Catchments and Creeks

P.O. Box 8085
Bargara, Queensland, 4670
Australia.

16 August 2023

A Solution to the Collatz Conjecture

1. Series functions:

Choose any positive whole number as a starting value, then apply the following steps:

- An Even number is followed by a 50% reduction in that number (i.e. 0.5X).
- An Odd number is multiplied by 3, then 1 is added (i.e. 3X+1).

Example (odd values shown in 'red'):

7—22—11—34—17—52—26—13—40—20—10—5—16—8—4—2—1(—4—2—1)

2. Task:

Demonstrate that a value of '1' is an inevitable outcome of this series, independent of the starting value.

3. What we know:

We know that the series does not approach a value of '1' if a starting value of infinity is adopted, even if we were to accept that infinity is an even number.

We know that if the starting value is taken as infinity, then the number of iterations (steps) in the series will also be infinite.

We can conclude that as the starting value of the series approaches infinity, the number of iterations in the series also approaches infinity.

We know that an Odd value must be followed by an Even value, but consecutive Even values can occur.

We know that as the number of iterations in the series increases, the potential for consecutive Even values also increases. In general, the probability of consecutive Even values is given by $(100\%)*(0.5)^n$ where 'n' = the number of consecutive Even values; however, the series would need to incorporate at least 'n' iterations.

We know that based on probability, any series is likely to consist of approximately 2/3 Even values and 1/3 Odd values. However, technically there could be an even number of Odd and Even values, as well as a series containing just one Odd value '1' (i.e. the halving series: 16, 8, 4, 2, 1). In the above example there are 6 Odd values (35%) and 11 Even values (65%).

We know that there can be more than one pathway to most values, but there is only one pathway from any value to a value of '1'. Consequently, every value that appears within a successful series must also satisfy the conjecture.

4. The difficulty in finding a numerical answer to this conjecture

The fact that any numerical solution must work for all starting values up to, but not including, infinity, causes difficulties in the formulation of a numerical solution. The reasons for such difficulties relate to the fact that independent of the starting value, the final stage of this numerical series always incorporates a condition where an Odd value cannot be generated until a value of '1' is achieved.

This means that in this final stage, the series reduces to a single numerical function, that being: *'An Even number is followed by a 50% reduction in that number'*. This means that no single numerical formula can address the entire series. In effect, it is a series in two parts.

The first part of the series incorporates both Odd and Even functions, while the second part of the series utilises only the Even number function.

The series must then follow a combination of the three possible numerical sequences:

- Odd—Even
- Even—Odd
- Even—Even (which is the only sequence utilised at the end of the series)

This means that after two iterations there are three possible outcomes:

- Odd—Even: X → (3X + 1) → 0.5(3X + 1)
- Even—Odd: X → (0.5X) → 3(0.5X) + 1
- Even—Even: X → (0.5X) → 0.5(0.5X)

Thus the likely outcome after two iterations would be:

- (1/3)(0.5(3X + 1) + (1/3)[3(0.5X) + 1) + (1/3)(0.25X)]
- (1/2)X + 1/3 + (1/2)X + 1/3 + (1/12)X
- (13/12)X + 2/3, which is greater than the starting value of 'X'.

Thus, based solely on probability, the likely outcome after any two iterations is an increase in the numerical value.

If someone were to analyse how many iterations it would take each starting value to reach a value of '1', then the analysis would be complicated by the fact that as the starting value approaches infinity, the number of iterations would also approach infinity.

However, we can avoid such difficulties by focusing on an analysis of the number of iterations required to generate a value that has already been demonstrated to satisfy the conjecture, which could include any value less than the starting value. This would now be an easier series to analyse because the number of iterations in such a truncated series would not approach infinity as the starting value approached infinity.

This approach is based on the concept that you don't need to prove what has already been proven. For example, testing for a starting value of '7' also tests for starting values of 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2 and 1.

5. Analysis of the numerical series:

The numerical series is effectively a mathematical interpretation of the game 'Snakes and Ladders'. In this numerical version of the game there is no winning 'top square' because the game extends to infinity. Instead, the game finishes when the numerical series collapses to a value of '1'.

Landing on an Odd number brings you to a climbing 'Ladder', but every ladder in this game climbs only one step. Landing on an Even number brings you to a descending 'Snake', which can descend multiple steps (iterations).

Importantly there are some special numbers within the numerical series that connect to special Snakes. We can refer to these values as gold, silver and bronze numbers. Landing on a gold or silver number traps the player on a Snake from which the player cannot escape, and which carries the player swiftly to a value of '1'. In fact all numbers trap a player on a pathway from which they cannot escape, but only gold and silver numbers carry a player swiftly to a value of '1'.

Figure 1 shows a schematic representation of the numerical series (real series chains have not been used). An analysis of over 200,000 series chains has shown that more than 90% of the series will arrive at the value '16' as the first gold number.

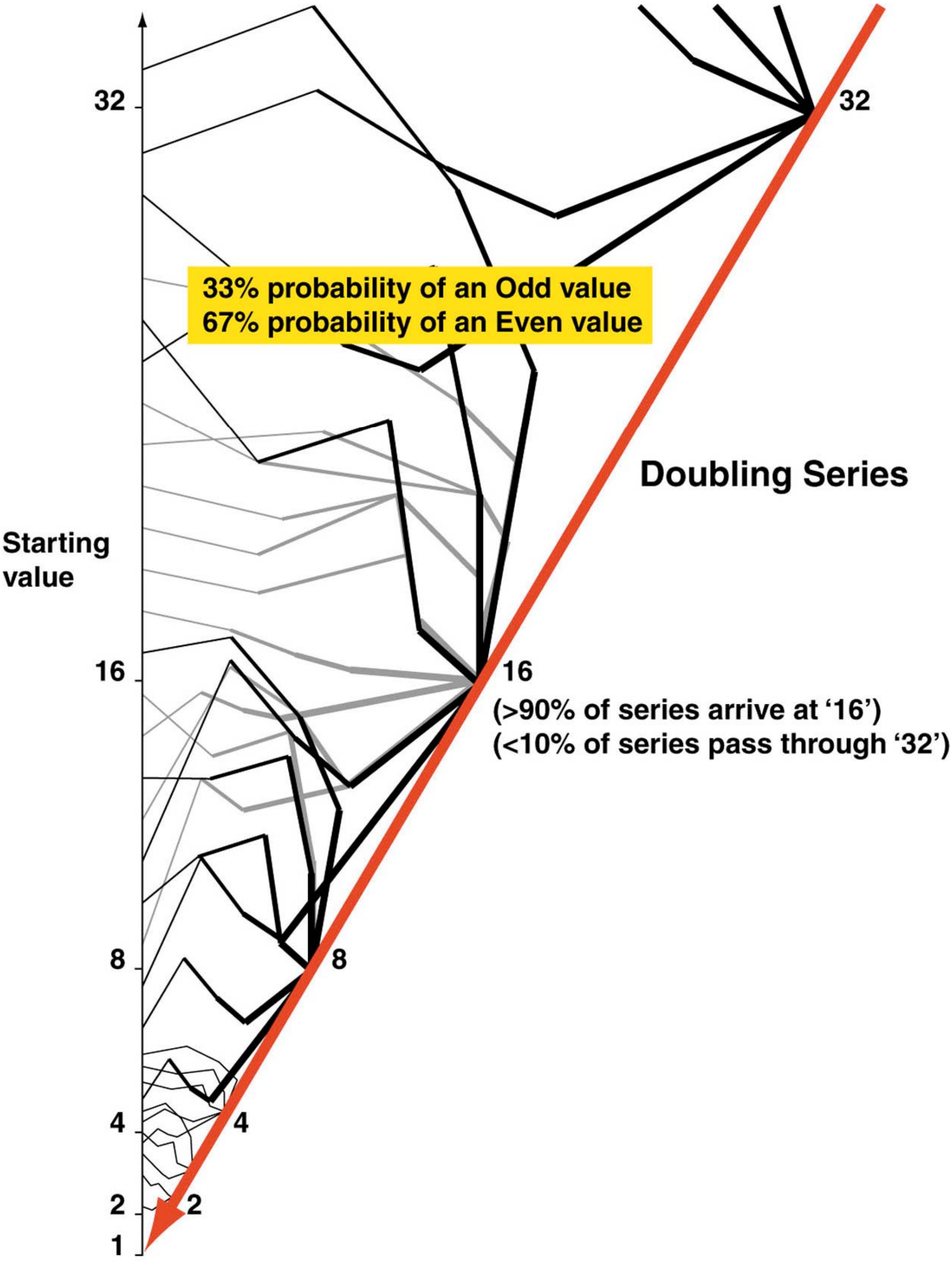

Figure 1: Schematic representation of the numerical series

A gold number will carry the player directly to the numerical value of '1'. Gold numbers consist of the doubling sequence starting with the value of '1':

Gold numbers: 1, 2, 4, 8, **16**, 32, n, 2n, 2(2n), 2(2(2n)), etc.

Example: 7—22—11—34—17—52—26—13—40—20—10—5—**16**—8—4—2—1

A silver number will carry the player through a series of Even values until they reach an Odd number from where the player will be just one iteration away from a gold number. Silver numbers consist of the doubling sequence starting with any Odd number that directly connects to a gold number.

Example: 7—22—11—34—17—52—26—13—**40**—**20**—**10**—**5**—16—8—4—2—1

Table 1: Gold and silver numbers

Gold	Function	Silver numbers					
16	(16-1)/3=	**5**	**10**	**20**	**40**	80	etc.
64	(64-1)/3=	21	42	84	168	336	etc.
256	(256-1)/3=	85	170	340	680	1360	etc.
1024	(1024-1)/3=	n=341	2n	2(2n)	2(2(2n))	2(2(2(2n)))	etc.
etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.

Landing on a bronze number will carry the player down a series of two or more Even values before arriving at an Odd value. Bronze numbers consist of the doubling sequence of any non-gold or non-silver number.

Example bronze numbers: **13**, **26**, **52**, 104, 208, 416, n, 2n, 2(2n), 2(2(2n)), etc.

Example sequence: 7—22—11—34—17—**52**—**26**—**13**—40—20—10—5—16—8—4—2—1

6. Demonstrating the inevitable outcome of the value '1'

The following proof is based on a three-part test:

- if all starting values up to a given value can be demonstrated to satisfy the conjecture, and;
- if it can be demonstrated that every series, with the exception of a starting value of '1', will eventually generate a value that is less than the starting value, then
- it can be concluded that all starting values from '1' to infinity, but excluding infinity, will satisfy the conjecture.

This process means that we do not need to demonstrate that all starting values will collapse to a value of '1'. Instead, we only need to demonstrate that all starting values (excluding '1') will collapse to a value less than the starting value. However, the second part of this test requires us to prove that a series-loop could not be generated above the 1–4–2–1 loop.

Case 1:

If all the values up to any given Even value can be proven to satisfy the conjecture, then that Even value (say, 2X) **must** also satisfy the conjecture because the first iteration would result in a value of (X), which has already been proven to satisfy the conjecture.

Case 2:

If we now consider the Odd number immediately greater than that Even number (i.e. 2X+1), and we let (X) be an Even number, then the following steps would occur:

- (2X+1) → 3(2X+1)+1 = (6X+4) which is an Even number, thus the next step:
- (6X+4) → 0.5(6X+4) = (3X+2) which is an Even number, thus the next step:
- (3X+2) → 0.5(3X+2) = (1.5X+1) which is less than the starting value of (2X+1).

Therefore, if all the numbers up to such an Odd number (2X+1) can be proven to satisfy the conjecture, and where (X) is an Even number, then that Odd number must also satisfy the conjecture because in two iterations a value less than that Odd number will be achieved.

Case 3:

If we take the above situation, but we let (X) be an Odd number, then the following steps would occur:

- (2X+1) → 3(2X+1)+1 = (6X+4) which is an Even number, thus the next step:
- (6X+4) → 0.5(6X+4) = (3X+2) which is an Odd number, thus the next step:
- (3X+2) → 3(3X+2)+1 = (9X+7) which is an Even number, thus the next step:
- (9X+7) → 0.5(9X+7) which could result in either an Even or Odd number:

If we first consider 0.5(9X+7) to be an Even number, then:

- 0.5(9X+7) → 0.5(0.5(9X+7)) = (9X+7)/4 which could be either Even or Odd

And so the process continues to develop an unresolved outcome because an Odd number can be repeatedly generated.

An analysis was performed on all Odd numbers up to a value of 223,787. From this analysis it was observed that it took a maximum of 220 iterations before a value was obtained that was less than the starting value. This occurred for the starting value of 35,655. A plot of this series is presented in Figure 2.

Figure 3 presents the data from the top 31 series that required the most iterations in order to generate a value less than the starting value.

Figure 4 presents the number of iterations required to generate a value less than the starting value for all Odd starting values from 1 to 223,787.

As can be seen from Figure 4, the maximum number of iterations required to generate a value less than the starting value declines as the starting value increases. The reason being that as the starting value increases, the maximum number of iterations also increases, which increases the probability of getting 7, 8, 9 or 10 consecutive Even values.

This means that as the starting value of a series approaches infinity, the probability of generating a value that is less than the starting value approaches 100% (i.e. a certain outcome).

However, there is the risk that a series could enter into a numerical loop, which would never result in a value less than the lowest value within that loop.

Figure 5 is a plot of the number of series (Y-axis) that required a specific number of iterations (X-axis) to reduce the series value to a number less than the starting value. This plot is based on the top 27,974 tests from an analysis of all the starting values up to 223,787.

Figure 5 suggests that the maximum number of iterations required to reduce any series to a value less than the starting value is likely to be less than 300. But this cannot be considered conclusive because the probability of a loop forming at a higher starting value has not been dismissed.

7. Possibility of a numerical loop occurring

We know that a numerical loop is possible because of the loop (1—4—2—1), but can it be demonstrated that no other loop could form within any series?

We know that if another loop does exist, it must occur for a very large starting value, because the series has already been tested by various people for all starting values up to some very high values. We also know that if another loop does exist, then such a loop would also be triggered by any value that exists within that loop, as well as every doubling series value generated by each of the values contained within that loop. This means that if another loop could be found, then there would in fact be an infinite number of starting values that would eventually enter such a loop.

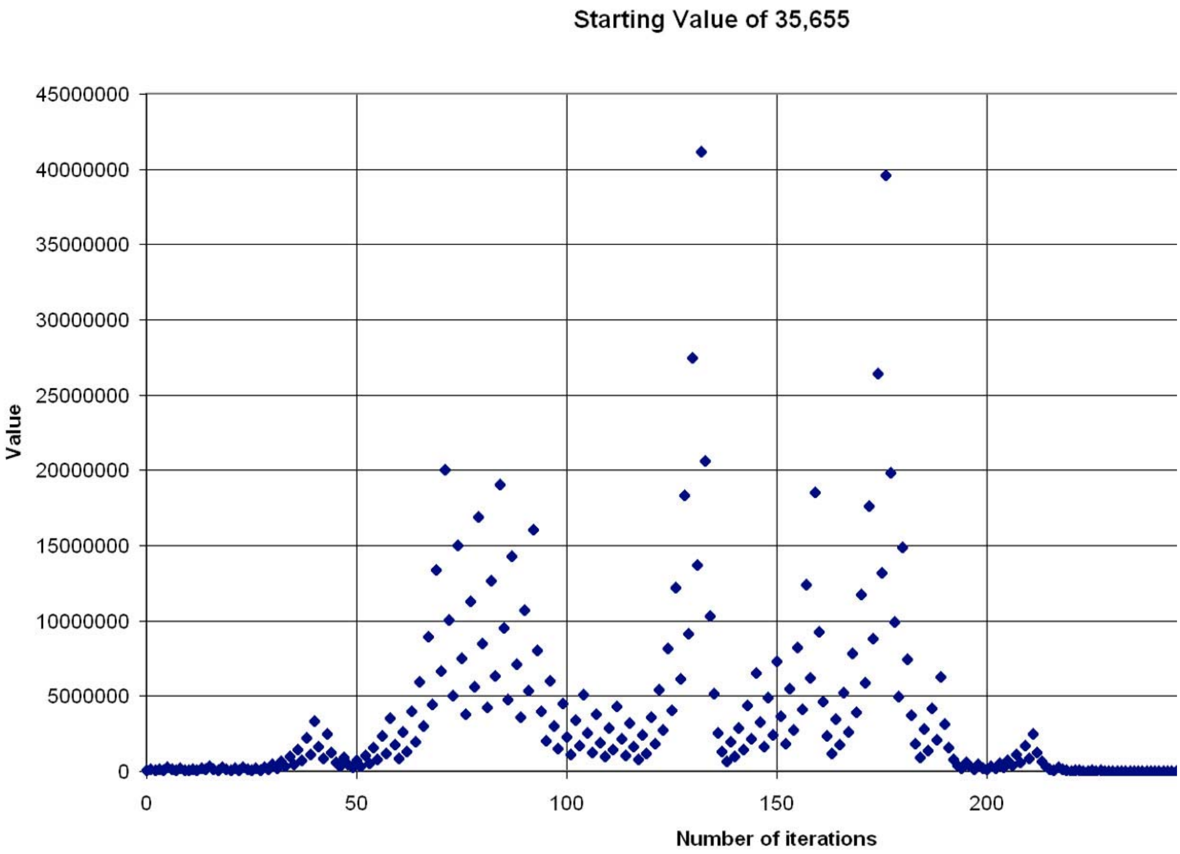

Figure 2: Progression of the series starting value of 35,655

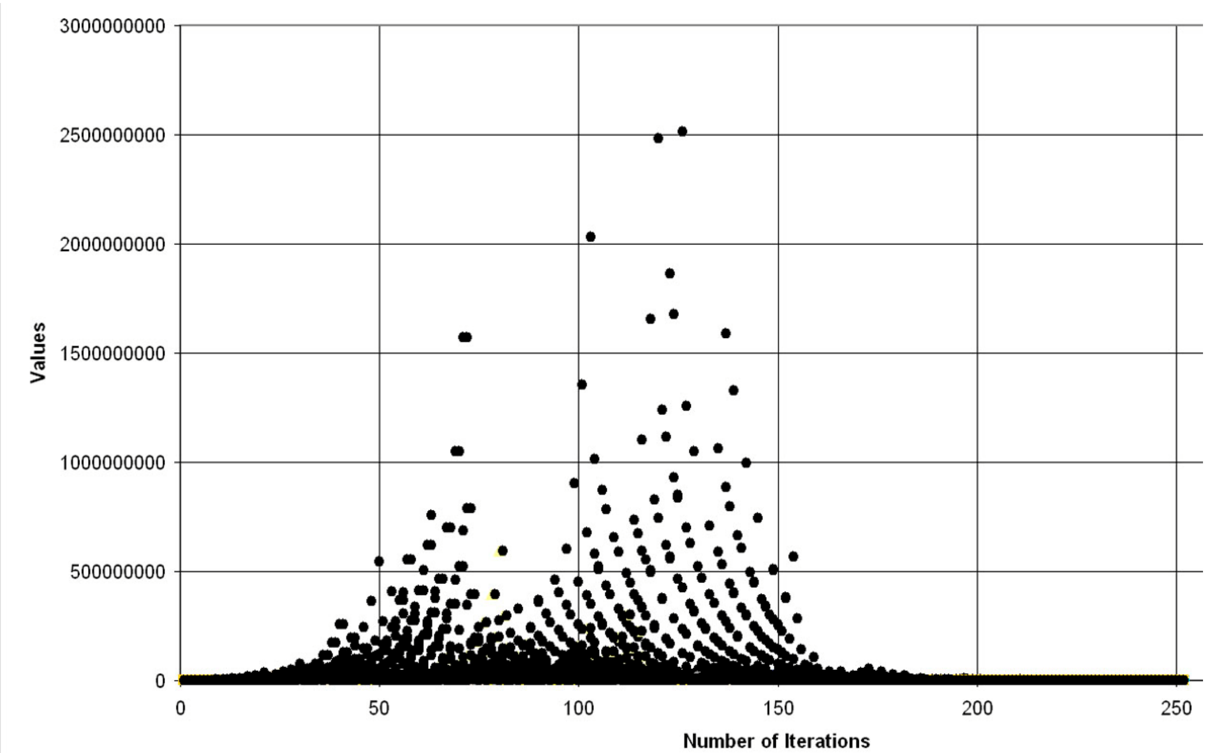

Figure 3: The collective data from the top 31 series (from 1 to 223,787) that required the most iterations in order to generate a value less than the starting value (all values that extend beyond 220 iterations are known to be less than the starting value for that series)

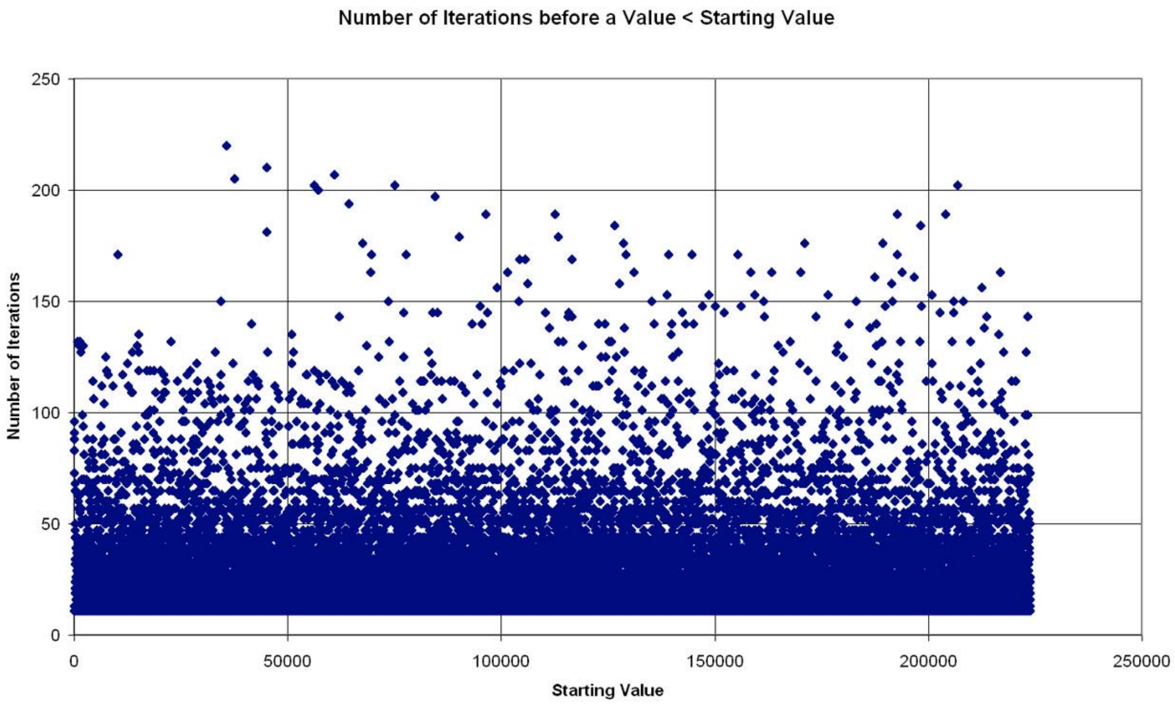

Figure 4: The number of iterations required to generate a value less than the starting value for all Odd starting values from 1 to 223,787

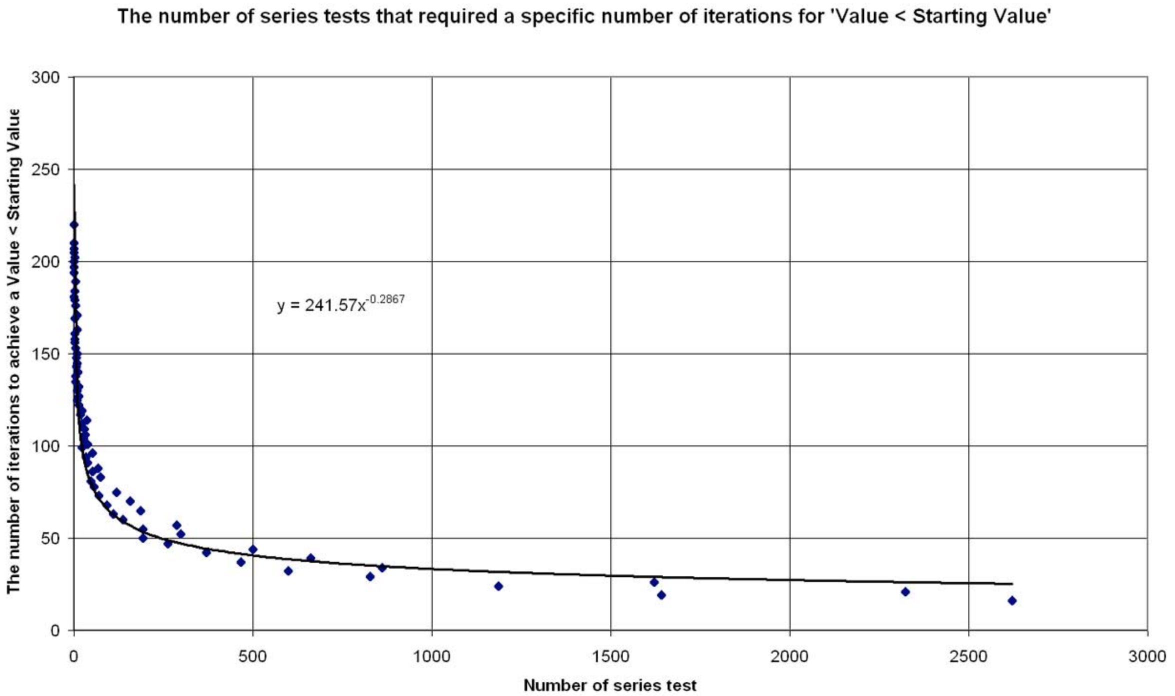

Figure 5: The number of series (Y-axis) that required a specific number of iterations (X-axis) to reduce the series value to a number less than the starting value (based on the top 27,974 tests from an analysis of 223,787 tests)

If the series starts with a very high starting value, the growth of the series, up until a value is generated that is lower than the starting value, could be approximated by the following conditions:

- An Even number is followed by a 50% reduction in that number (i.e. 0.5X).

- An Odd number is followed by that number multiplied by 3 (i.e. 3X); thus no ‘+1’.

Thus for any high-value series, the values within the series could be represented by Equation 1, but only for high values within the series.

$$(A)*(3^{m})*(0.5^{(n-m)}) = B \qquad (1)$$

where:

A = starting value

B = series value after ‘n’ iterations

m = number of Odd values within the series, but ‘m’ cannot exceed (n/2)

n = the number of iterations

In order for a loop to occur, ‘A’ must equal ‘B’, thus:

$(3^{m})*(0.5^{(n-m)}) = 1$

$(3^{m}) = 1/(0.5^{(n-m)})$

m(log(3) = -(n-m)log(0.5)

m(log(3) = m(log(0.5)) - n(log(0.5))

n(log(0.5)) = m(log(0.5)) - m(log(3))

n/m = (log(0.5) - log(3))/log(0.5) = 2.585

m = n/2.585, which does not exceed (n/2)

However, both ‘n’ and ‘m’ would need to be whole numbers, which means:

- 1/(((log(0.5) - log(3))/log(0.5))-2) = 1.71 would need to be a whole number, or very close to being a whole number, which it isn’t (a value ‘very close to being a whole number’ may work because of the assumptions associated with Equation 1, i.e. ignoring the ‘+1’).

Note: If a ‘whole number’ = X(Y.ZZZZZ), then 1/(Y.ZZZZZ - Y) must also be a whole number.

Consequently, a loop would not appear to be possible. To test this outcome, 112,245 high-value series conditions were tested, and no loops were identified. These tests involved series conditions based on Equation 1, and included all high-value series containing up to 760 iterations (i.e. well in excess of the expected 300 maximum iterations required to generate a value that is less than the starting value).

8. Alternative solution

Given the above demonstration that additional loops are not possible, and given that there is no limit to the size of the numerical values, or the number of iterations within a series, it can be concluded that every series from ‘1’ to infinity, but excluding infinity, must eventually generate a gold number, a silver number, or any number that exists within any series that satisfies the conjecture. This is the power of an infinite series of non-repeating values.

9. Conclusions

A solution to the Collatz Conjecture can be found not through an analysis of how many iterations is required to bring the series to a value of ‘1’, but by analysing:

- how many iterations are required to bring the series to a value less than the starting value; and
- demonstrating that a high-value loop cannot be generated; and
- demonstrating that if a solution can be found for a starting value of ‘1’, which it can, then a solution must be possible for a starting value of ‘2’, and subsequently for a starting value of ‘3’, and a starting value of ‘4’, etc. all the way up to, but not including, infinity.

Grant Witheridge