The Collatz Conjecture

Katherine Becking, Naveen Thomas Ninan, Rory Oden

Abstract

A basic introduction to the Collatz Conjecture and its variations is provided. A directed graph is explained. The expected value of the T and C functions are calculated and analyzed. The concept of escaping numbers is explained, and the principle that no number in the least escaping number's trajectory can be less than the least escaping number is used to make conclusions about possible forms of the least escaping number. The use of trees to map possible trajectory patterns and modular forms of the least escaping number is explored, and the limitations of this method are explained. Proofs are provided to demonstrate the trajectories of Mersenne numbers in the T function and to generalize these trajectories for variations on the Collatz function. The expected output is discussed, along with the methods used in the Collatz Calculator. The trajectories are analyzed and key words such as stopping time and minimum stopping time are defined. Finding maximum values using code are discussed. The attempt to understand the position of the longest ascent and its relation to the trajectory lengths is shown. Possible correlations and future steps for study are discussed.

I. Introduction

Intro to T and C with Examples:

Definition 1: The C map of the Collatz Conjecture states that starting with any positive integer, n, the iterations of this function will eventually produce the value 1, after you reach a power of two. If you continue you will end up in the loop: 4 → 2 → 1. Powers of 2 are always odd (16, 8, 4, 2, etc.) so once you hit a power of 2 you will always end up with 1.

$$C(n) = \begin{cases} 3n+1 & n \text{ is odd} \\ \frac{n}{2} & n \text{ is even} \end{cases}$$

Figure 1: C function

Definition 2: The T map of the Collatz Conjecture is similar to C, except it shortens the trajectories by dividing the 3n + 1 by 2 if n is odd.

$$T(n) = \begin{cases} \frac{3n+1}{2} & n \text{ is odd} \\ \frac{n}{2} & n \text{ is even} \end{cases}$$

Figure 2: T function

If we start with an initial input of 5 into *C*, then we would start by multiplying by three and adding one because five is odd. This results in an output of 16, which is a power of two. Successive iterates would generate a trajectory of

$$5 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1$$

If we start with an initial input of 5 into *T*, then we would start by multiplying by three and adding one because five is odd. We also have to divide by two in this step, giving us a value of 8. We skip the intermediate number 16, like we had in the C map of trajectories for 5. Successive iterates would generate a trajectory of

$$5 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1$$

The Collatz Conjecture states that for every starting value n, there is a number of iterations k such that C(n) applied k times is equal to one.

$$\forall n \in N\ \exists k \in N\ st\ C^\wedge k\ (n) = 1$$

This means that if C(n) is applied to a natural number n iteratively, there exists a number of iterations k that will produce a result of one. The Collatz Conjecture states that this is true for every n. The negation of this statement is:

$$\exists n \in N\ \forall k \in N\ st\ C^\wedge k\ (n) \neq 1$$

So, if the Collatz Conjecture is false, there must be a natural number n for which $C^k(\mathrm{n})$ is never equal to 1 for any number of iterations k. Such a number n is an escaping number. The hypothetical lowest such value is called the least escaping number (LEN) and is represented by n_0. All inputs up to

$5.78 * 10^{18}$ have been shown to produce a result of one, meaning the LEN, if it exists, must be very large[1].

II. Expected Value

Definition 3: Expected value is the quantity that results from summing each possible outcome multiplied by the likelihood that those outcomes will occur. The result is the expected average outcome from those choices. Expected value is very useful in determining what could happen and a very good use for expected value is stocks.

For example: Player 1 rolls a 6-sided die. What is the expected value of the dice roll?

$$EV = 1\left(\tfrac{1}{6}\right) + 2\left(\tfrac{1}{6}\right) + 3\left(\tfrac{1}{6}\right) + 4\left(\tfrac{1}{6}\right) + 5\left(\tfrac{1}{6}\right) + 6\left(\tfrac{1}{6}\right) = 3.5$$

The value 3.5 is the expected average based on the outcomes and their likelihoods. If this die was rolled a large number of times, then the average of all the numbers rolled would be a number close to 3.5.

Expected Value is a linear function[2]. In this instance, linearity is true if it follows:

$EV\ (X + Y) = EV(X) + EV(Y)$; X and Y are variables with finite expected values

$EV(cX) = cEV(X)$; c is any constant

The expected values could be calculated for each term separately and then combined together. This would equal the expected value of the both terms combined.

This is similar to the properties of integrals and derivatives such as:

$$\frac{d}{dx}(X+Y) = \frac{d}{dx}X + \frac{d}{dx}Y$$

$$\int_a^b (X+Y) = \int_a^b (X) + \int_a^b (Y)$$

Examples:

$$\begin{aligned} EV(T(m)) &= \frac{1}{2}\left(\frac{m}{2}\right) + \frac{1}{2}\left(\frac{3m+1}{2}\right) \\ &= \frac{m}{4} + \frac{3m+1}{4} \\ &= \frac{4m+1}{4} = m + \frac{1}{4} \end{aligned}$$

$$\begin{aligned} EV(C(m)) &= \frac{1}{2}\left(\frac{m}{2}\right) + \frac{1}{2}(3m+1) \\ &= \frac{m}{4} + \frac{3m+1}{2} \\ &= \frac{7m+2}{4} = \frac{7m}{4} + \frac{1}{2} \end{aligned}$$

The expected value of T(m) is a positive value. The T map is the modified version of the origin Collatz Conjecture that eliminates the repetitive step of dividing by 2 every time an odd number is the input. Therefore, $EV(T(m)) < EV(C(m))$ because the T function divides by 2 when odd whereas the C function does not. The C map goes up with a higher magnitude than the T map because it does not calculate the next iteration right away and that causes the expected value to be higher for the C function.

$$\begin{aligned} EV(T(m) - m) &= EV(T(m)) - EV(m) \\ &= m + \frac{1}{4} - m \\ &= \frac{1}{4} \end{aligned}$$

The expected value of $T(m) - m$ is a positive, constant value of ¼. This indicates that the function, on average, is on an increasing trajectory. As the number of iterates increase, the output of the function is also expected to increase. In the calculation above, the linearity of expected values is considered. Therefore, finding the expected values for the separate terms, T(m) and m, and subtracting would be equivalent to the expected value of the combined expression.

The reason why this is important is that this EV calculation tells us that the output of the function is greater than the input value. Additionally, this is only valid for $k \ll m$.

$$
\begin{aligned}
EV(\ln(T(m)) - \ln(m)) &= \frac{1}{2}\left(\ln\left(\frac{m}{2}\right) - \ln(m)\right) + \frac{1}{2}\left(\ln\left(\frac{3m+1}{2}\right) - \ln(m)\right) \\
&= \frac{1}{2}\ln\left(\frac{1}{2}\right) + \frac{1}{2}\ln\left(\frac{3m+1}{2m}\right) = \frac{1}{2}\ln\left(\frac{1}{2} * \frac{3m+1}{2m}\right) \\
&= \ln\left(\sqrt{\frac{3m+1}{4m}}\right) = \ln\left(\sqrt{\frac{3}{4} + \frac{1}{4m}}\right) \\
&\approx \ln\left(\sqrt{\frac{3}{4}}\right) = \ln\left(\frac{\sqrt{3}}{2}\right) \\
&\approx -0.144
\end{aligned}
$$

The expected value of $\ln(T(m)) - \ln(m)$ turns out to be a small negative number. This indicates that the log of the function, on average, is on a decreasing trajectory. Although it is expected to increase dramatically at first, it should eventually start decreasing more. This seems to be counterintuitive because ln() is an increasing function and the expected value is a negative number. It also seems contradictory that the previous calculation is a positive whereas this one is a negative. This indicates that the Collatz Conjecture is not too straightforward and can have a few gray areas. These opposite indications are a great example of this.

$$
\begin{aligned}
EV(C(m) - m) &= \frac{1}{2}\left(\frac{m}{2} - m\right) + \frac{1}{2}(3m + 1 - m) \\
&= \frac{m}{4} - \frac{m}{2} + \frac{3m+1}{2} - \frac{m}{2} \\
&= \frac{m+1}{2} + \frac{m}{4} \\
&= \frac{3}{4}m + \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
EV(F_5(n) - n) &= \frac{1}{2}\left(\frac{n}{2} - n\right) + \frac{1}{2}\left(\frac{5n+1}{2} - n\right) \\
&= \frac{n}{4} - \frac{n}{2} + \frac{5n+1}{4} - \frac{n}{2} \\
&= \frac{6n+1}{4} - \frac{2n}{2} = \frac{2n+1}{4} \\
&= \frac{1}{2}n + \frac{1}{4}
\end{aligned}
$$

The $F_5(n)$ function is like C(m) except that the input is multiplied by a 5 instead of a 3 and divides by 2 when odd. This closely related function has not been studied or researched to the extent of the Collatz function.

The two expected values above are for the functions $C(m) - m$ and $F_5(n) - n$ and they are both positive linear values. By comparison, it is indicated that C function has a larger expected increase than the F_5 function. This makes sense since F_5 cuts out the repetitive step of dividing by 2 after every iteration of an odd number; This is true even after replacing 3 with 5. Therefore, the increase in F_5 is cut down and $C(m) - m$ is predicted to have a larger value over one iteration.

$$\begin{aligned} EV\left(T^k(m) - m\right) &= E\left(T^k(m) - T^{k-1}(m) + T^{k-1}(m) - T^{k-2}(m) + T^{k-2}(m) + \ldots + - T(m) + T(m) - m\right) \\ &= E\left(T^k(m) - T^{k-1}(m)\right) + E\left(T^{k-1}(m) - T^{k-2}(m)\right) + \ldots + E(T(m) - m) \\ &= E\left(T\left(T^{k-1}(m)\right) - T^{k-1}(m)\right) + E\left(T\left(T^{k-2}(m)\right) - T^{k-2}(m)\right) + \ldots + \frac{1}{4} \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \ldots + \frac{1}{4} \\ &= \frac{k}{4} \end{aligned}$$

$T^k(m)$ represents the T map of the Collatz function that is run k number of times. The k is the number of iterations of that function. So, if $T(m) - m = \frac{1}{4}$, then it indicates that all the terms that follow that same pattern equals ¼. Therefore, for k iterations, the expected value of the T(m) function is $\frac{k}{4}$. This is done using a telescopic sum. This value indicates that the function keeps increasing whereas to prove the Collatz Conjecture, the number needs to get to 1.

III. Properties of the Least Escaping Number

Any output that results from applying C to an escaping number any k times must also be an escaping number, since a non-escaping result would cause the trajectory to end in one, making the original escaping number a non-escaping number, a contradiction. Hence, all of the outputs in an escaping number's trajectory must also be escaping numbers. For n_0, this means that all of the outputs in its trajectory must be larger than n_0, because all of the other escaping numbers are greater than n_0.

The piece of C that is applied to odd numbers (3n+1) will always result in an increased output, as will the odd piece of the T function ($\frac{3n+1}{2}$). However, the pieces of both the T and C functions that are applied to even numbers ($\frac{n}{2}$) will result in a decreased output. Therefore, n_0 cannot be even since the outputs in its trajectory cannot fall below its original value.

The principle that numbers in the LEN's trajectory must never fall below the original value can be used to map out possible trajectories for n_0. A certain ratio of odd iterations to even iterations must be maintained such that n_0 is never reduced to below its original value. For the C function, each odd iteration increases the value by approximately a factor of 3, while each even iteration decreases the value by a factor of two. (The addition of one in the odd function may be ignored because at large values such as n_0 must be, the difference of one number is not significant.) In the T function, the odd function increases the starting value by an approximate factor of 1.5, while the even function decreases it by a factor of 2. Hence, if k is the number of times the T function has been applied:

$$T^k(n_0) \approx \frac{3^{number\ of\ odd\ iterations\ that\ have\ been\ applied}}{2^k} n_0$$

Therefore, if $T^k(n_0)$ never falls below the original value n_0, it must follow a pattern of even and odd iterations of the T function such that:

$$3^{number\ of\ odd\ iterations} > 2^k$$

A. Possible Trajectories of the Least Escaping Number

The possible trajectories of n_0 through eight iterations of the T function are depicted in the following tree, with each level representing an iteration. A trajectory is eliminated as soon as it violates the condition $T^k(n_0) > n_0$. The trajectories are paths from the root of the tree to a leaf.

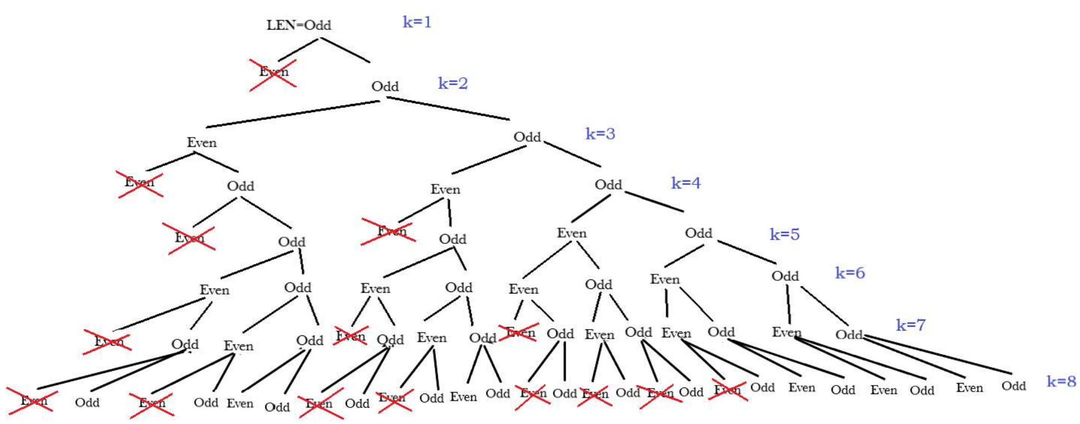

Figure 3: LEN trajectory tree

The next tree shows each modulus of 2^k, each representing a possible form of n_0. A modulus is an expression of the remainder a number will have when divided by another number. For example, $27mod64$ describes a number that has a remainder of 27 when divided by 64. Any number in the class $27mod64$ can be written as $64x + 27$, where x is some natural number. Note that $27mod64$ is a set of numbers with a common form, not a singular value.

In this tree, the branches represent subcategories, not paths. For example, $7mod16$ and $15mod16$ are both of the form $7mod8$. Each modulus that is crossed out violates the condition $3^{number\ of\ odd\ iterations} > 2^k$n . The following is an example of the calculations required to determine whether the aforementioned condition is met:

$$7mod64 = 64x + 7$$

$$2^k = 64$$

$$k = 6$$

$$T^6(64x+7) = T^5\left(\frac{3(64x+7)+1}{2}\right) = T^5(96x+11) = T^4\left(\frac{3(96x+11)+1}{2}\right) = T^4(144x+17) = T^3\left(\frac{3(144x+17)+1}{2}\right)$$
$$= T^3(216x+26) = T^2\left(\frac{216x+26}{2}\right) = T^2(108x+13) = T\left(\frac{3(108x+13)+1}{2}\right) = T(162x+20)$$
$$= \frac{162x+20}{2} = 81x+10$$

The above trajectory iterates the odd function four times and the even function twice.

$$3\text{^}(number\ of\ odd\ iterations) > 2\text{^}k$$

$$3^4 > 2^6$$

$$81 > 64$$

The condition is met, meaning that the LEN could be of the form $7mod64$. Notice that the last form in this trajectory is $10mod81$. Since the modulus is odd, there is no way of determining whether a given number of this form would be even or odd. Therefore, the trajectory for each class is only iterated k times. In this case, beyond 6 iterations, it is possible for a number of the form $7mod64$ to be either even or odd, meaning the form must be split into two cases ($7mod128$ and $71mod128$) in order to calculate the trajectory beyond k iterations.

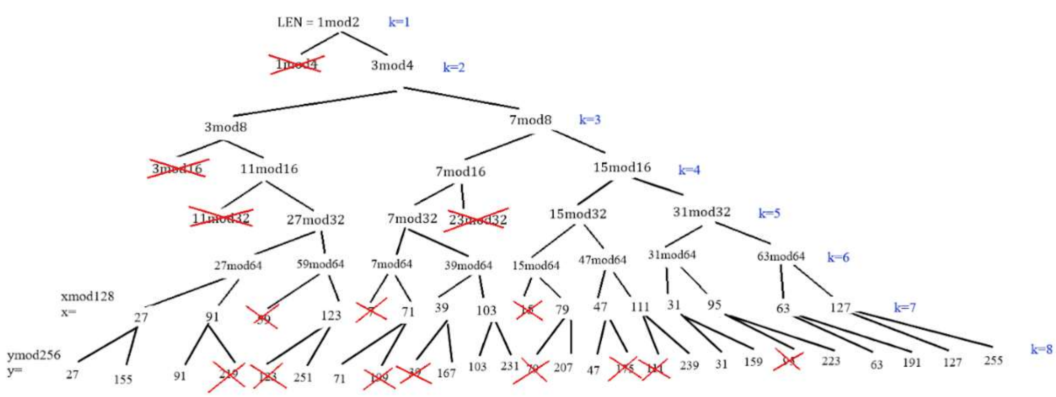

Figure 4: Modular form tree

As was expected, this modulus tree, Figure 4, has the same number of eliminations per level as the trajectory tree throughout the first eight levels, and the two trees have very similar structures. However, there are some differences. While the trajectory tree, Figure 3, only ever has eliminations on the left branch, the modulus tree sometimes has eliminations on the right, causing a difference in overall structure.

Although most levels on both trees have several eliminations, the overall number of branches is increasing. On each level, there are still infinitely many possibilities of trajectories or modulus forms for the LEN. The trees cannot be used to narrow down the possibilities to a single value. They do not reveal what the LEN is, but they reveal, to some extent, what it is *not*.

IV. Trajectories of Mersenne Numbers

The rightmost path on both trees represents number in the form $b2^k - 1$, which includes Mersenne numbers (numbers that are one less than a power of two, $b2^k - 1$). It can be inferred from the trees that Mersenne numbers are always odd for k iterations. The following proof demonstrates the resultant value when the T function is applied to a Mersenne number k times. The constant b is included in order to expand the application of this proof to other numbers.

Claim:

$$\forall k \in N \qquad T^k(b2^k - 1) = b3^k - 1$$

Proof:

Base Case:

$$T^1(b2^1 - 1) = T^1(2b - 1) = \frac{3(2b - 1) + 1}{2} = 3b - 1 = b3^1 - 1$$

Inductive Hypothesis: Assume that

$$k \in N, T^k(a2^k - 1) = a3^k - 1$$

Consider k+1 case:

$$T^{k+1}(b2^{k+1} - 1) = T\big(T^k(2b \cdot 2^k - 1)\big) = T(2b \cdot 3^k - 1) = \frac{3(2b \cdot 3^k - 1) + 1}{2} = b3^{k+1} - 1$$

A. Generalization of Mersenne Trajectories

This claim can be expanded to create a generalized formula for the paths of Mersenne numbers in the "variations" of the Collatz function, where 3 (in the $\frac{3n+1}{2}$ piece of the function) is replaced with $2^k + 1$. The constant b is also included in this proof to expand its application.

Claim: For an integer t such that $1 \leq t \leq n$,

$$F_{2^k+1}^{tk}(b2^{nk} - 1) = b(2^k + 1)^t(2^{nk-tk}) - 1$$

Proof:

Base Case: The first iteration uses the odd function, while the subsequent k-1 iterations use the even function.

$$F^{1}_{2^k+1}(b2^{nk}-1)=\frac{(2^k+1)(b2^{nk}-1)+1}{2}=b2^{nk+k-1}+b2^{nk-1}-2^{k-1}$$

$$F^{k-1}_{2^k+1}\left(b2^{nk+k-1}+b2^{nk-1}-2^{[k-1\}}\right)=\frac{b2^{nk+k-1}+b2^{nk-1}-2^{[k-1\}}}{2^{k-1}}=b2^{nk}+b2^{nk-k}-1$$

Therefore $F^{k}_{2^k+1}(b2^{nk}-1)=b(2^k+1)^s(2^{nk-sk})-1$

Inductive Hypothesis: Assume that

$$F^{sk}_{2^k+1}(b2^{nk}-1)=b(2^k+1)^s(2^{nk-sk})-1$$

Consider s+1 case:

$$F^{(s+1)k}_{2^k+1}(b2^{nk}-1)=F^{sk+k}_{2^k+1}(b2^{nk}-1)=F^{k}_{2^k+1}\left(F^{sk+k}_{2^k+1}(b2^{nk}-1)\right)=F^{k}_{2^k+1}(b(2^k+1)^s(2^{nk-sk})-1)$$

$$F^{1}_{2^k+1}(b(2^k+1)^s(2^{nk-sk})-1)=\frac{(2^k+1)(b(2^k+1)^s(2^{nk-sk})-1)+1}{2}=b(2^k+1)^{s+1}(2^{nk-sk-1})-2^{k-1}$$

$$F^{k-1}_{2^k+1}(b(2^k+1)^{s+1}(2^{nk-sk-1})-2^{k-1})=\frac{b(2^k+1)^{s+1}(2^{nk-sk-1})-2^{k-1}}{2^{k-1}}=b(2^k+1)^{s+1}\left(2^{nk-k(s+1)}\right)-1$$

Therefore $F^{(s+1)k}_{2^k+1}(b2^{nk}-1)=b(2^k+1)^{s+1}\left(2^{nk-(s+1)k}\right)-1$

Corollary: $t=n$

$$F^{nk}_{2^k+1}(b2^{nk}-1)=b(2^k+1)^n(2^{nk-nk})-1=2b(2^k+1)^n-1$$

V. Directed Graph

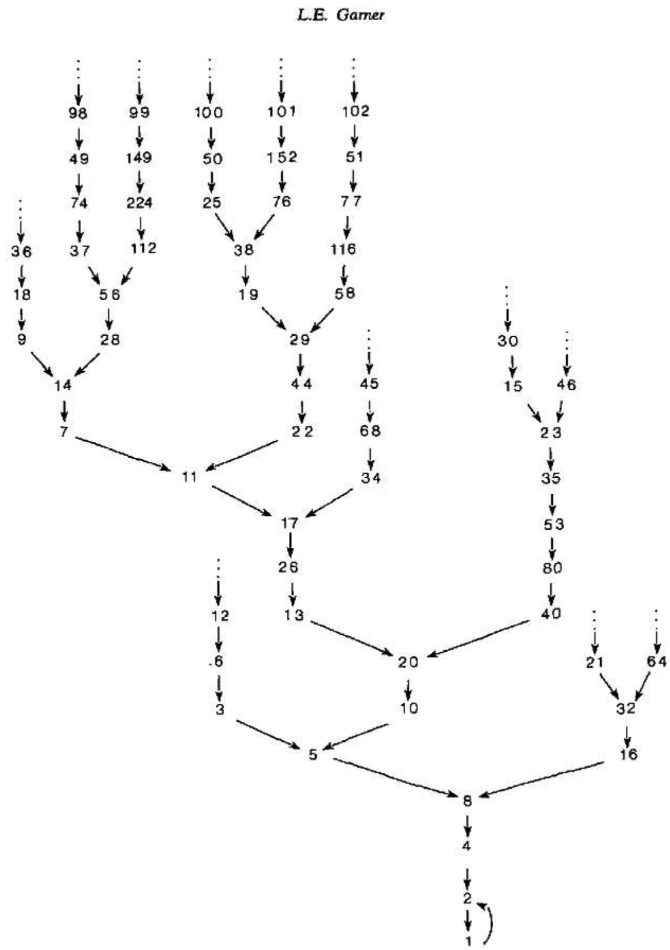

Figure 5: Directed graph of T(n)

Figure 5[3] is a directed graph of the T map of the Collatz Conjecture. It is a great way to visualize what is happening within the function and its trajectories. The arrows point to the output of the function when a number is inputted. There are also many branches on the tree/graph.

VI. Basic Code for the Collatz Conjecture in Java

Method 1:

```
/*
 * @purpose this is the normal collatz conjecture, 3n+1 if odd,
 * n/2 if even
 * @parameters integer n, which is the number to test
 * @return an int value
 */
public static int collatz(int n)
{
   if(n % 2 == 0)
   {
       return (n / 2);
   }else{
       return ((3*n) + 1);
   }

}
```

Method 2:

```
/*
 * @purpose this is the alternative collatz conjecture,
 * ((3n+1) / 2) if odd, n/2 if even
 * @parameters integer n, which is the number to test
 * @return an int value
 */
public static int altCollatz(int n)
{
   if(n % 2 == 0)
   {
       return (n / 2);
   }else{
       return (((3*n) + 1) / 2);
   }

}
```

Using this as a starting point we created a basic Collatz Calculator to test a few examples. The first method shows the C(n) function while the second method shows the T(n) function.

The calculator found the stopping time, minimum stopping time, and max value.

In Figure R3 the iterates of $n = 7$ are shown with the C map.

Definition 4: Stopping time[4] is the number of iterations before the output C(n) or T(n) is equal to 1.

Definition 5: Minimum stopping time[5] *is the number of iterations before the output C(n) or T(n) is less than n.*

```
-------------------------------------
  C Map - Collatz Calculator: 7
-------------------------------------
     #              Number
-------------------------------------
     1                7
     2               22
     3               11
     4               34
     5               17
     6               52
     7               26
     8               13
     9               40
    10               20
    11               10
    12                5
    13               16
    14                8
    15                4
    16                2
    17                1
-------------------------------------
  Stopping Time:     17
Min. Stop. Time:     12
      Max Value:     52 (#6)
```

Figure 6: Collatz Calculator- $C^k(7)$

```
-------------------------------------
  T Map - Collatz Calculator: 7
-------------------------------------
     #              Number
-------------------------------------
     1                7
     2               11
     3               17
     4               26
     5               13
     6               20
     7               10
     8                5
     9                8
    10                4
    11                2
    12                1
-------------------------------------
  Stopping Time:     12
Min. Stop. Time:      8
      Max Value:     26 (#4)
-------------------------------------
```

Figure 7: Collatz Calculator- $T^k(7)$

We found the maximum value in the list of trajectories, and our next step was to determine the relative maximums, or where the graphs of the trajectories peaked. We did this because we wanted our calculator to recognize more than just the maximum value, so it would allow for more analysis. The best example for this was the T map of T(63).

We used the calculator to compile the list of trajectories in an ArrayList. We then sorted through the ArrayList to determine if a value was a relative maximum. To be considered a local maximum, a trajectory of index *i* had to be greater than the trajectories in indices *i - 1* and *i + 1*.

For the input of 63 into *T*, this gave us a list of relative maximums as pictured in Figure R5. These maxima are viewed even more clearly on the graph in Figure R6.

```
-------------------------------------
  Stopping Time:          69

Min. Stop. Time:          56

      Max Value:        4616 (#44)

       Rel. Max:         728 (#7)
                         206 (#12)
                         350 (#16)
                         890 (#21)
                         668 (#23)
                         566 (#28)
                         638 (#31)
                        3644 (#38)
                        4616 (#44)
                         866 (#48)
                         650 (#50)
                         488 (#52)
                          92 (#56)
                          80 (#61)
                           8 (#66)
-------------------------------------
```

Figure 8: Collatz Calculator Output for $T(63)$

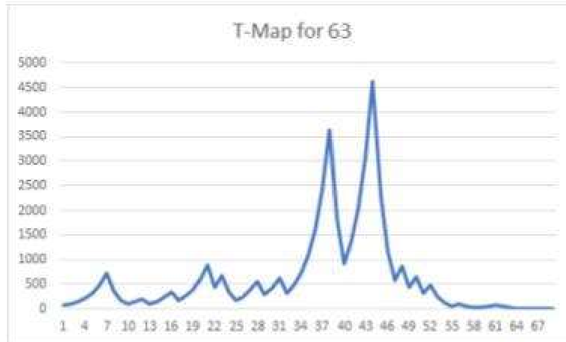

Figure 9: Height of output vs. no. of iterates

Our next step was to determine the length of the largest ascent, with the thought process being that the position of the largest ascent might have some correlation to the length of the trajectory.

To determine the largest ascent, a method was written in our code for the Collatz Calculator. It would accept a list of trajectories and return two values: the length of the ascent and the position. An ascent would be counted as a continued increase in the value of the trajectories. The length would terminate if the next value was less.

For T(63) the output was

```
Longest Ascent:          6 (#1)
```

This could be interpreted as seven values of the trajectory of T(63) were all increasing, or six steps. These seven values start at position #1, or the very first value of 63.

```
-------------------------------------
    T Map - Collatz Calculator: 63
-------------------------------------
          #                Number
-------------------------------------
          1                  63
          2                  95
          3                 143
          4                 215
          5                 323
          6                 485
          7                 728
          8                 364
```

We thought that there might be a correlation between the position of the longest climb and the trajectory length. So, we graphed the trajectory length versus the ratio of the position of the longest climb and the trajectory length for ten thousand numbers.

$$\text{Ratio} = \frac{\text{Position of Longest Ascent}}{\text{Trajectory Length}}$$

$$\text{Ratio for } T(63) = \frac{1}{6} = 0.166667$$

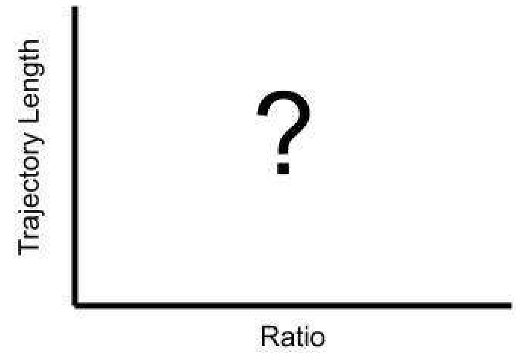

Figure 10: Trajectory Length vs. Ratio

Each dot represents a number like 63, though the coordinates are (0.166667, 69) or (Ratio, Trajectory Length).

There are some interesting possibilities for analysis in the graph in Figure R7. We could look for Mersenne numbers and see if they follow a pattern, or if another series of numbers had the same path. We could also see if the outliers have any common properties. It'd be great if we could zoom out per se and see the graph as a whole as the amount of tested numbers increase. Would we see a spiral or some other type of pattern?

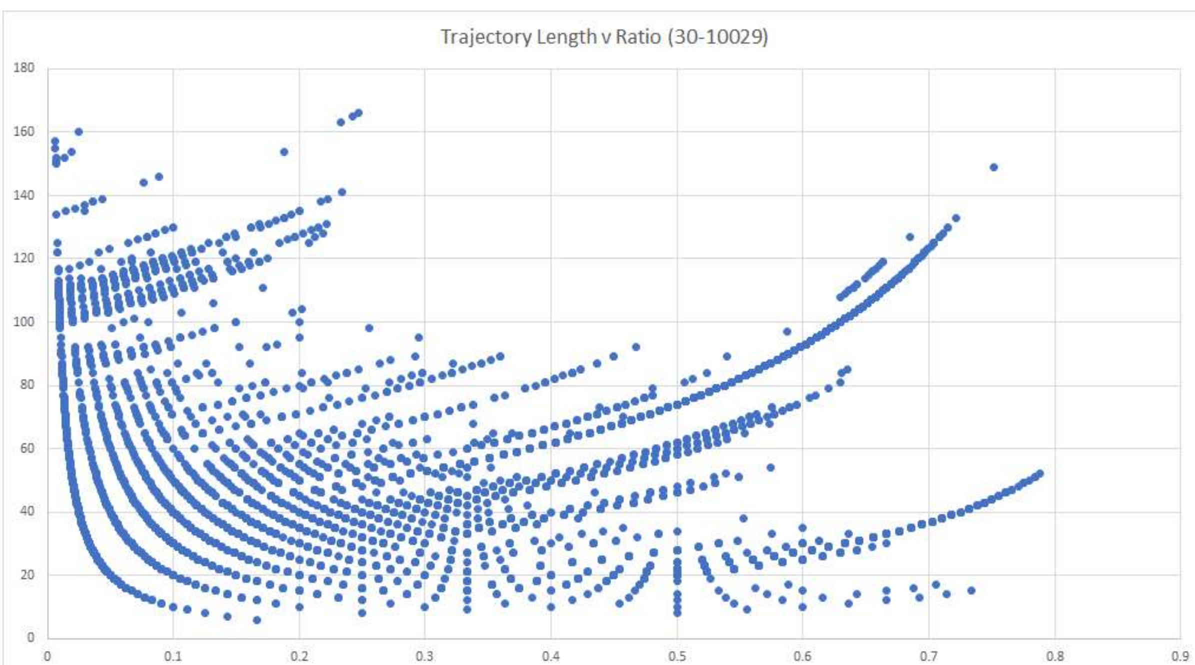

Figure 11: Trajectory Length vs. Ratio Graph

VII. Conclusion

This project has broadened our understanding of the Collatz Conjecture and its possible solutions. Some patterns in ascents and trajectories have become apparent through the use of coded calculators, and it has been determined what the least escaping number cannot be through the use of trees.

However, our investigations in this project have revealed more about the limitations of the problem than about the actual solution. A solution to the problem cannot be found computationally, since it is impossible to check every number. Neither can we use the process of elimination through modular trees to find the value of a single escaping number, since each level of such a tree leaves infinitely many possibilities open.

We have found several seemingly contradictory properties of the Collatz problem. For example, the expected value calculation for $T(m) - m$ is a positive number whereas the expected value for $ln(T(m)) - ln(m)$ is a negative value. This seems odd because the log function is an increasing function and the two values contradict each other. These opposite indications reveal that the Collatz problem has very contradictory properties and that is one of the reasons that it has been unable to be solved.

Another confounding aspect of this puzzle is its simplicity. It is easy enough for the average person to understand, and yet it has remained unsolved for decades. One possible explanation is that the necessary mathematics for understanding this problem simply does not exist yet. It is evident that this problem, if it is ever resolved, will be solved by thinking creatively and from a fresh perspective which would not be apparent to most people.

Ultimately, the Collatz Conjecture teaches us about the importance of theoretical mathematics. Through mathematics, new ways of thinking, analyzing, and observing are discovered. Some issues, in math and in the world, cannot be solved through computation, and require more complex theoretical solutions. Theoretical math may seem esoteric and removed from practical application, but it forms the basis of problem solving in every field and aspect of life.

References

[1] Tomás Oliveira e Silva, “Empirical verification of the 3x+1 and related conjectures, The ultimate challenge: the 3x + 1 problem”, 189207, Amer. Math. Soc., Providence, RI, **2010**.

[2] Charles M. Grinstead and J. Laurie Snell, *Introduction to Probability*, American Mathematical Society, pp. 241, Theorem 6.2, Chapter 6, **1 July 1997.**

[3] Lynn E. Garner, “On heights in the Collatz 3n + 1 problem”, Discret. Math., Vol. 55, pp 57-64, **1985.**

[4] Jeffrey C. Lagarias. “The 3x 1 Problem and Its Generalizations”, The American Mathematical Monthly, Vol. 92, No. 1, pp. 3-23. **Jan.1985**. doi:10.2307/2322189

[5] ibid