

# A Solution to the $3x + 1$ Problem

David C. Kay

Weaverville, NC, USA

Email: dck.gm09@yahoo.com

**How to cite this paper:** Kay, D.C. (2023) A Solution to the  $3x + 1$  Problem. *American Journal of Computational Mathematics*, 13, 371-377.

<https://doi.org/10.4236/ajcm.2023.132020>

**Received:** March 3, 2023

**Accepted:** June 27, 2023

**Published:** June 30, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

The author makes the claim to have a solution for the famous  $3x + 1$  problem. The key to its solution is a special proof that the term  $(3 + \varepsilon)^r$  is a non-integer, as well as the use of properties of extremal points of a Collatz sequence.

## Keywords

Collatz Sequences, Extreme Points, Pre-Extreme Sequence

---

## 1. Introduction

The article by Kay “Collatz Sequences and Characteristic Zero-One Strings: Progress on the  $3x + 1$  Problem” [1] provides the background needed here. Recall that a Collatz sequence  $C(n)$  generated by a positive integer  $n$  is defined by  $C(n) = \{n, T(n), T[T(n)], [[T[T^{\nu}]]], \dots\}$  for each integer  $n$ , where  $T(n) = n/2$  if  $n$  is even, and  $(3n + 1)/2$  if  $n$  is odd.

Solving the  $3x + 1$  problem means proving that all Collatz sequences contain the integer 1. This is a long-standing conjecture over 50 years old, that has stubbornly resisted solution. A number of prominent mathematicians even discouraged working on it. Paul Erdős commented “mathematics is not yet ready for this problem” (Lagarias [2]. p. 31).

## 2. Previous Research

The volume of research on this problem is truly incredible. In [2] Lagarias has listed over 300 - 400 research articles written in the 1900’s and 2000’s—all devoted to solving the problem. Most research ended up including related conjectures as difficult as the original problem, and much was in the area of dynamic processes and ergodic theory. Our approach using strings is not related directly to dynamic processes or ergodic theory. The solution here involves only standard

techniques of algebra, calculus, and elementary number theory.

We make use of a limited amount of past research. In particular, Lagarias [2] has pointed out that the  $3x + 1$  conjecture has been verified by special computer programs for all positive integers up to  $n^* = 20 \times 2^{58} = 5.8 \times 10^{18}$  (more recently to  $n^* = 8.7 \times 10^{18}$ ).

### 3. Zero-One Strings

A zero-one string is an infinite ordered string of zeroes and ones, written in the form

$$L = 0^{s_0} 1^{r_1} 0^{s_1} 1^{r_2} 0^{s_2} \dots 1^{r_i} 0^{s_i} \dots \tag{1}$$

where, for each  $i$ ,  $r_i$  is the number of ones in the  $i$ th group of consecutive ones, and  $s_i$  the number of zeroes in the  $i$ th group of consecutive zeroes. A positive integer  $n$  is said to generate a zero-one string  $L = x_1, x_2, \dots, x_i, \dots$ , iff a Collatz sequence  $C(n) = \{n_1, n_2, \dots, n_i, \dots\}$  exists such that  $n_i = x_i \pmod{2}$  for all  $i$ .

As in [1], we refer to  $n_i$  as the pre-resultant of a finite Collatz sequence  $C = C(n)$  where  $n_i$  is the last odd member of  $C$  in a group to all odd members, and  $n_{i+1} = (3n_i + 1)/2$  the resultant. The total number of ones in a finite string  $L$  will be denoted  $r$ , the total number of zeroes  $s$ , and the length,  $k = r + s$  (not to be confused with the concept of length as used by others. The norm of  $L$  is defined as

$$\|L\| = 3^r / 2^k \tag{2}$$

It was established in [1] that a generator can be constructed for a given finite string  $0^{s_0} 1^{r_1} 0^{s_1} 1^{r_2} 0^{s_2} \dots 1^{r_d} 0^{s_d}$ ; infinite strings do not always have generators, such as

$$111\dots 1\dots, 000\dots 0111\dots \text{ and } 010010001\dots$$

### 4. Formulas for Resultants

The advantage of using strings is that explicit formulas for resultants can be readily established. For example, in [1] a general formula was given for the resultant of a string in terms of a generator. A formula we will use here, also proved in [1], is the following: suppose  $n$  and  $m$  are the generator and resultant of the string  $L = 1^{r_1} 0^{s_1} 1^{r_2} 0^{s_2} \dots 1^{r_i}$  and that the Collatz sequence corresponding to  $L$  is  $C = \{n, n_1, n_2, \dots, n_b, m\}$ . Then

$$m = \frac{n^r}{2^k} \prod_{i=1}^r \left( 3 + \frac{1}{n'_i} \right) \tag{3}$$

where the product is taken over all the odd members  $n'_i$  of  $C$ .

For example, consider the Collatz sequence

$$C = \{89, 134, 67, 101, 152, 76, 38, 19\}$$

corresponding to the string  $L = 1^1 0^1 1^2 0^3 1^1$ , with  $s = 4$ ,  $k = 8$ , and resultant  $m = (3 \times 19 + 1)/2 = 29$ . The odd members of  $C$  are 89, 67, 101, and 19. Substituting into (3) produces

$$\begin{aligned}
 m &= \frac{89}{2^8} \left(3 + \frac{1}{89}\right) \left(3 + \frac{1}{67}\right) \left(3 + \frac{1}{101}\right) \left(3 + \frac{1}{19}\right) \\
 &= \frac{89 \cdot 134 \cdot 101 \cdot 152 \cdot 29}{2^4 \cdot 89 \cdot 67 \cdot 101 \cdot 19} \\
 &= \frac{89 \cdot 134/2 \cdot 101 \cdot 152/2^3 \cdot 29}{1 \cdot 89 \cdot 67 \cdot 101 \cdot 19} \\
 &= 89 \frac{67 \cdot 101 \cdot 19 \cdot 29}{89 \cdot 67 \cdot 101 \cdot 19} \\
 &= 29
 \end{aligned}$$

Applying the function  $f(x) = (3 + x)^r$  where  $r$  is a positive integer and using the mean-value theorem, it can be shown that there exists a positive real number  $\varepsilon < 10^{-18}$  such that

$$\prod_{j=1}^r \left(3 + \frac{1}{n_j}\right)^r = (3 + \varepsilon)^r \tag{4}$$

It follows from (3) that if  $m$  and  $n$  are odd integers and equal the respective resultant and generator of a Collatz sequence  $C$ , then

$$m = n(3 + \varepsilon)^r / 2^k \tag{5}$$

(It was shown in [1] that  $\varepsilon$  is irrational for cycles.)

Observe that if a non-trivial cycle exists,  $m = n$  and (5) produces the necessary condition for cycles  $(3 + \varepsilon)^r = 2^k$ .

### 5. The Term $(3 + \varepsilon)^r$

We finally prove that the term  $(3 + \varepsilon)^r$  is a non-integer. Suppose there are values for the positive integer  $n$  for which  $(3 + \varepsilon)^n$  is a positive integer. Take  $n$  to be the least of these; it follows that  $(3 + \varepsilon)^{n-1}$  is a non-integer. It follows that  $(3 + \varepsilon)^m$  is a non-integer for all  $m \leq n - 1$ .

If exactly one of  $(3 + \varepsilon)^{n-1}$  and  $3 + \varepsilon$  is rational, then  $(3 + \varepsilon)^n$  is irrational, a contradiction, so we conclude that either both are rational (Case (1)) or both are irrational (Case (2)). In Case (1), take  $(3 + \varepsilon)^{n-1} = x/y$  and  $3 + \varepsilon = u/v$  (all integers), and with the fraction  $xu/yv$  reduced to lowest terms. Then the product  $xu/yv = N$  or  $xu = yvN$ . No prime divisor of  $xu$  can appear in  $yv$ , so ultimately we must have  $xu \mid N$ . Then there exists an integer  $N'$  such that  $N = xuN'$  and  $xu = yvN'$  or  $1 = yvN'$ . One obtains  $y = v = 1$  which implies the contradiction that both  $(3 + \varepsilon)^{n-1}$  and  $(3 + \varepsilon)$  are integers.

In Case (2) when both  $(3 + \varepsilon)^{n-1}$  and  $(3 + \varepsilon)$  are irrational, these two may be approximated arbitrarily closely by rationals, say by  $x/y$  and  $u/v$ . It follows that their product is approximated arbitrarily closely to the result, which is the integer  $N$ . Hence,  $xu/yv = N$  or  $xu = yvN$ . Assuming  $(xu, yv) = 1$  as before, we obtain a contradiction as before, finally proving that  $(3 + \varepsilon)^n$  cannot be an integer.

This leaves  $N =$  a non-integer, a contradiction. The contradictions prove that the term  $(3 + \varepsilon)^n$  is a non-integer for all integers  $n$ , leading to the desired result.

It follows that  $(3 + \varepsilon)^r \neq 2^k$  and we have the proof that non-trivial cycles do not

exist. It remains to prove that a Collatz sequence not containing 1 also does not exist.

### 6. Extreme Points of a Collatz Sequence

The graph of any Collatz sequence reveals the existence of a sub-sequence consisting of extreme points  $J = \{n_i\} \cup \{m_i\}$ ,  $\{n_i\}$  the minimal extreme points and  $\{m_i\}$  the maximal extreme points. These numbers obey the basic inequalities  $m_i > n_i$  for each  $i \geq 1$ . Furthermore, each  $n_i$  is the generator of a Collatz sequence of all odd integers.

The following result provides more information concerning extreme points.

**Theorem A.** There exists a unique sequence of positive integers  $\{t_i\}$  (to be called the pre-extremal sequence) such that for each  $i$ ,

$$n_i = 2^{t_i} t_i - 1 \quad \text{and} \quad m_i = 3^{t_i} t_i - 1 \tag{6}$$

*Proof:* Since  $m_i$  is the resultant of the Collatz sequence of  $r$  odd numbers,  $r \geq 1$ , the following identity can be routinely established by induction on  $r$ :

$$m_j = \lceil 3^r n_i + 3^r - 2^r \rceil / 2^r \tag{7}$$

or

$$2^r (m_i + 1) = 3^r (n_i + 1) \tag{8}$$

Thus  $m_i + 1$  is divisible by  $3^r$  and  $n_i + 1$  is divisible by  $2^r$  and for certain positive integers  $p_1$  and  $p_2$ ,

$$m_i + 1 = 3^r p_1 \quad \text{and} \quad n_i + 1 = 2^r p_2$$

But from (6.3),  $2^r(3^r p_1) = 3^r(2^r p_2)$  and therefore  $p_1 = p_2 = t_i \geq 1$  for some unique  $t_i$ . Hence it follows that

$$n_i = 2^{t_i} t_i - 1 \quad \text{and} \quad m_i = 3^{t_i} t_i - 1.$$

as desired.

For a numerical example, consider the Collatz sequence  $C$  generated by  $n = 195$  and corresponding string  $L = 1^2 0^3 1^3 0^2 1^4 0^1 1^1$  written as  $L_1 K_1 L_2 K_2 L_3 K_3 L_4$ , where  $L_1 = 11$ ,  $K_1 = 000$ ,  $L_2 = 111$ ,  $K_2 = 00$ ,  $L_3 = 1111$ ,  $K_3 = 0$ , and  $L_4 = 1$ . Note that 195 is uniquely preceded by 390, so  $J$  begins with 195, a local minimum. The members of  $C$  can be written as follows, with the members of  $J$  in bold type:

$C(195) = \{\mathbf{195}, 293, \mathbf{440}, 220, 110, \mathbf{55}, 83, 125, \mathbf{188}, 94, \mathbf{47}, 71, 107, 161, \mathbf{242}, \mathbf{121}\}$   
and we have

$$\mathbf{195}(11)\mathbf{440}(000)\mathbf{55}(111)\mathbf{188}(00)\mathbf{47}(1111)\mathbf{242}(0)\mathbf{121}(1)$$

with

$$\begin{aligned} 195 &= 2^2(49) - 1 \\ 440 &= 3^2(49) - 1 \\ 55 &= 2^3(7) - 1 \\ 188 &= 3^3(7) - 1 \\ 47 &= 2^4(3) - 1 \\ 242 &= 3^4(3) - 1 \end{aligned}$$

$$121 = 2^1(61) - 1$$

In this case, the pre-extremal sequence is {49, 7, 3, 61}.

We observe the following algorithm that generates the sequence  $\{t_i\}$  Since  $n_i$  is odd, add 1 to make it even, and begin dividing by 2 until an odd number  $u$  appears. This odd number  $u$  can be taken as  $t_i$  since  $n_i + 1 = 2^v u$  or  $n_i = 2^v u - 1$ . The following result is also clear from this analysis:

**Theorem B.** The pre-extremal sequence consists of all odd integers.

### 7. A Formula for $t_{i+1}$ in Terms of $t_1$

Theorem A leads to a formula for  $t_{i+1}$  in terms of  $t_1$ . A preliminary result is needed first.

**Theorem C.** Let  $L = 1^n 0^{s_i}$  and  $n_i L m_i$ . Then

$$t_{i+1} = \frac{3^{n_i} t_i + 2^{s_i} - 1}{2^{n_i + s_i}} \tag{9}$$

*Proof:* By hypothesis,  $n_i L_i K_i m_i$ , and from (7) we obtain

$$m_i = \frac{3^{n_i} n_j + 3^{n_i} - 2^{n_i}}{2^{n_i + s_i}} \tag{10}$$

Since  $m_i$  is a generator of  $L_{i+1}$  and  $n_i$  is a generator of  $L_b$  by Theorem A

$$m_i = 2^{n_i+1} t_{i+1} - 1 \quad \text{and} \quad n_i = 2^{n_i} t_i - 1 \tag{11}$$

(10) becomes

$$2^{n_i+1} t_{i+1} - 1 = \frac{3^{n_i} (2^{n_i} t_i - 1) + 3^{n_i} - 2^{n_i}}{2^{n_i + s_i}}$$

$$2^{n_i+1} t_{i+1} = \frac{3^{n_i} (2^{n_i} t_i) - 2^{n_i}}{2^{n_i + s_i}} + 1 = \frac{3^{n_i} t_i - 1 + 2^{s_i}}{2^{s_i}}$$

which is equivalent to the desired result.

Define  $U_i = 3^{n_i} / 2^{n_i + s_i}$  and  $V_i = (2^{s_i} - 1) / 2^{n_i + s_i}$ . Then (9) becomes for all  $i$

$$t_{i+1} = U_i t_i + V_i$$

By successive substitution we obtain

$$t_{i+1} = \alpha_i t_1 + \beta_i \tag{12}$$

where

$$\alpha_i = U_1 U_2 \dots U_i \quad \text{and} \quad \beta_i = U_2 U_3 \dots U_i V_1 + U_3 U_4 \dots U_i V_2 + \dots + U_i V_{i-1} + V_i \tag{13}$$

A lemma will be needed later, based on these results.

**Lemma 1.** In the special case  $i = 1$ ,  $2^k \alpha_1 = 2^k U_1 = 2^{n_1} 3^{n_1}$  and  $2^k \beta_1 = 2^k V_1 = 2^{n_1} (2^{s_1} - 1)$ .

The proof is a routine result in algebra and will be omitted.

### 8. Divergent Strings

A divergent string is a string of zeros and ones not ending in 1010...

**Theorem D.** A divergent string has no generator.

*Proof:* Suppose  $L$  is a divergent string with a generator  $n$ , and with corresponding Collatz sequence  $C = C(n)$ . Then  $C$  cannot contain a 1 or  $L$  ends in 1010... Every member of  $C$  is then greater than  $n^*$  and there exists a least generator  $n' > n^*$  of  $C$ . Hence,  $n' = n_1$  is an (odd) minimal extreme point of  $C$ , followed by an (even) maximal extreme point  $m_1$ , which is the resultant of  $r_1$  odd members of  $C$  and generator of  $s_1$  even members of  $C$ , followed by  $n_2$ , generator of  $r_2$  odd members of  $C$ , ending with  $m_2$ , a maximal extreme point (even), the resultant of  $r_2$  odd members of  $C$ . There are a total of  $r_1 + s_1 + r_2 = k$  members of  $C$  listed here (in other words a smaller Collatz sequence  $C'$ , of which  $n_1$  is a generator and  $m_2$  the resultant). By (5)

$$m_2 = n_1(3 + \varepsilon)^r / 2^k$$

where  $r = r_1 + r_2$ . It will be convenient to let  $g = (3 + \varepsilon)^r - 3^r$ . We then obtain

$$2^k m_2 = n_1(g + 3^r)$$

By Theorem A,  $m_2 = 3^{r_2} t_2 - 1$  and  $n_1 = 2^{r_1} t_1 - 1$ . Thus

$$2^k (3^{r_2} t_2 - 1) = (2^{r_1} t_1 - 1)(g + 3^r)$$

$$2^k 3^{r_2} t_2 - 2^k = 2^{r_1} g t_1 + 2^{r_1} 3^r t_1 - g - 3^r$$

By (12),

$$2^k 3^{r_2} (\alpha_i t_1 + \beta_i) - 2^k = 2^{r_1} g t_1 + 2^{r_1} 3^r t_1 - g - 3^r$$

Using the Lemma we obtain

$$2^{r_1} 3^r t_1 + 2^{r_1} 3^r (2^{s_1} - 1) - 2^k = 2^{r_1} g t_1 + 2^{r_1} 3^r t_1 - g - 3^r$$

which reduces to

$$2^{r_1} 3^r (2^{s_1} - 1) - 2^k = g(2^{r_1} - 1) - 3^r$$

which can be put in the form

$$(3 + \varepsilon)^r n_1 = 2^{r_1} 3^r (2^{s_1} - 1) - 2^k + 3^r (n_1 + 1)$$

Since  $n_1$  is odd, the right side is an even integer, from which it follows that for some integer  $N$ ,

$$(3 + \varepsilon)^r n_1 = 2N$$

Again,  $n_1$  is not even so cannot contain 2 as a factor, so we have for some integer  $N'$

$$(3 + \varepsilon)^r = 2N'$$

which violates the result of Section 5, showing that  $L$  has no generator.

### 9. Conclusion

The  $3x + 1$  problem has been proved in the affirmative, establishing that a finite Collatz sequence always converges to 1.

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- [1] Kay, D. (2021) Collatz Sequences and Characteristic Zero-One Strings: Progress on the  $3x + 1$  Problem. *American Journal of Computational Mathematics*, **11**, 226-239. <https://doi.org/10.4236/ajcm.2021.113015>
- [2] Lagarias, J.C. (2010) *The Ultimate Challenge: The  $3x + 1$  Problem*. American Mathematical Society, Providence. <https://doi.org/10.1090/mbk/078>