A Brief Look into the Collatz Conjecture

Vedant Lohia

Abstract

The late Paul Erdős famously said that regarding the Collatz Conjecture, "Mathematics is not yet ready for such problems" and his words ring true to this day. This paper outlines the attempts made to prove or disprove the conjecture and refrains from restricting itself to a very mathematical audience. While surface level knowledge of such conjectures and their nature might satiate some, it is for those who wish to know more before a sea of notation and concepts overwhelms them that I dedicate this paper.

February 11, 2022

1 Introduction

The Collatz Conjecture, also known as the Syracuse Problem, Ulam's Problem and Hasse's Algorithm (although there are many other names) appears to be, in modern day mathematics, an unsolvable conjecture. Much like Fermat's Last Theorem, it is a problem that, at face value, exhibits a very innocent nature that can be easily explained using elementary mathematics, unlike perhaps the mathematics required in order to prove or indeed disprove it. It can be defined like so, performing operations on the positive integer n dependent on its parity:

$$C(n) = \begin{cases} 3n+1, \text{if} & n \equiv 1 \pmod 2 \\ n/2, \text{if} & n \equiv 0 \pmod 2 \end{cases}$$

The outputs of this function, which is commonly referred to as C(n) and is what for the rest of the paper I shall refer to the Collatz Function as, is the main appeal of the conjecture itself and it is what has plagued a great number of mathematicians over the years. To verbally express this function, if our starting integer is odd, we multiply it by 3 and add 1, if it is even, we divide it by 2. We then perform the function to the output of the first iteration and from this to obtain a list of integers if we keep going. There exists other papers that have served to summarise the progress made on this problem however they are primarily aimed at those who are already at a high level of mathematics see (Lagarias, 2010).

1.1 The Conjecture Itself

To demonstrate, we may start out with our positive integer n as 7. Sparing any unnecessary calculation we obtain the following set of integers upon repetition of the function:

$$\{7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1\}$$

Notice what occurs as we reach the final collection of integers {4, 2, 1} One may ask why there is even a set at all as it can be tempting at first glance to assume that this function produces a set of integers that contains infinitely many terms. To an extent, that is true, however for convenience we recognise that the final three integers repeat endlessly. To put it in terms of the function C(n), C(C(C(1))) = 1, that is that C(1) = 4, C(4) = 2 and C(2) = 1. It is this phenomenon that concerns yet simultaneously fascinates mathematicians alike. So naturally, the conjecture that can be derived, having tested numerous positive integers, is that if every single positive integer eventually leads to the {4,2,1 } loop. And so far, there has been no number tested, despite this function having been carried out for all numbers up to 2^{68}, that is roughly 295 quintillion(Tao, 2020), that has broken this pattern.

1.2 The Aims of this Paper

Most literature concerning maths and moreover problems, conjectures and theorems like the Collatz Conjecture can be difficult to grasp without extensive background knowledge. It is very easy to get deterred from finding out more on such complex mathematical ideas when there are a multitude of prerequisites to learn, understand and put into practice beforehand. Some may not have the resources, time or energy to do so. Taking this into account, the first goal of this paper will be to compound all current and past approaches in solving the Collatz Conjecture in an easily understandable format. I will very shortly discuss a bit about its background and will then proceed in Section 3 to go over these approaches. From each method then, denoted by their respective subsections, you can feel free to research more and broaden your knowledge on specific topics that have been considered and used to attempt to prove the conjecture. It is safe to say that some may not want to go on themselves to work on the Collatz Conjecture and that there are some who wish to learn about it at its surface. This paper will mainly be aimed at those who may not know extensively about the advanced mathematics required to understand the details of approaches but would anyways like to discover more about the problem for whatever reason. The second aim of this paper is to justify the need to prove or disprove the Collatz conjecture. It is great from the perspective of a pure mathematician to dwell on such a problem but without a clear contribution and meaningfulness to mathematics as a whole, one can step back from it and question what the purpose is. I will discuss some potential uses in Section 4, aptly named Relevance and Potential Uses, and will shed some light as to whether this conjecture is one that need be shelved for the time being or if it is a feasible problem to solve given the current mathematical tools we have available. Without any further delay, let's move on to the history of the Collatz Conjecture.

2 Early History

Despite the conjecture falling under multiple names, it is the German mathematician Lothar Collatz who's name it takes after. As hinted at previously, the conjecture has been around for a while and it is generally accepted that it was spread in the 1950's at the International Congress of Mathematicians in Cambridge, Massachusetts. It is here that Collatz shared around the problem among the mathematicians in attendance with the likes of Shizuo Kakutani, Stanislaw Ulam and Harold Scott MacDonald Coxeter; most of whom would work to some degree on the problem thereafter. It is unclear however as to the exact origin of the conjecture or who created it first as there have been many contestants to claim it themselves (most notably the English mathematician Sir Bryan Thwaites). While it was brought to the attention of other mathematicians in 1950, Collatz allegedly formulated it in 1937 having studied graph theory, specifically concerning graphical depictions of iterating functions such as C(n), just a couple of years prior and after having attended the Congress in 1950, in 1952 he then went on to spread it to another one of his colleagues attributed to the problem: Helmut Hasse, hence the alternative name Hasse's algorithm.

By the late 1950's the problem was still fairly unknown although it had spread considerably to other notable institutions e.g Cambridge University. Due to the problem remaining relatively hidden and the dominant prevailing 'Bourbaki-style' mathematics, which did not favour solving such problems, in the 1960's, as stated by Jeffrey C. Lagarias (Lagarias, 2010) in his works pertaining to the conjecture, significant publication and literature concerning it was more or less suspended until the 1970's. 1972 was the year of its popularisation with the conjecture being published in a Scientific American column by Martin Gardner and from there on it has worked its way into history as one of the most unreachable and unattainable problems in number theory.

3 Approaches and Partial Results

Now we approach the list of partial results and methods that have tried to solve this great problem over the years. The list is compiled and assorted mostly by the names of those who have made attempts bar the first in no particular order.

3.1 A Heuristic Argument

There is a lot to be said about the nature of the Collatz function and its outputs. It does not appear follow a strict pattern and as such, the outputs and the sequences of these outputs have often been labelled as "Hailstone Numbers". This description of their nature alludes to the rapid rise or fall in value that tends to occur in what is known as a "Collatz orbit", referring to the circulation of numbers that we pass through on our way to 1. It is best to observe this in action by referring back to Section 1.1 with our example of our starting integer 7 and one can surmise that this rise and fall becomes more erratic as we choose larger and larger starting integers. The Heuristic Argument concerning the Collatz function which although has, to my knowledge, no clear originator, serves to demonstrate that, given successive odd iterations of the Collatz function, they are expected to decrease each time by a factor of 3/4. This, while not wholly proving the conjecture, strongly advocates for its validity. Before we delve into the argument and its derivation, we must define the function T(n), a simplified model of C(n) that essentially skips an unnecessary step in the process as we know that if our initial integer n is odd, it must undergo 3n + 1 which by a simple proof below is bound to be even hence we divide it by 2 to give (3n + 1)/2 as our operation for the odds. This "shortcut" Collatz function is particularly helpful not only in expressing the problem in a more simple form but also when it comes to studying what are known as the parity sequences of a Collatz orbit which is essentially a list of whether the numbers are odd or even when regarding outputs of the function.

Assume an odd integer 2k + 1 for a constant k such that $k \in W$.We substitute 2k + 1 into the 3n+1 to give 3(2k + 1) + 1 = 6k + 4. Hence we see that 6k+4 can be expressed as a multiple of 2 or in the form of 2 multiplied by a constant as 2(3k + 2); the definition of an even number. Below is our simplified version of C(n).

$$T(n) = \begin{cases} (3n+1)/2, \text{ if } n \equiv 1 \pmod 2 \\ n/2, \text{ if } n \equiv 0 \pmod 2 \end{cases}$$

What we look to do when trying to prove the Collatz conjecture, with the assumption that it is true, is to determine whether for any number n, its sequence pertaining to its orbit converges, at some point, to 1. This is what is considered the total stopping time and can be expressed as such with respect to positive integer k such that $T^k(n) = 1$. The positive integer k can be deemed as the number of "steps" of calculation required to reach 1 and can be expressed more concisely as $\sigma(n) = k$. For our example of the number 7 we can write, $T^{16}(n) = 7$ or $\sigma(7) = 16$. With this out of the way, we make an assumption based on the probability of acquiring an odd or even number from the operation (3n + 1)/2 where n is an odd integer. We say that given n as an odd integer, there is a 1/2 likelihood of obtaining an even number and a 1/2 likelihood of obtaining an odd number. We also assume that for any given Collatz orbit, the number of evens in the parity sequence will be k/2 and ditto for the odds. We observe that for odd integer n, the probability of getting an odd output from (3n + 1)/2 is 1/2, the probability of obtaining an odd output from this (ie (3n + 1)/4) is 1/2*1/2 = 1/4 and the probability of gaining an odd output from this is 1/2*1/2*1/2 = 1/8 and so on. This is where we can calculate what is known as an expected growth factor which has been proven to be 3/4. What this essentially boils down to is evaluating the sum below, you may be able to spot a pattern as to how it relates.

$$\left(\frac{3}{2}\right)^{0.5}\left(\frac{3}{4}\right)^{0.25}\left(\frac{3}{8}\right)^{0.125}\left(\frac{3}{16}\right)^{0.0625}...$$

Again, without getting into too much calculation and detail, this sequence can be represented more concisely as shown below and we split it into two parts and prove that they both converge. After evaluating each, we multiply them to give our final answer of 3/4.

$$\prod_{i=1}^{\infty}\left(\frac{3}{2^i}\right)^{\frac{1}{2^i}}=\prod_{i=1}^{\infty}(3)^{\frac{1}{2^i}}\cdot\prod_{i=1}^{\infty}\left(\frac{1}{2^i}\right)^{\frac{1}{2^i}}$$

$$\prod_{i=1}^{\infty}(3)^{\frac{1}{2^i}}=3$$

$$\prod_{i=1}^{\infty}\left(\frac{1}{2^i}\right)^{\frac{1}{2^i}}=\frac{1}{4}$$

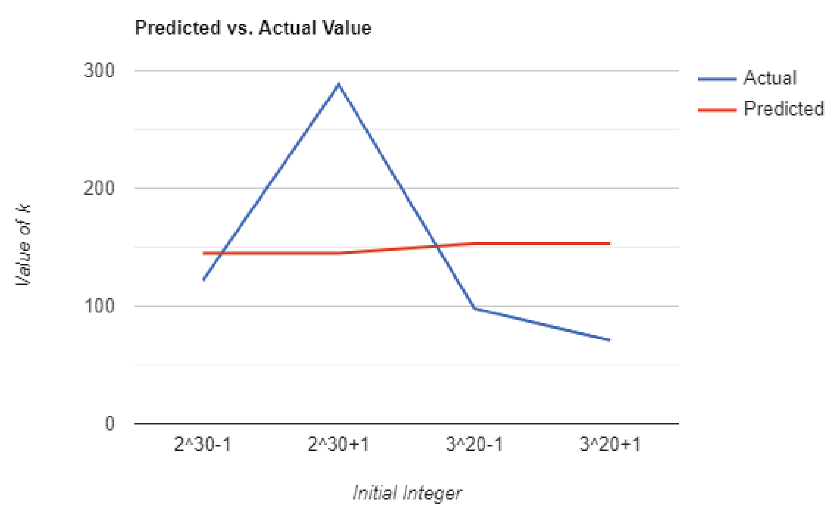
$$\prod_{i=1}^{\infty}\left(\frac{3}{2^i}\right)^{\frac{1}{2^i}}=\frac{3}{4}$$

This result has been used, with varying success, to predict the number k in a given Collatz trajectory. Looking back at our initial assumption that T(n) will take an equal distribution of even and odd numbers, we can create an equation involving k that we can use to approximate.

$$\left(\frac{3}{4}\right)^{\frac{k}{2}}n=1$$

k, after a bit of calculation involving logarithms, turns out to be proportional to log(n). Plotting this on a graph yields interesting but near fruitless results in an accurate prediction of k. Nevertheless, the table below, with values taken from the paper: 3n + 1 problem: scope, history and results by T. Ian Martiny (Martiny, 2015), shows the behaviour of this model and its accuracy.

Initial Integer	**Actual Value**	**Predicted Value**
n	$\sigma(n)$	k
$2^{30}-1$	122	145
$2^{30}+1$	288	145
$3^{20}-1$	98	153
$3^{20}+1$	71	153



3.2 Terras, Allouche and Korec

While it should be mentioned that each individual listed in this subsection has not collaboratively worked on the Collatz Conjecture, their respective partial results all fall in the same vein. To summarise beforehand, Terras had shown in his 1976 paper that almost all initial values n on which we perform our Collatz function T conclusively iterate to a value that is less than n. Allouche and Korec have improved this result whereby they proved that for an initial value n, it iterates to a value less than $n^{0.869}$ and more improved to a value that is less than $n^{0.7925}$ respectively. To begin, we will look at the initial derivation that Terras made in his paper ”A stopping time problem on the positive integers” (Terras, 1976).

3.2.1 Terras

Terras initiates his paper by laying out the fundamental theorem he wishes to prove, as any paper should, which is that the limit of the function below as k tends to infinity is 0.

$$F(k) = \lim_{m\to\infty} \left(\frac{1}{m}\right) \#\{n \leq m | \chi(n) \geq k\}$$

$$\lim_{k\to\infty} F(k) = 0$$

What this function essentially denotes is the probability of, in the range of specified numbers from 1 to m, choosing a number that has a stopping time $\chi(n)$ greater than the total stopping time k. The # in this case denotes the cardinality (the number of elements) of the set of numbers from n to m which meet our given condition. If we prove hence that the limit of this function with respect to k as k tends to infinity is 0, simply put it says that the probability of getting a number that has an infinite total stopping time (i.e a number that disproves the conjecture) is virtually 0. From this we can derive that *almost all* numbers iterate to a value that is less than their original. Almost all here is equivalent to saying that for example almost all numbers are not prime despite us knowing that they exist. The percentage of primes out of natural numbers tends towards although that isn't proof that no primes exist.

To begin to make sense of the method Terras uses to prove this, we first define function $\tau(n)$. This new function can be seen as another way to define the stopping time of an integer n and is key in proving the existence and result of the limit. Terras first creates a set denoting the sum of the numbers in the Collatz orbit according to their parity. If we map the even numbers to the value 0 and the odd numbers to 1, we can sum it all to give what Terras denotes as S_i. We then define a function $\lambda_i(n)$ as:

$$\lambda_i(n) = 2^{-i}3^{S_i(n)}$$

Without going into too much detail, our new stopping time function $\tau(n)$ is defined hence as equal to k whereby $k > 0$ if k is the smallest integer such that $\lambda_k(n) < 1$ (the first instance of when in a given Collatz orbit a number below the starting integer is reached). Terras proceeds then to show that for sufficiently large values of n, the $\tau(n)$ stopping time function is equivalent to the $\chi(n)$ stopping time function. Having proved this, he then substitutes $\tau(n)$ into our original function F(k). The development of the new stopping time function is what allows Terras to express the original function in a slightly different manner and prove that a limit for it even exists. With our condition now being whether $\tau > k$, Terras proves that you need a certain amount of odd numbers in the sequence defined below (which is equivalent to $\tau > k$) for the condition to be met. If we recall from the beginning, we are trying to prove that the probability of this condition being met is in fact 0. γ in the below case is equal to $\frac{ln(2)}{ln(3)}$ and the X(n) function (not to be confused with stopping time $\chi(n)$) denotes what the parity of a given nth integer is in a Collatz orbit.

$$X_0(n) + X_1(n)... + X_{i-1}(n) > i\gamma$$

Then Terras defines a sequence ϵ which has an arbitrary sequence of odd and evens (0s and 1s) and then denotes n(a, k) as the number of such defined sequences which contain specifically a 0s , i.e an a amount of evens. With this being said, we can now express our original probability F(k) as the number of sequences that meet the above condition against the total numbers of sequences which when expressed with respect to n(a,k) looks like this:

$$F(k) = \sum_{a=0}^{k} n(a,k)/2^k$$

Terras then shows that n(a, k) can be at most $\binom{a}{k}$ and thus we obtain the inequality below

$$F(k) = \sum_{a=0}^{k} n(a,k)/2^k \leq \sum_{a=0}^{k} \binom{a}{k}/2^k$$

It is at this point that the limit of F(k) can be evaluated with what is known as the central limit theorem to prove that F(k) converges to 0.

3.2.2 Allouche and Korec

Allouche's contribution (Allouche, 1978-1979) has been mentioned in the starter to this section however it is unable to be explained at least by myself due to the paper in question written in another language. Nevertheless, to summarise the paper briefly, Allouche proves that all "almost all" values iterate to a value less than $n^{0.869}$ and also states that not just asymptotic behaviour is required in order to determine the periodicity of the function with periodicity referring to whether there are repeating points in the function and the intervals between them. The ideas used in Allouche's paper build on those used used by Terras in his original proof and so does the subsequent paper by Ivan Korec (Korec, 1994).

The initial claim to prove in Korec's paper is that for an arbitrary starting value y, it without fail iterates at some point in its Collatz orbit to a value less than $y^{\approx 0.79..}$, if and only if $S_m(y) < md$ for a sufficiently large value of m where $S_k(y)$ denotes for any Collatz orbit of y until term k the number of the odd numbers, in essence:

$$S_k(y) = X_0(y) + X_1(y) + X_2(y) + ...X_{k-1}(y)$$

The overall theorem to be proved here is that for c > $\log_4 3$, the expression below has an "asymptotic density" of 1 essentially meaning that the probability of picking a value that holds the condition is so high that it is essentially one hundred percent or 1. N denotes the set of non-negative integers. The backwards "E" here is equivalent to the phrase "there exists".

$$M_c = \{y \in N | (\exists n)(T^n(y) < y^c)\}$$

Another function we need to define in order to complete the proof is U which operates on variables m and d. In the first claim Korec is proving that if these conditions in the function below hold, the starting y value iterates to less than y^c.

$$U(m, d) = \#\{y \in N | 0 \leq y < 2^m \text{and} S_m(y) \leq md\}$$

Again, # denotes the cardinality of the set. To prove theorem one, we must effectively prove what is below for a given ϵ greater than 0 and less than 1. The best case scenario for the below statement is if we consider what happens if ϵ is itself 0. It would imply that that both sides are equal, in turn stating that the number of the numbers less than a that iterate to a value less than y^c is equal to a itself.

$$\#\{y \in M_c | y < a\} \geq (1 - \epsilon) \cdot a$$

To put it briefly, Korec forms several inequalities based on the one below. We know that $T^p(y)$ must be greater than or equal to $y/2^p$ since $y/2^p$ would imply that our inital value y was being divided by 2 multiple times which would be the best possible scenario but very unlikely. If even one of the numbers in the Collatz orbit of $T^p(y)$ was odd, it would be greater than $y/2^p$. Note, $p < m$. From this he obtains the inequality below that where k $= S_m(y)$.

$$T^p(y) \geq \left(\frac{y}{2^p}\right) > \frac{m \cdot 2^m}{2^m} = m$$

$$T^m(y) = y \cdot \frac{T^1(y)}{T^0 y} \cdot \frac{T^2(y)}{T^1 y} \cdot \frac{T^3(y)}{T^2 y} \cdots \frac{T^m(y)}{T^m - 1y} < y \cdot \left(\frac{3m+1}{2m}\right)^k \cdot \left(\frac{1}{2}\right)^k$$

After a bit of algebraic manipulation, we obtain the inequality below which we can then shorten to the inequality right underneath it for sufficiently large values of m as m grows quicker than $\log_2 m$ hence the limit of the expression would be 0.

$$\frac{k}{m} < \frac{c}{\log_2 3} - \frac{1 + 2(1-c)\log_2 m}{m}$$

$$\frac{k}{m} < \frac{c}{\log_2 3}$$

Hence we now know that for a large value of m and a and an initial value y given the conditions $m \cdot 2^m \leq y < a$ and $S_m(y) < md$, $y \in M_c$. The next step in this proof is to prove that "almost all" numbers do this. Korec splits up the numbers in the bound $m \cdot 2^m \leq y < a$ into "pairwise disjoint" sets (denoted by L(a)) which have 2^m integers in each with a small but negligible remainder at the end which he uses the floor function to get rid of (e.g 3.1415926... would turn into 3). He then uses a result from Terras' paper shown below and forms a final inequality to finish the proof.

$$d > \frac{1}{2}, \lim_{m \to \infty} \frac{U(m,d)}{2^m} = 1$$

$$\#\{y \in M_c | y < a\} \geq L(a) \cdot U(m,d) \geq \left(1 - \frac{\epsilon}{2}\right) \cdot \frac{a}{2^m} \cdot \left(1 - \frac{\epsilon}{2}\right) \cdot 2^m > (1-\epsilon) \cdot a$$

If we recall our definition of U(m, d), it is essentially the number of numbers that iterate to a value less than y^c as we just proved and hence $L(a) \cdot U(m,d)$ represents the number of numbers in $m \cdot 2^m \leq y < a$ that do so. This inequality proves theorem one as we know that the very left side of the inequality is a subset of the list of integers a and is, if we recall from way back in the beginning, less than a. Considering the limit as epsilon tends to 0, both sides of the inequality if we divide it all by a would be 1 as both sides would be a. Hence by what is known as the sandwich lemma, whatever is inbetween also must be equal to 1, and thus the proof is complete.

3.3 Eliahou

Eliahou's main result (Eliahou, 1993) sets a lower bound on the lowest possible cycle length for values above 2^{40} at 17,087,915 which would strongly suggest that no counter example to the Collatz conjecture exists. He first considers the below inequality whereby Ω denotes the amount of elements within a given Collatz orbit, Ω_1 represents all of the odd elements within that orbit, M is the highest value in the orbit and m is the lowest value in the orbit. This inequality is based on the idea that the proportion of odd elements in a cycle is effectively equal to $1/log_2 3$.

$$log_2(3 + M^{-1}) < \frac{\#\Omega}{\#\Omega_1} < log_2(3 + m^-1)$$

He then improves the inequality further with the addition of the variable μ $=(\#\Omega_1) \cdot (\sum n \in \Omega_1 n^{-1})$.

$$\frac{\#\Omega}{\#\Omega_1} < log_2(3 + \mu)$$

He then asserts that the product of all terms in a Collatz orbit starting from n will be the same as the product of all the terms starting from T(n) as the sequence is cyclical.

$$\prod_{n \in \Omega} n = \prod_{n \in \Omega} T(n)$$

Thus the below statement is inferred.

$$\prod_{n \in \Omega} \frac{T(n)}{n} = 1$$

From which by our definition of T(n) we can substitute and obtain the result below (we recall from Terras' paper that k is the stopping time).

$$\prod_{n \in \Omega_1} (3 + n^{-1}) = 2^k$$

k_1 is now assigned as the number of odd elements in a Collatz orbit and using the above result the initial inequality is rephrased.

$$log_2(3 + M^{-1}) < \frac{k}{k_1} < log_2(3 + m^{-1})$$

This can be further simplified whereby the M is removed as 1/M tends towards zero as M increases.

$$log_2(3) < \frac{k}{k_1} < log_2(3 + m^{-1})$$

Eliahou then defines functions K(m) and L(m). Let k_1 be equal to a variable l, then the function K(m) denotes the smallest positive integer k for which the above inequality holds and the function L(m) denotes the smallest positive integer l for which the above inequality holds. Using the corollary defined below and the inequality above, a computer is used to calculate the result that a cycle length must at least be about 17 million elements long.

$$\#\Omega \geq K(min\Omega)$$
$$\#\Omega_1 \geq L(min\Omega)$$

3.4 Tao

Tao's contribution to the Collatz conjecture (Tao, 2019) in late 2019 is the biggest breakthrough in recent years towards the problem. His main result in his paper "Collatz Orbits Attain Almost Bounded Values" states that, for any given function f(n) such that when n tends to infinity f(n) also tends to positive infinity, the minimum term within a given Collatz orbit of n will be less than f(n) for almost all values of n. The methods used to prove this statement are well beyond the scope of this paper but it is worth noting, if one were to read and try to understand his proof, that he breaks up the problem by defining three new functions Syrac, Geom and Log. The Syrac function which operates on the odd numbers finds the largest odd number that divides into a given n and the Geom function finds a geometric random variable of the mean n. The Log function is a random variable that takes in values of a finite subset of the natural numbers with a logarithmically uniform distribution. The essence of the paper is treating the Collatz function as a sort of system much those relating to partial differential equations that model everyday phenomena. The idea of the paper is to find a set of numbers that fairly exhibit how the Collatz function behaves for several iterations and then to extract how many of the numbers reached 1 or close to 1.

4 Other notable contributions

While the above mentioned are the most well-known and significant, the papers and results I am about to mention are also of note when it comes to attempts at tackling the conjecture. To begin, I would like to give credit to C.J Everett who, independentley of Terras, arrived at the same result as depicted in his original 1976 paper in his paper "Iteration of the number thoeretic function" (Everett, 1977). Heights of Collatz orbits have been frequently studied and the two main results to stem from this are from Guo-Gang Gao's paper "On consecutive numbers of the same height in the Collatz problem" (Gao, 1993) and from Lynn. E Garner's paper "On heights in the Collatz 3n + 1 problem"(Garner, 1985). The former proves that for a given orbit, there exists infinitely many such orbits with the same amount of terms (ie consecutive integers) and the same total stopping time. The latter proves that there are infinitely many pairs of neighbouring integers (e.g 3,4) which when the Collatz function is applied, have the same maxima (heights) and total stopping time. Parity of our starting integers appears to wildly change the outputs of the function however Hong Bo Yang's paper "About 3X + 1 problem" proves the claim that if there exists an odd integer for which the Collatz conjecture fails, there must exist infinitely many such odd integers. A bit more trivial but still to note is the result obtained from Ray P. Steiner's paper "A Theorem on the Syracuse Problem" (Steiner, 1977) which asserts that the only such "circuit" (a circuit being a certain number of odd integers followed by the same number of even integers in an orbit) that exists for the Collatz function is that of 1,2. A descent in a Collatz orbit is defined as consecutive term of the sequence being less than the previous. Thomas Brox's paper "Collatz cycles with few descents" (Brox, 2000) proves that d(C) (the number of descents) must be less than 2log$|C|$ where $|C|$ is the number of odd elements in the sequence. Further tightening tightening the criteria for Collatz cycles is Diego Domenici's "A few observations on the Collatz problem" (Dominici, 2009). The main result of this paper is that for the Collatz conjecture to hold, each of the numbers n in an orbit has to be able to take the form below:

$$n = \frac{2^m}{3^l} - \sum_{k=1}^{1} \frac{2^(b_k)}{3^k}$$

where $1 \leq l \leq m - 3 and 0 \leq$b$_1 < b_2 < ... < b_l \leq m - 4$

Similar, to Steiner's result, John L. Simons' "A simple (inductive) proof for the non-existence of 2-cycles for the 3x + 1 problem" (Simons, 2007) shows that there exists no periodic orbits whereby there are two groups of consecutive odd elements. Lastly, a result of more observation and interest rather than practicality, Jeffery C. Lagarias and K. Soundararajan's paper "Benford's Law for the 3x + 1 function" (Lagarias & Soundararajan, 2006) proves that Benford's law does indeed hold for the majority of starting values and their respective sequences.

5 Relevance and Potential Uses

The Collatz conjecture has no immediate significance and there is not a huge sense of urgency to prove nor disprove it. This may be due to change in the future however as seemingly unrelated areas of maths are linked to it in search of an answer. Much like the case of Fermat's Last Theorem, elliptic curves were not the first resource when working on the problem and the same logic can be applied to the Collatz conjecture. Despite this, there are some fields in which the conjecture, if proven, is predicted to have an impact. Ergodic theory namely is an often understated area of maths that relates to observing dynamical systems and their outcomes down the line, for example the climate. This would naturally have links to chaos theory and would provide insight into how the sensitive dependence on numerical initial inputs for a function in number theory, such as the Collatz function, could completely change their orbit/outcome even if the number differed by 1. Links between numbers which differ by one could be made with respect to their prime factorisations too. The most significant operation on a given number concerning its prime factors when applying the Collatz function is adding 1, as multiply by 3 and dividing by 2 do not drastically change much. Aside from expected advancements in these regions of mathematics, there is not much incentive currently to pursue a proof other than a modest sum of money and/or an intellectual challenge. For all we know the Collatz conjecture may be quite literally unsolvable from our base mathematical axioms (referring to Godel's Incompleteness Theorem) but nevertheless it serves as a reminder to the vast expanse of mathematics that is yet to be explored.

References

Allouche, J.-P. (1978-1979). Sur la conjecture de ”syracuse-kakutani-collatz”. *Seminaire de Théorie des Nombres de Bordeaux*, *8*, 1-16. Retrieved from `http://eudml.org/doc/182044`

Brox, T. (2000). Collatz cycles with few descents. *Acta Arithmetica*, *92*, 181–188.

Dominici, D. (2009). A few observations on the collatz problem. *International Journal of Applied Mathematics & Statistics*, *14*(J09), 97–107.

Eliahou, S. (1993). The 3x+ 1 problem: new lower bounds on nontrivial cycle lengths. *Discrete mathematics*, *118*(1-3), 45–56.

Everett, C. (1977). Iteration of the number-theoretic function f (2n)= n, f (2n+ 1)= 3n+ 2. *Advances in Mathematics*, *25*(1), 42–45.

Gao, G.-G. (1993). On consecutive numbers of the same height in the collatz problem. *Discrete mathematics*, *112*(1-3), 261–267.

Garner, L. E. (1985). On heights in the collatz 3n+ 1 problem. *Discrete mathematics*, *55*(1), 57–64.

Korec, I. (1994). A density estimate for the $3x + 1$ problem. *Mathematica Slovaca*, *44*(1), 85-89. Retrieved from `http://eudml.org/doc/32414`

Lagarias, J. C. (2010). The 3x+ 1 problem: An overview. *American Mathematical Society*.

Lagarias, J. C., & Soundararajan, K. (2006). Benford's law for the function. *Journal of the London Mathematical Society*, *74*(2), 289–303.

Martiny, T. (2015). *The 3n+ 1 problem: Scope, history and results* (Unpublished doctoral dissertation). University of Pittsburgh.

Simons, J. L. (2007). A simple (inductive) proof for the non-existence of 2-cycles of the 3x+ 1 problem. *Journal of Number Theory*, *123*(1), 10–17.

Steiner, R. P. (1977). A theorem on the syracuse problem. In *Proc. 7th manitoba conference on numerical mathematics* (pp. 553–559).

Tao, T. (2019). Almost all orbits of the collatz map attain almost bounded values. *arXiv preprint arXiv:1909.03562*.

Tao, T. (2020). *The notorious collatz conjecture.*

Terras, R. (1976). A stopping time problem on the positive integers. *Acta Arithmetica*, *3*(30), 241–252.