower base terms: ***24m+4 (2), 24m+10 (1),*** and ***24m+22 (1)***. The binary series will be used to show that there are no unending branches.

Section 1

Defining the "binary series" of a branch.

Only ***$24h+3$, $24h+9$,*** **and** ***$24h+15$*** first terms appear in branches with binary series. These three groups of branches are characterized by their first term ***$24h+3$, $24h+9$*** or ***$24h+15$*** and a binary series of ***1***'s and ***2***'s (see ***2,1,1,2*** below) counting the divisions by two on their tower base terms ***$24m+4$ (2), $24m+10$ (1),*** or ***$24m+22$ (1)*** and a last term ***$24k+16$.*** The length of a binary series is the number of tower base terms in the branch. Let ***r*** be the length of the binary series. If the sum of ***r 1's*** and ***2's*** in the binary series is ***s,*** there are three different groups of branches each having the same binary series, whose first terms differ by ***$(24)(2^s)$***. The first terms are ***$24h+3+(p-1)(24)(2^s)$, $24h+9+(p-1)(24)(2^s)$,*** and ***$24h+15+(p-1)(24)(2^s)$, $p=1,2,3\ldots$ $h \geq 0$.*** All groups end with ***$24k+16+(p-1)(24)(3^{r+1})$, $p=1,2,3\ldots$ $k \geq 0$. $24h+21$*** has no binary series. However, there are a group of branches that begin with ***$24h+21+(p-1)(24)$*** followed immediately by the branch last term ***$(24)(3h+2)+16+(p-1)(24)(3)$***.

A binary series *(2,1,1,2)* counts divisions by two on the even terms *$24m+4$ (2), $24m+10$ (1), $24m+22$ (1)*.
The sum of this binary series is six. These are a series of branches whose first terms differ by ***$(24)(2^6)=1536$.***
The first term sequence is ***$9+(p-1)(24)(2^6)$ 9, 1545, 3081…*** The last terms differ by ***$(24)(3^5)=5832$***.
The last term sequence is ***$40+(p-1)(24)(3^5)$ 40, 5872, 11704…***

The first branch is ***9, 28, 7, 22, 11, 34, 17, 52, 13, 40.***
The second branch is ***1545, 4636, 1159, 3478, 1739, 5218, 2609, 7828, 1957, 5872.***
The third branch is ***3081, 9244, 2311, 6934, 3467, 10402, 5201, 15604, 3901, 11704.***

Proving the formula for branches with the same binary series.

Start with ***$24h+3$,*** and ***$24h+3+(p)(24)(2^s)$.*** Multiplying by three and adding one ***($2j+1 \rightarrow 6j+4$)*** gives two terms that differ by ***$(p)(24)(2^s)(3)$ $72h+10$*** and ***$72h+10+(p)(24)(2^s)(3)$.*** A total ***$r+1$*** applications of ***$2j+1 \rightarrow 6j+4$*** to ***$24h+3$,*** and its odd successors, and ***s*** divisions by two on ***$72h+10$*** and its even successors, which cause a ***$24k+16$*** term to appear, are mirrored in ***$24h+3+(p)(24)(2^s)$*** so that a ***$24k+16+(p)(24)(3^{r+1})$*** term appears. The same proof holds for groups with the first terms ***$24h+9+(p-1)(24)(2^s)$,***
and ***$24h+15+(p-1)(24)(2^s)$, $p=1,2,3\ldots$ $h \geq 0$.***

Calculating the proportion of all *$24h+3$, $24h+9$, $24h+15$* and *$24h+16$* terms in branches.

We will prove that all ***$24h+3$, $24h+9$, $24h+15$*** and ***$24h+16$*** terms are in branches. We will show that there are branch binary series of all lengths ***r*** with all ***2^r*** possible combinations of ***1***'s and ***2***'s for every value of ***r.***

There are three series of branches whose binary series sums to ***s*:**

$24h+3+(p-1)(24)(2^s)$, $24h+9+(p-1)(24)(2^s)$, **and** ***$24h+15+(p-1)(24)(2^s)$, $p=1,2,3\ldots$ $h \geq 0$.***

If all ***$24h+3$, $24h+9$,*** and ***$24h+15$*** terms are put in three separate ascending sequences, terms with the same binary series occur every ***2^s*** terms: ***$1/2^s$*** proportion of the sequence terms. ***$h < 2^s$*** for each sequence first term.

$24h+3 \rightarrow 72h+10 \rightarrow 36h+5 \rightarrow 108h+16$, shows the first two ***$24h+3$*** binary series are ***(1)*** if ***h*** is even and ***(1,2)*** if ***$h=3$.*** All other binary series begin with ***(1,2,…).*** The proportion of ***$24h+3$*** terms with binary series ***(1)*** is ***$1/2^1$.*** We will prove by an induction argument that the proportion of ***$24h+3$*** terms with binary series length ***$r \geq 2$*** is ***$3^{r-2}/2^{2r-1}$.*** The proportion of ***$24h+3$*** terms with binary series ***(1,2)*** is ***$1/2^3$*** verifying the formula for ***$r=2$***. The ***$r+1$*** position of every binary series of that length contains either ***(1)*** one or ***(2)*** two divisions by two.

The proportion of ***$24h+3$*** terms of binary series length ***$r+1$*** is

$$(1/2^1)(3^{r-2}/2^{2r-1})+(1/2^2)(3^{r-2}/2^{2r-1}) = (3/4)(3^{r-2}/2^{2r-1}) = 3^{r-1}/2^{2(r+1)-1}.$$

Starting with ***1/2*** for length one and summing the geometric series for length ***$r \geq 2$***

$$1/2+1/8+3/32+9/128+\ldots= 1/2+(1/8)/(1-3/4)=1.$$

For ***$h=3$*** the ***$24h+9$*** branch binary series binary series is ***(2)***. All other binary series begin with ***(2,…).***
We will prove by an induction argument that the proportion of ***$24h+9$*** terms with binary series length ***$r \geq 1$*** is ***$3^{r-1}/2^{2r}$.*** The proportion of ***$24h+9$*** terms with binary series ***(2)*** is ***$1/2^2$*** verifying the formula for ***$r=1$.***

The proportion of ***$24h+9$*** terms with length ***$r+1$*** is ***$(1/2)(3^{r-1}/2^{2r})+(1/4)(3^{r-1}/2^{2r})=3^r/2^{2(r+1)}$.***
Summing the geometric series for length ***$r \geq 1$*** gives ***$1/4+3/16+9/64+\ldots= (1/4)/(1-3/4)=1$.***

$24h+15 \rightarrow 72h+46 \rightarrow 36h+23 \rightarrow 108h+70 \rightarrow 54h+35 \rightarrow 162h+106$ shows that for $h=3$ the first binary series is $(1,1)$. All other binary series begin with $(1,1,\ldots)$. We will prove by an induction argument that the proportion of $24h+15$ terms with binary series length $r \geq 2$ is $3^{r-2}/2^{2r-2}$. The proportion of $24h+15$ terms with binary series $(1,1)$ is $1/2^2$ verifying the formula for $r=2$.

The proportion of $24h+15$ terms with length $r+1$ is $(1/2)(3^{r-2}/2^{2r-2})+(1/4)(3^{r-2}/2^{2r-2})=3^{r-1}/2^{2(r+1)-2}$.

Summing the geometric series for length $r \geq 2$ gives $1/4+3/16+9/64+\ldots=(1/4)/(1-3/4)=1$.

Calculating the proportion of $24k+16$ terms are in branches.

The formula for the last term in a group of branches with the same binary series of length r is $24k+16+(p-1)(24)(3^{r+1})$ $p=1,2,3\ldots$ $k < 3^{r+1}$. Put all $24k+16$ terms in an ascending sequence. The proportion of all terms with the same binary series of length r is $1/3^{r+1}$ of the terms in the sequence. The last terms of the 2^r branches with binary series of length r are $2^r/3^{r+1}$ proportion of all terms in the sequence.

The total proportion of $24k+16$ terms of length $r \geq 0$ is $1/3+2/9+4/27+\ldots = (1/3)/(1-2/3)=1$.

The total proportions of one indicate that every $24h+3$, $24h+9$, $24h+15$ and $24h+16$ term are in branches. That they are all in branches with binary series of all 2^r combinations of r 1's and 2's for every value of r indicates there are no unending branches.

Section 2

The repeating binary series structure of towers.

Within a tower if the sum of r 1's and 2's in the binary series of a branch is s, there are three groups of branches having the same binary series. The first begins with $24h+3+(2^s)(24k+16)(4^{(x)(p-1)}-1)/3^{r+1}$, and ends with $(24k+16)(4^{(x)(p-1)})$, $x=3^{r+1}$, $p=1,2,3\ldots$ The other two groups that begin with $24h+9\ldots$ and $24h+15\ldots$ have the same form as $24h+3\ldots$ **$r+1$ applications of $2j+1 \rightarrow 6j+4$** applied to $24h+3$ and its odd successors and applied to $(2^s)(24k+16)(4^{(x)(p-1)}-1)/3^{r+1}$ and s divisions by two applied to $72h+10$ and its even successors and applied to $(2^s)(24k+16)(4^{(x)(p-1)}-1)/3^r$ gives

$$(24k+16)+(24k+16)(4^{(x)(p-1)}-1) = (24k+16)(4^{(x)(p-1)}).$$

A branch with no binary series starts with $24h+21+((24)(3h+2)+16)(4^{(3)(p-1)}-1)/3$ and ends with $((24)(3h+2)+16)(4^{(3)(p-1)})$.

Link between the formulas for branch and tower first terms.

For some t, $24h+3+(t-1)(24)(2^s) = 24h+3+(2^s)(24k+16)(4^{(x)(p-1)}-1)/3^{r+1}$.

For $x=3^{r+1}$ every power of three in $4^{(x)(p-1)}-1=(3+1)^{(x)(p-1)}-1$ has a coefficient divisible by 3^{r+1}. $(24k+16)(4^{(x)(p-1)}-1)/3^{r+1}$ is a multiple 24. The same is true for the forms beginning with $24h+9\ldots$, $24h+15\ldots$, and $24h+21\ldots$ Each tower's branch binary series structure is a microcosm of the total branch binary series structure. $4^{(x)(p-1)}$, $x=3^{r+1}$ replaces 3^{r+1}. In each case the last terms of tower branches with the same binary series occur in intervals of 3^{r+1}. $2^r/3^{r+1}$ is the proportion of the 2^r last terms of tower branches with a binary series of length r.

For length $r \geq 0$ $1/3+2/9+4/27\ldots= 1$ is the total proportion.

There are tower branches with binary series of all 2^r combinations of r 1's and 2's for every value of r. The first branch with a binary series of length r comes within the first 3^{r+1} branches in the tower.

Section 3

All terms in branches and towers have unique predecessors and successors.

All even terms $24m+2y$, $y=0$ to 11 and all three odd terms $6n+1$, $6n+3$, and $6n+5$ are accounted for.
The predecessors of the first term in a branch are $(2^m)(6n+3)$ $24m+0$, $24m+6$, $24m+12$, and $24m+18$.
See appendix 1 for branch successors of $6n+1$, $6n+3$, $6n+5$ and predecessors of $24m+4$, $24m+10$, $24m+16$ and $24m+22$. $24n+4 \rightarrow 12n+2$ $(24m+2, 24m+14) \rightarrow 6n+1$. $12n+10$ $(24m+10, 24m+22) \rightarrow 6n+5$. q is the successor of each even term $2q$ in a branch. As shown above in the description of a tower, all $12k+8$ $(24m+8, 24m+20)$ are in towers.

Every positive integer is in a branch or a tower exactly once. Every $6n+3$ term is at the beginning of a branch, and every $24k+16$ term is the last term in a branch. Multiplying each of $24m+4$, $24m+10$, and

24m+22 by four gives a ***24k+16*** term so all ***24m+4, 24m+10,*** and ***24m+22*** terms are tower bases. As shown above, all ***6n+1, 6n+5,* and *12n+2 (24m+2, 24m+14)*** are in branches. Since there are no unending branches, all ***$(2^j)(6n+3)$ 24m+0, 24m+6, 24m+12,*** and ***24m+18*** terms appear above ***6n+3*** terms in the Collatz Structure. Since all ***24k+16*** are branch last terms, all ***12k+8 (24m+8, 24m+20)*** terms appear in towers.

There can be no duplicate terms in a branch. All the predecessors of a duplicate pair of terms would be duplicates. This would require ***24h+3, 24h+9,*** or ***24h+15*** to be a duplicate term, and those terms only appear at the beginning of a branch. There can be no duplicates in a tower. They are strictly increasing sequences. Since they all start with a different base, no duplicates can appear in different towers. Finally, no duplicate terms can appear in different branches. From the second term forward until the last term is reached all terms in branches have unique predecessors and successors.

Section 4

The Collatz Structure containing all positive integers is a connected structure. There are no circular or unending Collatz sequences. To prove this we need to define a new item that is a part of all Collatz sequences. An L_8 begins with a ***24k+16*** term in a secondary tower. The Collatz algorithm is applied until a ***24k+16*** term appears in an adjoining tower. A chain of adjoining L_8 moves through Collatz Structure until reaching a ***24k+16*** Trunk Tower term. An L_8 chain binary series is built from the number of divisions by two on all the tower base terms in the L_8 chain. The usage factor for the L_8 chain binary series is calculated by inverting the powers of two in the even factors of the tower base terms and summing the resulting geometric series. We will prove by induction that the usage factor of an L_8 chain with a binary series of **length *r*** is $3^r/4^r$. The binary series of an L_8 chain with one tower base term is ***1,*** or ***2.*** The usage factor is $1/2^1+1/2^2= 3/4$ verifying the formula for $r = 1$. For **length *r+1*** the usage factor is $(1/2)(3^r/4^r)+(1/4)(3^r/4^r) = 3^{r+1}/4^{r+1}$**.**

Every *24m+4, 24m+10, and 24m+22* tower base term in an L_8 chain is in a branch with a first term of *24h+3, 24h+9, or 24h+15.* The sum of the geometric series of the L_8 chain binary series is ***3/4+9/16+27/64+…= (3/4)/(1 – 3/4)=3.*** This equals the total proportion of ***24h+3, 24h+9,*** and ***24h+15*** terms in branches. This total proportion is the sum of three geometric series, which are based on the powers of two of even factors in tower base terms. The equality between the total proportion of ***24h+3, 24h+9,*** and ***24h+15*** terms in branches and the L_8 chain usage factor shows that **every *24j+4, 24j+10,* and *24j+22* tower base term in all *24h+3, 24h+9,* and *24h+15* branches appears in an L_8 chain.** Since every L_8 chain ends in a Trunk Tower term, no L_8 chain can be part of a circular or unending Collatz sequence. Since all ***24h+3, 24h+9,*** and ***24h+15*** branches are part of some L_8 chain, the Collatz Structure containing all positive integers is a connected structure. Thus, every positive integer forms a Collatz sequence with unique terms terminating in the number one.

Appendix 1. A branch cannot have more than two consecutive even terms, and only the even terms *24m+4, 24m+10, 24m+16,* or *24m+22* are the immediate successors of odds terms.

6n+1 → 18n+4
If *n = 4j,   18n+4 = 72j+4   (24m+4,   m=3j) → 36j+2 → 18j+1.*
If *n = 4j+1, 18n+4 = 72j+22 (24m+22, m=3j) → 36j+11.*
If *n = 4j+2, 18n+4 = 72j+40 (24m+16, m=3j+1)* Last term in the branch.
If *n = 4j+3, 18n+4 = 72j+58 (24m+10, m=3j+2) →36j+29*
6n+3 → 18n+10.
If *n = 4j,   18n+10 = 72j+10 (24m+10, m=3j) → 36j+5.*
If *n = 4j+1, 18n+10 = 72j+28 (24m+4,   m=3j+1) → 36j+14 → 18j+7*
If *n = 4j+2, 18n+10 = 72j+46 (24m+22, m=3j+1) → 36j+23.*
If *n = 4j+3, 18n+10 = 72j+64 (24m+16, m=3j+2)* Last term in the branch.
6n+5 → 18n+16.
If *n = 4j,   18n+16 = 72j+16 (24m+16, m=3j)* Last term in the branch.
If *n = 4j+1, 18n+16 = 72j+34 (24m+10, m=3j+1) → 36j+17.*
If *n = 4j+2, 18n+16 = 72j+52 (24m+4,   m=3j+2) → 36j+26 → 18j+13.*
If *n = 4j+3, 18n+16 = 72j+70 (24m+22, m=3j+2) → 36j+35.*

Appendix 2. **Collatz structure details.**

Groups of similar Collatz sequence segments. If a Collatz sequence segment has a first term $\boldsymbol{a}$ and a last term $\boldsymbol{b}$ with $\boldsymbol{r}$, $2j+1 \to 6j+4$ and $\boldsymbol{s}$ divisions by two, there is a series of Collatz sequence segments containing the same number of terms and the same number of adjoining $\boldsymbol{L_8}$ of the same size and structure with a first term $a+(p-1)(24)(2^s)$ and last term $b+(p-1)(24)(3^r)$, $p=1,2,3\ldots$

The average branch binary series length: $3r=(1)(3/4)+(2)(9/16)+(3)(27/64)+\ldots$ $3r-(3)(3/4)r=3$, $r=4$. The binary series usage factor is three. Three lengths are being calculated. $3/4$ is the proportion of length one. $9/16$ of length two…Multiply the equation by $3/4$ and subtract. $3r-(3)(3/4)r=3/4+9/16+\ldots.=3$.

The average branch binary series sum: $((2,1,1,1)+(2,2,1,1)+(2,1,1,1))/3=(5+6+5)/3=5.333\ldots$ There are twice as many binary series components with one division by two $24j+10$ (1), $24j+22$ (1) than there are components with two divisions by two $24j+4$ (2). Three binary series of length four with twice as many ***1's*** as ***2's*** make up the computation.

Calculating the decrease in term size for $\boldsymbol{L_8}$ with the fewest $\boldsymbol{24k+16}$ terms.
$2/3\ (1-1/3)$ of the branches in a tower have binary series of length one or more. $4/9\ (1-1/3-2/9)$ have binary series of length two or more. The geometric series terms are increased by $3/2$ to base the calculation on the branches that have binary series. The average length of the L_8 binary series is:

$(1)(3/2)(2/3)+(2)(3/2)(4/9)+(3)(3/2)(8/27)+\ldots$

$(1+(2)(2/3)+(3)(4/9)+\ldots-(2/3)(1+(2)(2/3)+(3)(4/9)+\ldots)=1+2/3+4/9+\ldots=3 \quad (3)(3)=9$

Adjusting the proportion of branches with binary series from three to one. $9/3=3$.
The average $\boldsymbol{L_8}$ binary series sum is $(1,1,2)=4$.

$1/3$ of all branches have no binary series. The average number of divisions by two to reach the tower base term is $2+4+2=2.67$. Let $2j+1\to 6j+4$ be represented by an increase of 1.56 multiples of two. The average decrease in $\boldsymbol{L_8}$ term values is $-2.67-2+1.56-1+1.56-1+1.56=-2$. The ratio between the initial $24j+16$ term in an $\boldsymbol{L_8}$ with minimum number of tower terms and the last $24j+16$ term is on average $4/1$.

A circular sequence $1\to 4\to 2\to 1$ can be used to generate a sequence of arbitrary length with the same number and positions of $2j+1\to 6j+4$ and divisions by two. The binary series of length s is $(2,2,2,\ldots)$
$1+(2^{2s})(24)(p-1)$ is the beginning term and $1+(3^s)(24)(p-1)$ end term.
For $s=3$, $p=2$, $1537\to 4612\to 2306\to 1153\to 3460\to 1730\to 865\to 2596\to 1298\to 649$.

$\boldsymbol{24k+16}$ first term sequence segments

s=1 2 3 4 5 6 $(2^{s-1}-1)(24)+16+(p-1)(24)(2^s)$ The binary series is $(1,1,1,\ldots)$ The length $r=s-3$.

k=0	1	3	7	15	31
2	5	11	23	47	95
4	9	19	39	79	159

first term → last term					last term formula	
16 →8	40→10	88→11	184→35	376→107	$s=1,2,3$	$8,10,11+(24)(p-1)$
64→32	136→34	280→35	528→107	1144→323	$s\geq 4$	$11+s=4$ to $m\sum(24)(3^{s-4})+(24)(3^{s-3})(p-1)$

Thanks for your interest in this paper. If you wish to make comments send them to Jim Rock at collatz3106@gmail.com.

© 2018 James Edwin Rock.
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.