

# Partitioning the Positive Integers with the Collatz Algorithm

## Jim Rock

**Abstract.** Collatz sequences are formed by applying the Collatz algorithm to a 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. Eventually you get back to one. The 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.

**History of the Collatz conjecture.** The 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 Erdős said about the Collatz conjecture: "Mathematics may not be ready for such problems." [https://en.wikipedia.org/wiki/Collatz\\_conjecture](https://en.wikipedia.org/wiki/Collatz_conjecture)

**Describing the Collatz Structure.** The Collatz Structure is made of up branches and towers (defined in the next two paragraphs).  $m \rightarrow n$  indicates that  $n$  is the successor of  $m$  in a Collatz sequence.  $m \leftarrow n$  indicates that  $m$  is the predecessor of  $n$ .

**Describing a branch.** Terms of the form  $6n+3$  ( $24h+3$ ,  $24h+9$ ,  $24h+15$ , and  $24h+21$ ,  $h \geq 0$ ) always 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 ( $24k+16=4^j a$ ,  $a=24m+4$ ,  $24m+10$ , or  $24m+22$ ,  $j \geq 1$ ,  $m \geq 0$ ) appears. There can be no more than two consecutive even terms in a branch (appendix 1) so  $24k+16$  terms appear at the end of a branch.  $9 \rightarrow 28 \rightarrow 14 \rightarrow 7 \rightarrow 22 \rightarrow 11 \rightarrow 34 \rightarrow 17 \rightarrow 52 \rightarrow 26 \rightarrow 13 \rightarrow 40$ . There are no unending branches, as explained below.

**Describing a tower.** A tower is comprised of a base term ( $a=24m+4$ ,  $24m+10$  or  $24m+22$ ,  $m \geq 0$ ) and  $24k+16=4^j a$  terms,  $j=1,2,3 \dots$ . Let  $24k_n+16$  be the  $n$ -th term within a tower.  
 $24k_n+16 \rightarrow 12k_n+8 \rightarrow 6k_n+4$ .  $k_n = 4k_{n-1}+2$ ,  $6k_n+4 = 24k_{n-1}+16 \rightarrow 12k_{n-1}+8 \rightarrow 6k_{n-1}+4 \dots$   
 $24k_2+16 \rightarrow 12k_2+8 \rightarrow 6k_2+4$ .  $k_2 = 4k_1+2$ ,  $6k_2+4 = 24k_1+16 \rightarrow 12k_1+8 \rightarrow 6k_1+4$ .  
 If  $k_1=4m$ ,  $6k_1+4 = 24m+4$  or  $k_1=4m+1$ ,  $6k_1+4=24m+10$ . If  $k_1=4m+3$ ,  $6k_1+4=24m+22$ .

**Constructing the Collatz Structure.** The Collatz Structure starts with the Trunk Tower. Each  $(4^j)(4)$ ,  $j=1,2,3 \dots$  Trunk Tower term is the last term in a branch. At every  $a=24m+4$ ,  $24m+10$ , and  $24m+22$  base term in these Trunk Tower branches is a  $4^j a$ ,  $j=1,2,3 \dots$  secondary tower. Each  $4^j a$  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  $4^j a$ ,  $j=1,2,3 \dots$  secondary tower. Each  $4^j a$  is the last term in a branch. This process is repeated indefinitely. The position of all other terms in the Collatz Structure is explained below.

### Defining the "binary series" of a branch.

**The binary series is used to show that all  $24h+3$ ,  $24h+9$ , and  $24h+15$  terms appear in branches.**

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 the even terms  $24m+4$  (2),  $24m+10$  (1), or  $24m+22$  (1) giving odd terms (see appendix 1 for details) and a last term  $24k+16$ . 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  $24h+15+(p-1)(24)(2^s)$ ,  $p=1,2,3 \dots$   $h \geq 0$ . All groups end with  $24k+16+(p-1)(24)(3^{s+1})$ ,  $p=1,2,3 \dots$   $k \geq 0$ .  $24h+21$  has no binary series. However, there are a group of branches that begin with  $24h+21+(p-1)(24)$  followed immediately by  $(24)(3h+2)+16+(p-1)(3)(24)$ .

**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 first branch is 9, 28, 14, 7, 22, 11, 34, 17, 52, 26, 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.

The first term sequence is  $9+(p-1)(24)(2^6)$  9, 1545, 3081...  
The last term sequence is  $40+(p-1)(24)(3^3)$  40, 5872, 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 gives two terms that differ by  $(3)(p)(24)(2^s)$ . The terms are divisible by the same power of two. 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  $(p)(24)(2^s)$  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  $24h+15+(p-1)(24)(2^s)$ ,  $p=1,2,3... h \geq 0$ .

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

All branches have different first and last terms, since (as shown below) in a branch each term's predecessor and successor are unique. 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;  $(24h+3+(p-1)(24)(2^s) p=1,2,3...)$   $1/2^s$  proportion of the sequence terms.  $h < 2^s$  for the first term in the sequence whose binary series sum is  $s$ .

$24h+3 \rightarrow 72h+10 \rightarrow 36h+5 \rightarrow 108h+16$ , shows the first two  $24h+3$  binary series are  $(1)$  and  $(1,2)$ . All other binary series begin with  $(1,2,...)$ . The proportion of  $24h+3$  terms with binary series  $(1)$  is  $1/2^1$ .

The proportion of  $24h+3$  terms with binary series  $(1,2)$  is  $1/2^3$ . Assume the proportion of  $24h+3$  terms with binary series length  $r \geq 2$  is  $3^{r-2}/2^{2r-1}$ . The  $r+1$  position of every binary series of that length contains either  $(1)$  one division by two 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  $r \geq 2$  gives a total proportion of one.  $1/2+1/8+3/32+9/128+... = 1/2+(1/8)/(1-3/4)=1$ .

The first  $24h+9$  binary series is  $(2)$ . All other binary series begin with  $(2,...)$ . The proportion of  $24h+9$  terms with binary series  $(2)$  is  $1/2^2$ . If the proportion of  $24h+9$  terms with binary series length  $r \geq 1$  is  $3^{r-1}/2^{2r}$ , the proportion 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 a total proportion of one.  $1/4+3/16+9/64+... = (1/4)/(1-3/4)=1$ .

The first  $24h+15$  binary series is  $(1,1)$ . All other binary series begin with  $(1,1,...)$ . The proportion of  $24h+15$  terms with binary series  $(1,1)$  is  $1/2^2$ . If the proportion of  $24h+15$  terms with binary series of length  $r \geq 2$  is  $3^{r-2}/2^{2r-2}$ , the proportion with length  $r+1$  is  $(1/2)(3^{r-2}/2^{2r-2})+(1/4)(3^{r-2}/2^{2r-2})=3^{r-1}/2^{2(r+1)-2}$ . For length  $r \geq 2$  the total proportion is one.  $1/4+3/16+9/64+... = (1/4)/(1-3/4)=1$ .

**Calculating the proportion of  $24k+16$  terms in branches.** Put all  $24k+16, k=0,1,2... terms in an ascending sequence. The proportion of all terms with the same binary series is  $1/3^{r+1}$  of the terms in the sequence.$

$24k+16+(p-1)(24)(3^{r+1}) p=1,2,3... k < 3^{r+1}$  for the first term in the sequence whose binary series has  $r$  terms. The last terms of the  $2^r$  branches with binary series of length  $r$  are  $2^r/3^{r+1}$  of all terms in the sequence.

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

**All terms of the form  $24h+3, 24h+9$ , and  $24h+15$  are in branches, and there are branch binary series with all  $2^r$  possible combinations of 1's and 2's for every value of  $r$ . There are no unending branches. Since all  $24h+16$  terms appear in branches, all  $24h+21 \rightarrow 24(3h+2)+16$  also appear in branches.**

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

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 binary series and tower binary series.**

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  $24h+9...$ ,  $24h+15...$ , and  $24h+21...$  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 branches with the same binary series occur in intervals of  $3^{r+1}$ .  $2^r/3^{r+1}$  is the proportion of the  $2^r$  tower branches with a binary series of length  $r$ . For  $r \geq 0$   $1/3+2/9+4/27...=1$  is the total proportion. **There are tower branches with binary series of all  $2^r$  combinations of  $r$  1's and 2's for all  $r$ . The first branch with a binary series of length  $r$  comes within the first  $3^{r+1}$  branches in the tower.**

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

**There are twelve even terms  $24m+2y$ ,  $y=0$  to 11 and three odd terms  $6n+1$ ,  $6n+3$ , and  $6n+5$ .**

$24n+4 \leftarrow 12n+2$  ( $24m+2$ ,  $24m+14$ )  $\leftarrow 6n+1$ .  $12n+10$  ( $24m+10$ ,  $24m+22$ )  $\leftarrow 6n+5$ .  $q$  is the successor of each even term  $2q$  in a branch. 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  $24m+22$ . 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  $24h+3$ ,  $24h+9$ ,  $24h+15$ , and  $24h+21$  term is 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 the last terms in branches, 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.

**Definition of an  $L_8$ , an  $L_8$  chain binary series, and its usage factor.** 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 based on 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. The binary series of one tower base term is 1, or 2. The usage factor is  $1/2^1+1/2^2=3/4$ . The binary series of two tower base terms is 11, 12, 21, 22. The usage factor is  $1/4+1/8+1/8+1/16=9/16$ . If the usage factor of a binary series with  $r$  tower base terms is  $3^r/4^r$ , for  $r+1$  tower base terms 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.

**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.**

**$6n+1 \rightarrow 18n+4$**

If  $n = 4j$ ,  $18n+4 = 72j+4$  ( $24m+4$ ,  $m=3j$ )  $\rightarrow 36j+2 \rightarrow 18j+1$ .

If  $n = 4j+1$ ,  $18n+4 = 72j+22$  ( $24m+22$ ,  $m=3j$ )  $\rightarrow 36j+11$ .

If  $n = 4j+2$ ,  $18n+4 = 72j+40$  ( $24m+16$ ,  $m=3j+1$ ) Last term in the branch.

If  $n = 4j+3$ ,  $18n+4 = 72j+58$  ( $24m+10$ ,  $m=3j+2$ )  $\rightarrow 36j+29$

**$6n+3 \rightarrow 18n+10$ .**

If  $n = 4j$ ,  $18n+10 = 72j+10$  ( $24m+10$ ,  $m=3j$ )  $\rightarrow 36j+5$ .

If  $n = 4j+1$ ,  $18n+10 = 72j+28$  ( $24m+4$ ,  $m=3j+1$ )  $\rightarrow 36j+14 \rightarrow 18j+7$

If  $n = 4j+2$ ,  $18n+10 = 72j+46$  ( $24m+22$ ,  $m=3j+1$ )  $\rightarrow 36j+23$ .

If  $n = 4j+3$ ,  $18n+10 = 72j+64$  ( $24m+16$ ,  $m=3j+2$ ) Last term in the branch.

**$6n+5 \rightarrow 18n+16$ .**

If  $n = 4j$ ,  $18n+16 = 72j+16$  ( $24m+16$ ,  $m=3j$ ) Last term in the branch.

If  $n = 4j+1$ ,  $18n+16 = 72j+34$  ( $24m+10$ ,  $m=3j+1$ )  $\rightarrow 36j+17$ .

If  $n = 4j+2$ ,  $18n+16 = 72j+52$  ( $24m+4$ ,  $m=3j+2$ )  $\rightarrow 36j+26 \rightarrow 18j+13$ .

If  $n = 4j+3$ ,  $18n+16 = 72j+70$  ( $24m+22$ ,  $m=3j+2$ )  $\rightarrow 36j+35$ .

**Appendix 2. Collatz structure details.**

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

**The average branch binary series length:**  $3r=(1)(3/4)+(2)(9/16)+(3)(27/64)+...$   $3r - (3)(3/4)r = 3$ ,  $r=4$ .  
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.  $3r - (3)(3/4)r = 3/4 + 9/16 + .... = 3$ .

**The average branch binary series sum:**  $((2,1,1,1)+(2,2,1,1)+(2,1,1,1))/3 = (5+6+5)/3 = 5.333...$

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  $L_8$  with the fewest  $24k+16$  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:

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

$(1+(2)(2/3)+(3)(4/9)+... - (2/3)(1+(2)(2/3)+(3)(4/9)+...))=1+2/3+4/9+...=3$   $(3)(3)=9$

Adjusting the proportion of branches with binary series from three to one.  $9/3=3$ .

The average  $L_8$  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  $2j+1 \rightarrow 6j+4$  be represented by an increase of  $1.56$  multiples of two. The average decrease in  $L_8$  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  $L_8$  with minimum number of tower terms and the last  $24j+16$  term is on average  $4/1$ .

**A circular sequence  $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$  can be used to generate a sequence of arbitrary length with the same number and positions of  $2j+1 \rightarrow 6j+4$  and divisions by two. The binary series of length  $s$  is  $(2,2,2,...)$**

$1+(2^{2s})(24)(p-1)$  is the beginning term and  $1+(3^s)(24)(p-1)$  end term.

For  $s=3$ ,  $p=2$ ,  $1537 \rightarrow 4612 \rightarrow 2306 \rightarrow 1153 \rightarrow 3460 \rightarrow 1730 \rightarrow 865 \rightarrow 2596 \rightarrow 1298 \rightarrow 649$ .

**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,...) 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](mailto:collatz3106@gmail.com).**

© 2018 James Edwin Rock.

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).