On a Generalization of the Collatz Conjecture

John Lesieutre

under the direction of
Mr. Zuoqin Wang
Massachusetts Institute of Technology

Research Science Institute
July 31, 2007

Abstract

In this paper we analyze a generalized version of the Collatz conjecture proposed by Zhang Zhongfu and Yang Shiming. We present a heuristic argument in favor of their conjecture and generalize a number of fundamental theorems from the original problem. We then obtain results related to properties of the generalized conjecture which do not arise in the original.

1 Introduction

The Collatz conjecture is one of the long-standing open problems of mathematics. Apparently first posed by Lothar Collatz in the 1930s, it has since withstood every attempt at proof. Since the problem was passed around orally for many years before any articles were written about it, it goes by a number of names, including Kakutani's problem, Hasse's algorithm, Ulam's problem, and the Syracuse problem. The problem is related to a wide range of topics in mathematics, including number theory, computability theory, and the analysis of dynamic systems. The statement of the Collatz conjecture involves the mapping $T : \mathbb{N} \to \mathbb{N}$ where

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

Let $T^{(i)}(x)$ denote the result of i iterations of T on x. We call the sequence of $T^{(i)}(x)$ the *trajectory* or x. The Collatz conjecture states that for all x there is some i such that $T^{(i)}(x) = 1$. For example, if we start with $x = 7$, the iteration goes $7 \to 22 \to 11 \to 34 \to 17 \to 52 \to 26 \to 13 \to 40 \to 20 \to 10 \to 5 \to 16 \to 8 \to 4 \to 2 \to 1$. The conjecture has been verified for initial values of x up to 2^{50}, but proof remains elusive [5]. However, there exist many partial results; one theorem shows that asymptotically, "nearly all" numbers do eventually reach 1, though there still may be an infinite number of exceptions. Other theorems estimate the number of iterations a typical number takes to reach 1 or show that the conjecture holds for certain broad classes of integers.

Here we deal with a generalization of the Collatz conjecture proposed by Zhang Zhongfu and Yang Shiming [8]. We examine the mapping $T_n : \mathbb{N} \to \mathbb{N}$ where

$$T_n(x) = \begin{cases} \dfrac{x}{p_{i_1} p_{i_2} \cdots p_{i_k}} & \text{where the } p_i \text{ are exactly the primes} \leq p_n \text{ dividing the numerator,} \\ p_{n+1}x + 1 & \text{if no prime } p_i \leq p_n \text{ divides } x. \end{cases}$$

For example, with $n = 3$, we multiply by 7, add 1 and then divide out by 2, 3, and 5. Taking $x = 37$ as our initial value, the trajectory of x under iteration of T_n is $37 \to 26 \to 13 \to 92 \to 46 \to 23 \to 162 \to 27 \to 9 \to 3 \to 1$. Here we investigate the properties of this function under iteration. We note that our general case contains non-trivial cycles, which are conjectured not to occur in the original problem, and thus formulate our motivating conjecture:

Conjecture 1. *For any p_{n+1} and x_0, repeated iteration of $T_n(x)$ on x_0 will eventually yield $T_n^{(i)}(x_0) = 1$ or $T_n^{(i)}(x_0)$ will enter one of a finite number of non-trivial periodic cycles.*

These non-trivial cycles are discussed more fully in Section 3.4. The simplest example occurs when $p_{n+1} = 11$, when we have the periodic trajectory $17 \to 47 \to 37 \to 17 \to \cdots$.

The Collatz conjecture is very similar to the case $n = 1$ in our more general statement, except that it also excludes the possibility of non-trivial cyclic trajectories. A proof of the general conjecture would not resolve the Collatz conjecture as there could still exist non-trivial cycles. It is known that if such cycles exist they must have period at least 275,000 [5].

While a great body of work pertaining to the original conjecture (our $n = 1$) exists, the only paper concerning this generalization is the paper by Zhang Zhongfu and Yang Shiming in which the generalized form is suggested. They do not analyze in it much depth, only demonstrating the existence of non-trivial cycles and discussing a slight further generalization.

Using a wide range of techniques, mathematicians have proved numerous results about the Collatz conjecture, from demonstrating that it holds for "almost all" large x to showing the truth of certain cases to demonstrating its connections to rational approximations of $\log_2 3$. It has been attacked with graph theory, analysis, elementary number theory, and ergodic theory, with each of these approaches yielding new results. For a more complete overview, the reader may consult Lagarias's excellent survey article [5].

2 Probabilistic Analysis of Trajectories

Our conjecture is equivalent to the assertion that there are only a finite number of cycles and that the mapping contains no divergent trajectories, that is, there is no x_0 for which $T_n^{(i)}(x_0) \to \infty$ as $i \to \infty$. If we assume that for large initial values, the function acts sufficiently randomly that residues are equally distributed modulo $\left(p_n^{\#}\right)^k = (p_1 p_2 \dots p_n)^k$, we may calculate expected trends in the size of the numbers in our trajectory and then supply some supporting evidence for part our conjecture. We specifically investigate the expected ratio between consecutive terms not divisible by primes $\leq p_n$ in trajectories. For a large integer n_0, we may estimate the ratio between n_0 and n_1, the next term coprime to $p_n^{\#}$.

Heuristic 1.

$$\frac{n_1}{n_0} = p_{n+1} \prod_{j=1}^{n} p_j^{-p_j/(p_j-1)^2}$$

Proof. Before proceeding, we must establish two preliminary lemmata which are necessary to our argument.

Lemma 2. *If set A is a reduced residue system modulo $p_n^{\#} = p_1 p_2 \dots p_n$, then any p_i, $(i \leq n)$ divides precisely $\frac{\#(A)}{p_i - 1}$ elements of the set $kA + 1$ for any k not divisible by any prime less than or equal to p_n.*

Proof. Consider the set of residues of A taken mod p_i. There are no 0 residues by definition of A, since $p_i \mid p_n^{\#}$. However, the residues must be evenly distributed over the other $p_i - 1$ possibilities. Suppose that some residue k_1 in A, occurs more often than some other residue k_2. When we multiply A by a constant relatively prime to the modulus $p_n^{\#}$, we get the same residue system. But there is one such number C so that $k_1 \equiv C k_2 \pmod{p_n^{\#}}$. Therefore k_1 and k_2 must occur with equal frequency among the elements of A taken mod p_i. So the residue -1 occurs $\frac{\#(A)}{p_i - 1}$ times in the set A taken mod p_i. Now, the set kA is the same residue system as A for $(k, p_n^{\#}) = 1$. So the residue -1 occurs there this same number of times.

Now, in the set $kA+1$ the residue 0 occurs exactly this many times and exactly that many elements of $kA+1$ are divisible by p_i. $\square$

Lemma 3. *We have*

$$\prod_{i=1}^{\infty}(n^{-i})^{(n^{-i})} = n^{-n/(n-1)^2}$$

for $n > 1$.

Proof. First, note that

$$\prod_{i=1}^{\infty}(n^{-i})^{(n^{-i})} = \exp\left(\log\prod_{i=1}^{\infty}(n^{-i})^{(n^{-i})}\right) = \exp\left(\sum_{i=1}^{\infty} n^{-i}\log n^{-i}\right) = \exp\left(-\log n\sum_{i=1}^{\infty} in^{-i}\right).$$

So now let $k = \sum_{i=1}^{\infty} in^{-i} = \frac{1}{n} + \frac{2}{n^2} + \frac{3}{n^3} + \ldots$ so that

$$\begin{aligned} nk &= 1 + \frac{2}{n} + \frac{3}{n^2} + \frac{4}{n^3} + \ldots \\ &= \left(1 + \frac{1}{n} + \frac{1}{n^2} + \frac{1}{n^3} + \ldots\right) + \left(\frac{1}{n} + \frac{2}{n^2} + \frac{3}{n^3} + \ldots\right) \\ &= \frac{1}{1-\frac{1}{n}} + k \end{aligned}$$

whence $k = \frac{n}{(n-1)^2}$. Thus our product is equal to

$$\exp\left(-\log n\sum_{i=1}^{\infty} in^{-i}\right) = (e^{\log n})^{\left(-n/(n-1)^2\right)} = n^{-n/(n-1)^2}.$$

$\square$

Now we return to the heuristic. The first step is to determine the expected factor by which a given prime p_i will decrease a number when we divide out by small factors. Consider a prime $p_i \leq p_n$. The probability that p_i will divide $p_{n+1}x+1$ exactly m times is equal to the probability that it will divide it once, times the probability that it will divide it $m-1$ more times, times the probability that it will divide it no more times after that. By Lemma 2, with

$k = p_{n+1}$ and $n_0 \in A$, the chance that $p_{n+1}n_0 + 1$ is divisible by p_i once is $\frac{1}{p_i - 1}$. The chance that $p_{n+1}n_0 + 1$ will be divisible by p_i exactly m times is then $\left(\frac{1}{p_i-1}\right)\left(\frac{1}{p_i}\right)^{m-1}\left(\frac{p_i-1}{p_i}\right) = \frac{1}{p_i^m}$, as described above. It may seem irrational that we treat the first division differently than the others, but to see why this is plausible, consider the $3x + 1$ mapping of the Collatz conjecture; when we apply $3x + 1$ to an odd number we may always take out a factor of 2 immediately, but after that we treat it as random. Now, when p_i divides a number m times, the factor of reduction is $\frac{1}{p_i^m}$. We then expect p_i to reduce $p_{n+1}x + 1$ by an average factor of

$$\prod_{m=1}^{\infty} \left(\frac{1}{p_i^m}\right)^{\left(\frac{1}{p_i^m}\right)} = \prod_{m=1}^{\infty} (p_i^{-m})^{(p_i^{-m})} = p_i^{-p_i/(p_i-1)^2}$$

by Lemma 2. The expected total factor of reduction due to all primes up to p_n is therefore

$$\prod_{j=1}^{n} p_j^{-p_j/(p_j-1)^2}.$$

The expected ratio is then found as

$$\frac{n_1}{n_0} = p_{n+1} \prod_{j=1}^{n} p_j^{-p_j/(p_j-1)^2}.$$

□

For $n = 1$, the Collatz conjecture, the expected ratio is 3/4, which suggests that on a large scale, trajectories tend to decrease and divergence is unlikely. For $n = 2$, the expected ratio is $\frac{5\sqrt[4]{3}}{12} \approx 0.54836$, which suggests that trajectories will tend to decrease even faster. Appendix A gives a graph of the expected ratio for the first 10,000 primes. We might expect that this ratio would tend to 0 as n increases without bound since we are dividing out by an ever-larger number of primes. We claim that this ratio converges: note first that

$$
\begin{aligned}
p_{n+1}\prod_{j=1}^{n} p_j^{-\frac{p_j}{(p_j-1)^2}} &= \exp\left(\log p_{n+1}\sum_{j=1}^{n}\log p_j^{-p_j/(p_j-1)^2}\right)\\
&= \exp\left(\log p_{n+1}-\sum_{j=1}^{n}\frac{p_j}{(p_j-1)^2}\log p_j\right)\\
&= \exp\left(\log p_{n+1}-\sum_{j=1}^{n}\left(\frac{1}{p_j}+\frac{1-2p_j}{p_j(p_j-1)^2}\right)\log p_j\right)\\
&= \exp\left(\log p_{n+1}-\sum_{j=1}^{n}\frac{\log p_j}{p_j}-\sum_{j=1}^{n}\frac{(1-2p_j)\log p_j}{p_j(p_j-1)^2}\right)
\end{aligned}
$$

We show in Appendix B that $\sum_{p\le N}\frac{\log p}{p} = \log N + C_1 + o(1)$ for some constant C_1. Some simple calculus shows that the second sum converges over all the naturals and therefore its sum over the primes is some constant C_2. With x as p_{n+1}, the limit of our product is

$$
e^{\log p_{n+1}-\log p_{n+1}-C_1-C_2-o(1)} = e^{-(C_1+C_2+o(1))}
$$

which ensures convergence to some constant C. Calculating the value over the first $10,000$ primes shows that $C \approx 0.520$. We also note that the expected ratio seems to be less than 1 for all p_{n+1}, and our apparently tight error bound combined with our calculations over the first 10,000 p_{n+1} provide strong evidence for this.

3 Structure of Trajectories

It is very difficult to directly analyze the trajectory of a given number, but we may still obtain some results. To begin, we present generalizations of several basic results of Terras and others presented in Lagarias's paper [5].

3.1 The Encoding Matrix

One of the most important tools in the study of the Collatz conjecture has been Terras's encoding vector (called the parity vector by Lagarias), which contains entries corresponding to the action of the function for each iteration (whether the image of x is $3x+1$ or $x/2$) [7]. In our general case we define $A_k(n)$ to be the matrix with (i,j)-entry $0 \le m_{ij} < p_i \equiv T_n^{(k)}(x)$ (mod p_i). The matrix represents the action of the function on the initial number under repeated iteration. For example, the iteration of T on 13 with $p_{n+1} = 7$ yields $13 \to 92 \to 46 \to 23 \to 162 \to 27 \to 9 \to 3 \to 1$. As an encoding matrix, we write

$$A_9(13) = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 1 \\ 3 & 2 & 1 & 3 & 2 & 2 & 4 & 3 & 1 \end{pmatrix}$$

For example, the 3$^{\text{rd}}$ column corresponds to 46, which is congruent to 0 mod 2, 1 mod 3, and 1 mod 5. The final column represents 1. We denote the n^{th} column by $\alpha_n(x)$ and the vector of the n^{th} row as $\beta_n(x)$. We may use the encoding matrix to study the trajectory of x. We may derive an important formula for $T_n^{(k)}(x)$ in terms of $A_k(x)$.

Explicitly, letting α_i denote the i^{th} column of A_k, β_i the i^{th} row of A_k, and defining

$$a(A_k, j) = \#\{\alpha_i : j \le i \le k, \alpha_i \text{ has all nonzero entries}\},$$

$$b(\beta_i, j) = \#\{m_{ie} : j \le e \le k, m_{ie} = 0\},$$

and

$$c(\alpha_i) = \begin{cases} 0 & \text{if some entry of } \alpha_i \text{ is } 0, \\ 1 & \text{otherwise,} \end{cases}$$

we can write

Theorem 4.

$$T_n^{(k)}(x) = \lambda_k(x)x + \rho_k(x)$$

where

$$\lambda_k(x) = \frac{p_{n+1}^{a(A_k,1)}}{\prod_{j\leq n} p_j^{b(\beta_j,1)}}$$

and

$$\rho_k(x) = \sum_{i=1}^{k} \frac{c(\alpha_i)p_{n+1}^{a(A_k,i+1)}}{\prod_{j\leq n} p_j^{b(\beta_j,i)}}$$

We give a proof of this in Appendix C.

Essentially, λ_k measures the overall multiplicative factor and ρ_k makes small corrections for the +1s. Note that we may sometimes write $\lambda_i(A_k)$ and $\rho_i(A_k)$, and here A_k is the matrix representing the trajectory of x. We may prove various other theorems about the densities and structures of trajectories, but first, we require some definitions.

3.2 Matrices: Attainable, Admissible, and Inflating

Definition 1. *We call a matrix B_k attainable if there exists some n_0 such that $A_k(n_0) = B_k$*

To see why not all matrices are attainable, note that no encoding matrix may contain the sequence

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

since this would correspond to taking an odd number, dividing it by 3, and getting an even number. If two vectors α_1 and α_2 may occur as consecutive α_i and α_{i+1} we say that α_2 is a *successor* of α_1, and we denote the number of possible successors of a vector α by $n(\alpha)$. A matrix is attainable if and only if α_{i+1} is a successor of α_i for all $0 \leq i < k$.

We may further classify the attainable matrices. Following Lagarias, we say

Definition 2. *A matrix A_k is admissible with length k if it satisfies* *(i)* $\lambda_k(A_k) < 1$

(ii) $\lambda_i(A_k) > 1$ *for* $1 \le i < k$

Essentially, this means that the total multiplicative factor is greater than 1 for the first $k-1$ columns and less than 1 after k columns. Note that λ_k is never equal to 1, since it is a power of p_{n+1} over a product of powers of smaller primes. Similarly, we define

Definition 3. *A matrix A_k is inflating with length m if it satisfies $\lambda_i(A_k) > 1$ for $1 \le i \le m$.*

This means that the multiplicative factor is always greater than 1 for the first m columns of the matrix.

It is clear that there are $(p_n^{\#})^k$ different matrices of length k for a given p_n. But how many of these are attainable? It is difficult to analyze this in the general case but not altogether too difficult to solve the case $n = 2$ ($p_{n+1} = 5$). We find that the number of attainable matrices grows exponentially with k. A derivation of the rate of growth and proof may be found in Appendix D, where we prove the following result.

Theorem 5. *The number of attainable matrices of length k in the case $n = 2$ is asymptotic to m^k where m is the unique real solution to $x^5 - x^4 - 2x^3 - 2x^2 - x - 1 = 0$.*

While this theorem does not apply to the general case, it suggests how it may be solved and also that the frequency of allowed matrices should decrease as k increases.

3.3 Some Fundamental Results on Encoding Matrices

The encoding matrix possesses a number of interesting properties that make it useful in the study of our problem. We generalize two important theorems of Terras below [7].

Our first result states that the numbers whose encoding matrices of length k are identical fall in arithmetic sequences.

Theorem 6. *Let* $S(A_k) = \{x : A_k(x) = A_k\}$ *for an attainable matrix* A_k*. Then* $S(A_k) = \{x_0(A_k) + ip_n^{\#} \prod_{j=1}^{k-1} n(\alpha_j)\}$ *where* $n(\alpha_i)$ *is the number of possible successors to* α_i *and* $x_0(A_k)$ *is the least* x *satisfying* $A_k(x_0) = A_k$*.*

Proof. We proceed by induction on k, the length of A. Clearly, when $k = 1$ our theorem holds. By the Chinese Remainder Theorem, some x_0 exists in the interval $[1, p_n^{\#}]$ satisfying $A_1(x_0) = A_1$. Also, it is clear that $A_1(x_0) = A_1(x)$ if and only if $x = x_0 + p_n^{\#} i$. Further, we see that after we take the image of the members of this sequence under A_1 they must be evenly distributed over the possible successors of α_1 when taken modulo $p_n^{\#}$. To see this, consider two cases: when we divide by some product of primes $\leq p_n$ we get a sequence of numbers with some fixed residues modulo each p_j by which we did not divide, and all possible choices for the residue for each p_i by which we did divide, and these are equally distributed. The number of these possibilities is exactly $n(\alpha_1)$. And when we map by $p_{n+1}x + 1$ we have only one possible residue modulo $p_n^{\#}$ which is still the equidistrubtion we seek.

Now, assume the theorem holds for a length k. That is, we assume $S(A_k) = \{x_0(A_k) + ip_n^{\#} \prod_{j=1}^{k-1} n(\alpha_j)\}$. After we take the image of some x under the mapping described by A_k, we must be left with one of the $n(\alpha_k)$ possible residues. None of these possible residues occurs more often than any other by an argument analogous to that above. Our sequence must therefore be split in $n(\alpha_k)$ new arithmetic sequences with the same common difference. Therefore the common difference is multiplied by this factor $n(\alpha_k)$. There must also still exist some x_0 for each A_k. These facts imply our inductive step and thereby our theorem. □

We generalize two functions of Terras [7]. Let $\sigma(x)$, the *stopping time* function, be the smallest k so that $T_n^{(k)}(x) < x$, that is, the number of iterations it takes for a number to reach a number less than itself. In the Collatz problem, a proof that every number has a finite stopping time would amount to a proof of the conjecture since we could inductively show that each number > 1 must reach a smaller number, which we would already know by the induction reaches 1. Also let $\omega(x)$, the *coefficient stopping time* function, be the smallest k such that $\lambda_k(x) < 1$. Then let $H_k = \{x : \sigma(x) = k\}$ and $I_k = \{x : \omega(x) = k\}$. It is clear that $H_k \subseteq I_k$. Further, we may show

Theorem 7. *$I_k \setminus H_k$ is a finite set.*

Proof. Our proof closely follows that of Terras' for the Collatz problem [7]. To begin, note that when A_k is an admissible matrix of length k and $x \in S(A)$, then for $1 \leq i \leq k-1$ we have

$$T_n^{(i)}(x) \geq \lambda_k(A_k)x \geq x$$

by the definition of admissibility. Thus, all x with length-k matrix A_k have a stopping time $\sigma(x) \geq k$. Now define $\epsilon_k(A_k) = 1 - \lambda_k(A_k)$ for admissible A_k. Then we have $T_n^{(k)}(x) = x + (\rho_k(A_k) - \epsilon_k(A_k)x)$. x has stopping time exactly k whenever $\rho_k(A_k) - \epsilon_k(A_k)x < 0$, that is, $x > \frac{\rho_k(A_k)}{\epsilon_k(A_k)}$. Since this depends only on A_k, there may be only a finite number of $x \in S(A_k)$ for which $\sigma(x) \neq \omega(x)$. So if A_k is admissible then all sufficiently large x in $S(A_k)$ have stopping time k.

Now suppose that A_k is not admissible. If no initial segment of A_k is admissible then A_k is inflating and $\lambda_k(A_k) > 1$ so no members of $S(A_k)$ have stopping time k. If some initial segment of A_k is admissible then both $\omega(A_k)$ and $\sigma(A_k)$ are less than k for sufficiently large elements of $S(A_k)$ and only a finite number of elements of $S(A_k)$ may have stopping time k. This, taken with our first argument and noting that there are only a finite number of matrices of length k, proves our theorem. □

Corollary 8. *H_k and I_k have the same asymptotic density. Further, this asymptotic density is given by*

$$F(k) = \sum_{\substack{A \text{ admissible} \\ \text{length}(A) \leq k}} \text{weight}(A)$$

where

$$weight(A) = \left(p_n^{\#} \prod_{j=1}^{k-1} n(\alpha_j) \right)^{-1}$$

.

Proof. H_k and I_k are both infinite sets. They share all but a finite number of elements, which implies equal asymptotic density. Our weight function is derived from Theorem 6 and makes sense on a fairly intuitive level; when all possible successor matrices to some matrix are still admissible, our weight remains unchanged. But when some of these are not, our weight corresponding to that particular branch is reduced by some factor. The sum of the weights over all admissible vectors of length k must be the density of numbers with stopping time $\leq k$. □

Conjecture 2. $\lim_{k\to\infty} F(k) = 1$

We conjecture that the density of numbers with finite stopping time is 1. This is a known result in the Collatz conjecture. We have attempted to resolve this conjecture by using combinatorial methods to bound the sizes of the sets of attainable, admissible, and inflating vectors of length k and then applying a bound on the weight function. However, our methods have not achieved strict enough bounds to ensure convergence to 1.

3.4 Non-trivial Cycles

One conjecture about the Collatz mapping is that for our $n = 1$, there are no non-trivial cycles (i.e., no numbers go into repeating loops that do not include 1). However, in our general case, many such cycles exist and we may find them fairly easily. As noted in the introduction, the simplest such cycle occurs with $p_{n+1} = 11$ when we have the loop $17 \to 47 \to 37 \to 17 \to \cdots$. A given p_{n+1} may lead to multiple such cycles; $p_{n+1} = 61$ seems to be the smallest such case, though we have not proven this. We may show that a given p_{n+1} may have a only a finite number of non-trivial cycles:

Theorem 9. *There exist only a finite number of cycles for a given p_{n+1}.*

Proof. Our theorem is a consequence of Theorem 7. Consider the set

$$C = \{x : x \text{ is the smallest element of a cycle}\}$$

. All elements of C have some finite $\omega(x)$ since when $T_n^{(p)}(x) = x$, $\lambda_p(A_p(x)) < 1$ because $\rho_p(A_p(x))$ must be positive. However, no $x \in C$ has a finite $\sigma(x)$ since x is the smallest element of the cycle. By Theorem 7, C must be a finite set. □

While Theorem 3.4 proves part of our conjecture (since we know there may not be an infinite number of cycles), it is difficult to find any stronger results. We do not know exactly how many cycles we expect for a given p_{n+1}, or even whether large p_{n+1} contain more or less cycles than smaller ones. $p_{n+1} = 61$ seems to be the only prime less than 300 for which more than one cycle exists, but we can not be sure that there are other cycles that eluded our computer search. A list of a few illustrative cycles and the C++ code of the program used for our search may be found in Appendix E.

4 A Further Generalization

We may consider a more general version of our problem by studying a version where instead of mapping with $p_{n+1}x+1$ we use $p_{n+1}x+K$ with $(K, p_n^{\#}) = 1$. We still divide out by primes $\leq p_n$ as in our original problem. These generalization is also posed and briefly discussed by Zhang and Yang [8]. We may derive several simple theorems based on this map relating cycles between two different K in certain cases.

4.1 Basic Results

Theorem 10. *A cycle in the mapping using $p_{n+1}x+k$ may be transformed to one in $p_{n+1}x+kq$ when q is not divisible by any prime $\leq p_n$.*

Proof. Consider a trajectory $m_1 \to m_2 \to \cdots \to m_i \to \cdots$ that eventually becomes cyclic and is in its cycle for $i \geq s$. We have $m_{s+1} = (p_{n+1}m_s+k)/F_{s+1}$ where the F_i are divisible only by primes $\leq p_n$. So now $(qm_{s+1}) = (p_{n+1}(qm_s)+kq)/F_i$ and we see that $(qm_s, qm_{s+1}, \ldots)$ is a cycle in the mapping with $p_{n+1}x+kq$. □

We may also obtain the following stronger but less general result:

Theorem 11. *Consider a mapping where we replace $p_{n+1}x+1$ by $p_{n+1}x+p_{n+1}^a k$. For a given k and p_{n+1}, all mappings of this form contain the same number of cycles.*

Proof. As above, consider a trajectory $m_0 \to m_1 \to m_2 \to \cdots \to m_i \to \cdots$ so that $p_{n+1}m_{i-1}+p_{n+1}^a k = F_i m_i$ with F_i containing only prime factors $\leq p_n$ and m_i containing no such factors. Now define l_i and n_i so that $m_i = p_{n+1}^{l_i} n_i$ with $p_{n+1} \nmid n_i$. Now consider two cases:

Case 1: If $l_i \geq a$:

$$m_{i+1} = (p_{n+1}m_i + p_{n+1}^a k)/F_{i+1} = (p_{n+1}(p_{n+1}^{l_i} n_i) + p_{n+1}^a k)/F_{i+1}$$
$$p_{n+1}^{l_{i+1}} n_{i+1} = (p_{n+1}^a (p_{n+1}^{l_i-a+1} + k))/F_{i+1}$$

and we have $l_{i+1} = a$.

Case 2: If $l_i < a$:

$$m_{i+1} = (p_{n+1}m_i + p_{n+1}^a k)/F_{i+1} = (p_{n+1}(p_{n+1}^{l_i} n_i) + p_{n+1}^a k)/F_{i+1}$$
$$p_{n+1}^{l_{i+1}} n_{i+1} = (p_{n+1}^{l_i+1}(n_i + p_{n+1}^{a-l_i-1} k))/F_{i+1}$$

and so $l_{i+1} = l_i + 1$. Combining these two cases, we see that for some t, $p_{n+1}^a \mid m_j$ and $p_{n+1}^{a+1} \nmid m_j$ for all $j \geq t$. So now note that we have

$$m_{t+1} = (p_{n+1}m_t + p_{n+1}^a k)/F_{t+1}$$

$$\frac{m_{t+1}}{p_{n+1}^a} = \left(p_{n+1}\frac{m_t}{p_{n+1}^a} + k\right)/F_{t+1}$$

so our cycle in the $p_{n+1}x + p_{n+1}^a k$ gives us a cycle in $p_{n+1}x + k$, which is the case $a = 0$. By the preceding theorem, taking $q = p_{n+1}$, we may use induction and see that our cycle in $p_{n+1}x + p_{n+1}^{a_1} k$ corresponds to exactly one cycle in $p_{n+1}x + p_{n+1}^{a_2} k$ for any choice of a_1 and a_2 and both of these mappings contain exactly the same number of cycles. □

4.2 Prime 2-Cycles in the Case $n = 2$

Looking at examples we notice that the $3x + K$ problem sometimes contains cycles of length 2, with $x_1 \to x_2 \to x_1 \to \cdots$, and in some of these $(x_1, x_2) = 1$. We call such cycles with coprime x_1, x_2 and length 2 *prime 2-cycles*. For example, with $K = 7$, we have $5 \to 11 \to 5 \to \cdots$. We completely characterize these cycles in the case $n = 1$. A list of a few typical cycles in the case $3x + K$ may be found in Appendix F.

Theorem 12. *All prime 2-cycles in T_2 are of the form $x_1 = \frac{2^a+3}{(2^a+3, 2^b+3)}$, $x_2 = \frac{2^b+3}{(2^a+3, 2^b+3)}$, and $K = \frac{2^{a+b}-9}{(2^a+3, 2^b+3)}$.*

Proof. We seek solutions to the equations $\frac{3x_1+K}{2^a} = x_2$ and $\frac{3x_2+K}{2^b} = x_1$. So we have

$$\frac{3\frac{3x_1+K}{2^a} + K}{2^b} = x_1$$

$$2^{a+b}x_1 = 9x_1 + 3K + 2^a K$$

$$(2^{a+b} - 9)x_1 = (2^a + 3)K$$

. Then we see that x_1 must be some constant multiple of $\frac{2^a+3}{(2^a+3,2^{a+b}-9)}$. But for our cycle to be prime, it must be equal to exactly that or x_1 and K have a common factor. Also note that $(2^a+3, 2^{a+b}-9) = (2^b(2^a+3), 2^{a+b}-9)$ since 2^b is coprime to both these odd numbers. Now $(2^a+3, 2^b(2^a+3)-(2^{a+b}-9)) = (2^a+3, 3(2^b+3)) = (2^a+3, 2^b+3)$. Then we have $x_1 = \frac{2^a+3}{(2^a+3,2^b+3)}$ and we may substitute in to get $K = \frac{2^{a+b}-9}{(2^a+3,2^b+3)}$ and $x_2 = \frac{2^b+3}{(2^a+3,2^b+3)}$. In substituting, we also note that 2^a and 2^b are in fact the largest powers of two by which we may divide. Thus all solutions are of our form and all numbers of our form are actual solutions. □

4.3 Generalization of Other Results

We may also generalize most of our earlier results to this case, though we omit these proofs. A simple extension of Lemma 2 allows us to use the same heuristic and arrive at the same argument against the existence of divergent trajectories. Our formula for $\lambda_k(x)$ remains unchanged (since it is simply a multiplicative factor and does not consider the added terms), and we change $\rho_k(x)$ so that

$$\rho_k(x) = K \sum_{i=1}^{k} \frac{c(\alpha_i) p_{n+1}^{a(A_k,i+1)}}{\prod_{j \leq n} p_j^{b(\beta_j,i)}}$$

This result is clearly analogous to our original formula for $\rho_k(x)$ except it is now multiplied by a constant factor of K. Using this result and others, we may use proofs essentially identical to those above to show that there are a finite number of cycles for $p_{n+1}x + K$. One interesting difference is that since we require $x > \frac{\rho_k(A_k)}{\epsilon_k(A_k)}$ and ρ_k has been multiplied by a factor of K, we may have a larger set of x for which $\omega(x) \neq \sigma(x)$ and therefore possibly more cycles. Theorem 10 implies that $K = pq$ must contain at least as many cycles as $K = p$, which provides further evidence that the larger K, the more cycles we expect. However, it remains difficult to analyze the predict number of cycles for a given K.

5 Conclusion

We have analyzed a generalization of the Collatz conjecture and demonstrated extensions of various fundamental results in the original problem to our version. We have given a heuristic argument against the existence of divergent conjectures. Further, we have shown that only a finite number of cycles may exist for a given p_{n+1} and have extended this result to a further generalization with $p_{n+1}x + K$. However, many questions remain. It is known that in the Collatz mapping almost all integers eventually iterate to 1. We desire a proof of an analogous result in the general case. Many problems also remain in the analysis of non-trivial cycles, especially in estimating their number for a given p_{n+1} and K. Of course, the greatest open problem is the generalized conjecture itself, a resolution of which would provide a strong new result on the Collatz conjecture and better our understanding of such dynamic systems.

6 Acknowledgements

I am most grateful to my mentor, Mr. Zuoqin Wang of MIT, for his constant guidance, encouragement, and assistance. He located a problem for me to work on, found resources to read, suggested lines of research, and helped me along when I encountered difficulties. I am also indebted to my tutor, Chris Mihelich, for his helpful suggestions for improvements and careful editing, as well as for his mathematical help, especially the proof in Appendix B. I am also obliged to Kartik Venkatram whose advice was invaluable during the revision process. I thank the Center for Excellence in Education, the Research Science Institute, and the RSI staff for making it possible for me to pursue this research. Finally, I thank my fellow RSI students for their support and for showing (or at least feigning) interest in my research.

References

[1] Paul J. Andaloro. The $3x + 1$ problem and directed graphs. Fibonacci Quarterly 40 (2002), no. 1, 43–54.

[2] Stefan Andrei, Manfred Kudlek, and Radu Niculescu. Some results of the Collatz problem. Acta Informatica 37 (2000), 145–160.

[3] Ranan B. Banerji. The $3x + 1$ problem and number representation. Available at `http://www.sju.edu/~rbanerji/rb/papers/paper1.htm` (2004/6/30).

[4] Barry Brent. $3x + 1$ dynamics on rationals with fixed denominator. Available at `arXiv:math.DS/0203170` (2004/7/21).

[5] Jeffrey C. Lagarias. The $3x + 1$ problem and its generalizations. Available at `http://www.cecm.sfu.ca/organics/papers/lagarias/` (2004/6/29).

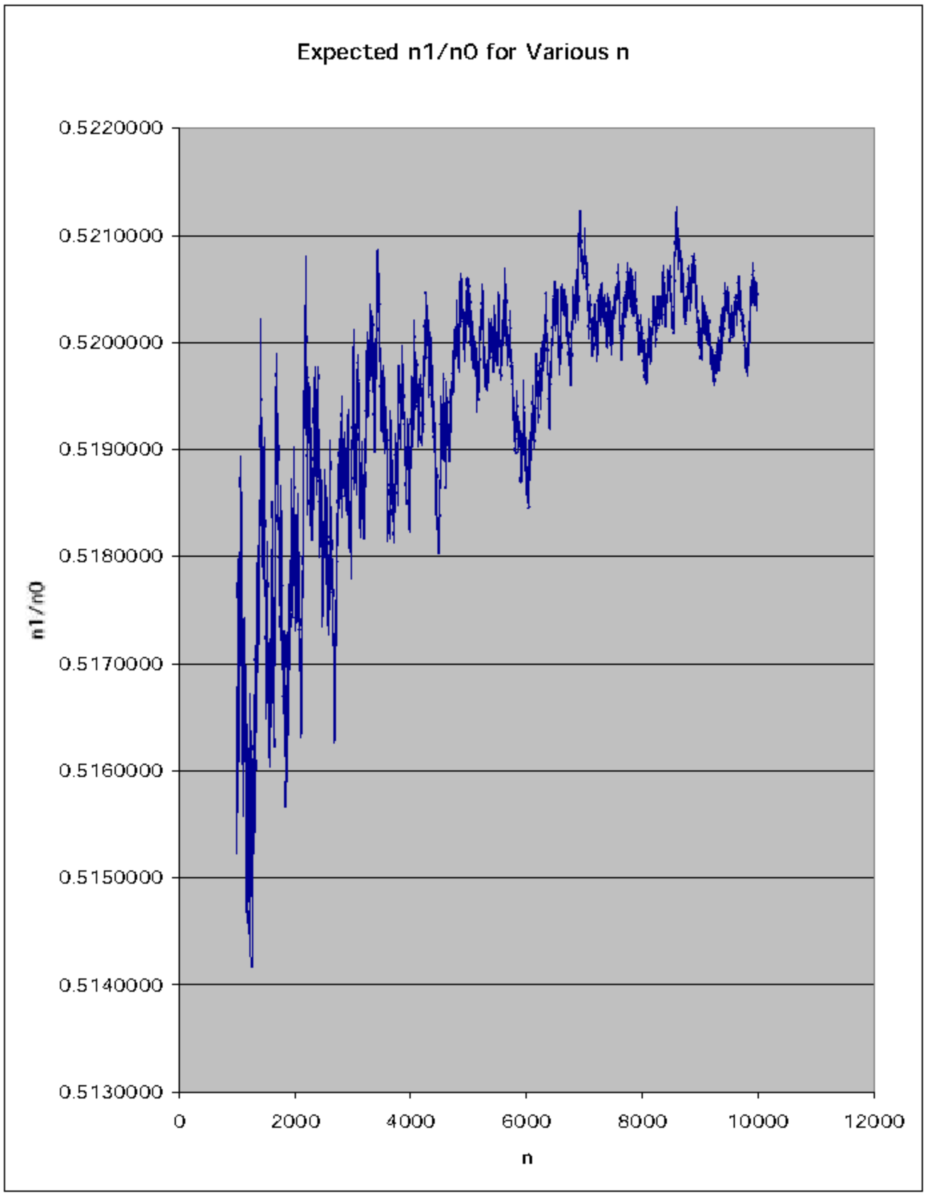
[6] K.R. Matthews. The generalized $3x + 1$ mapping. Available at `http://www.numbertheory.org/pdfs/survey.pdf` (2004/6/28).

[7] Riho Terras. A stopping time problem on the postive integers. Acta Arithmetica 30 (1976), 241–252.

[8] Zhang Zhongfu and Yang Shiming. Ying She Shu Lie Wen Ti. Shu Xue Chuan Bo 22 (1998), no. 2, 76–88.

A Graph of Expected n_1/n_0 for $1000 < n < 10000$

Our heuristic argument shows that there is some certain finite expected ratio between consecutive terms coprime to $p_n^\#$ in the $p_{n+1}x+1$ mapping and that this ratio converges to some constant C. The following graph illustrates the expected value of this ratio over the first $10,000$ primes.



B Proof of Formula for $\sum_{p\le N} \frac{\log p}{p}$

Lemma 13.

$$\sum_{p\le N} \frac{\log p}{p} = \log N + c + o(1)$$

This result is a routine exercise in analytic number theory and I thank Chris Mihelich for the argument.

Proof. We proceed using summation by parts. To begin, we have

$$\sum_{p\le N} \frac{\log p}{p} = \sum_{2\le n\le N} \left(\frac{\log n}{n}\left(\pi(n) - \pi(n-1)\right)\right) = \sum_{2}^{N+1} \left(\frac{\log n}{n}\Delta E^{-1}\pi(n)\delta n\right).$$

The first equality follows since $\pi(n) - \pi(n-1)$ is 1 if and only if n is prime, and the second follows from definitions. Then, applying summation by parts, our sum is equal to

$$\left[\frac{\log n}{n}\pi(n-1)\right]_2^{N+1} - \sum_{2}^{N+1} \pi(n)\Delta\left(\frac{\log n}{n}\right)\delta n$$

The prime number theorem and the series expansion of the logarithmic integral tell us that

$$\pi(n) = \frac{n}{\log n} + \frac{n}{(\log n)^2} + O\left(\frac{n}{(\log n)^3}\right).$$

So we we have

$$\begin{aligned}
\Delta\left(\frac{\log n}{n}\right) &= \frac{\log(n+1)}{n+1} - \frac{\log n}{n} = \frac{\log n + \log(1+1/n)}{n(1+1/n)} - \frac{\log n}{n} \\
&= \left\{\frac{\log n}{n}\left(1 - \frac{1}{n} + O\left(\frac{1}{n^2}\right)\right) + \frac{1}{n}\left(\frac{1}{n} + O\left(\frac{1}{n^2}\right)\right)\left(1 + O\left(\frac{1}{n}\right)\right)\right\} - \frac{\log n}{n} \\
&= -\frac{\log n}{n^2} + \frac{1}{n^2} + O\left(\frac{\log n}{n^3}\right)
\end{aligned}$$

So now we deal with

$$
\begin{aligned}
&\sum_{2}^{N+1} \pi(n)\Delta\left(\frac{\log n}{n}\right)\delta n \\
=&\sum_{2}^{N+1}\left(\frac{n}{\log n}+\frac{n}{(\log n)^2}+O\left(\frac{n}{(\log n)^3}\right)\right)\left(\left(\frac{1}{n^2}\right)\left(1-\log n+O\left(\frac{\log n}{n}\right)\right)\right)\delta n \\
=&\sum_{2}^{N+1}\frac{1}{n\log n}\left(1+\frac{1}{\log n}+O\left(\frac{1}{(\log n)^2}\right)\right)\left(1-\log n+O\left(\frac{\log n}{n}\right)\right)\delta n \\
=&\sum_{2}^{N+1}\frac{1}{n\log n}\left(-\log n-1+1+O\left(\frac{1}{\log n}\right)\right)\delta n \\
=&\sum_{2}^{N+1}\left(-\frac{1}{n}+O\left(\frac{1}{n(\log n)^2}\right)\right)\delta n.
\end{aligned}
$$

Returning to our original sum, we have

$$
\begin{aligned}
&\left[\frac{\log n}{n}\pi(n-1)\right]_2^{N+1}-\sum_{2}^{N+1}\pi(n)\Delta\left(\frac{\log n}{n}\right)\delta n \\
=&1+o(1)-\sum_{2\le n\le N}\left(-\frac{1}{n}+O\left(\frac{1}{n(\log n)^2}\right)\right) \\
=&1+o(1)+\log N+c_1+o(1)+c_2+o(1)=\log N+c+o(1).
\end{aligned}
$$

□

C Proof that $T_n^{(k)}(x)=\lambda_k(x)x+\rho_k(x)$

We stated in Section 3.1 that

$$T_n^{(k)}(x)=\lambda_k(x)x+\rho_k(x)$$

for

$$\lambda_k(x) = \frac{p_{n+1}^{a(A_k,1)}}{\prod_{j\le n} p_j^{b(\beta_j,1)}}$$

and

$$\rho_k(x) = \sum_{i=1}^{k} \frac{c(\alpha_i) p_{n+1}^{a(A_k,i+1)}}{\prod_{j\le n} p_j^{b(\beta_j,i)}}$$

Proof. We proceed by induction on k. First, we show the case $k = 1$. If α_1 contains all nonzero entries, so that $a(A_1, 1) = 1$ and x is divisible by no prime $\le p_n$, the image of x is $T_n^{(1)}(x) = p_{n+1}x + 1$. Our formula gives

$$\lambda_1(x) = \frac{p_{n+1}^1}{1} = p_{n+1}$$

and

$$\rho_1(x) = \sum_{i=1}^{1} \frac{p_{n+1}^{a(A_k,i+1)}}{\prod_{j\le n} p_j^{b(\beta_i,j)}} = \frac{p_{n+1}^0}{\prod_{j\le n} p_j^0} = 1$$

Then we have $T^{(1)}(x) = \lambda_1(x)x + \rho_1(x)$ in this case. Now consider the case when α_1 has some entry equal to 0. λ_k has a numerator of 1 since $a(A_1, 1) = 0$ and a denominator equal to the product of all primes $\le p_n$ dividing x. ρ_k is 0 since $c(\alpha_1) = 0$. Then $T_n^{(1)}(x)$ is x divided by these primes, as it should be. So our base case holds.

So now assume that we have $T_n^{(k)}(x) = \lambda_k(x)x + \rho_k(x)$ for all k. When α_{k+1} has all nonzero entries, we should have $T_n^{(k+1)} = p_{n+1}T_n^{(k)} + 1$. We have $\lambda_{k+1} = p_{n+1}\lambda_k$ because $a(A_{k+1}, 1) - a(A_k, 1) = 1$ and the denominator remains unchanged. We also get $\rho_{k+1} = \sum_{i=1}^{k+1} \frac{c(\alpha_i)p_{n+1}^{a(A_k,i+1)}}{\prod_{j\le n} p_j^{b(\beta_i,j)}} = p_{n+1}\rho_k + 1$ from the definitions of a, b, and c. So $T_n^{(k+1)}(x) = \lambda_{k+1}(x)x + \rho_{k+1}(x) = p_{n+1}\lambda_k(x)x + p_{n+1}\rho_k(x) + 1 = p_{n+1}(\lambda_k(x)x + \rho_k(x)) + 1 = p_{n+1}T_n^{(k)} + 1$ as it should. Now consider the case when α_{k+1} has entries equal to 0. $\lambda_{k+1}(x)$ is just $\lambda_k(x)$ divided by the primes dividing the most recent iterate $T_n^{(k)}(x)$. Additionally, $\rho_{k+1}(x)$ is $\rho_k(x)$ divided by these primes, since $c_{k+1} = 0$ and our bottom changes by this factor. So we have $T_n^{(k+1)}(x) = T_n^{(k)}(x)$ divided out by small primes. This completes our proof. □

D The Number of Attainable Matrices for $n = 2$

As noted in section 3.2, we may estimate the number of attainable matrices of length k in the case $n = 2$ with the following theorem:

Theorem 14. *The number of attainable matrices of length k in the case $n = 2$ is asymptotic to m^k where m is the unique real solution to $x^5 - x^4 - 2x^3 - 2x^2 - x - 1 = 0$.*

Proof. We have noted that a matrix is admissible if and only if each α is a successor of its predecessor. Therefore, in estimating the number of attainable matrices, we need only consider the last α. We can then easily derive relations between the number of attainable matrices for k and $k + 1$. To begin, we consider all possible pairs of successors. In our case, these are

$$\begin{pmatrix}0\\0\end{pmatrix} \to \begin{pmatrix}0\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}$$
$$\begin{pmatrix}0\\1\end{pmatrix} \to \begin{pmatrix}0\\2\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}$$
$$\begin{pmatrix}0\\2\end{pmatrix} \to \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}$$
$$\begin{pmatrix}1\\0\end{pmatrix} \to \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}$$
$$\begin{pmatrix}1\\1\end{pmatrix} \to \begin{pmatrix}0\\0\end{pmatrix}$$
$$\begin{pmatrix}1\\2\end{pmatrix} \to \begin{pmatrix}0\\2\end{pmatrix}$$

If we call the number of $\binom{0}{0}$s ending attainable matrices of length k a_k, the number of $\binom{0}{1}$s ending them b_k, up to the number of $\binom{1}{2}$s ending them f_k, we may derive the following relations:

$$a_{k+1} = a_k + e_k$$
$$b_{k+1} = a_k + c_k$$
$$c_{k+1} = a_k + b_k + f_k$$
$$d_{k+1} = a_k + d_k$$
$$e_{k+1} = a_k + c_k + d_k$$
$$f_{k+1} = a_k + b_k + d_k$$

These Fibonacci-like recurrences guarantee that our sum $a_i + b_i + c_i + d_i + e_i + f_i$ will tend to grow at an exponential rate. Assuming the factor by which the sum multiplies each time is some m and that the $a_k \dots f_k$ tend toward some constant relative frequencies, we must solve the system

$$a_k + e_k = ma_k$$
$$a_k + c_k = mb_k$$
$$a_k + b_k + f_k = mc_k$$
$$a_k + d_k = md_k$$
$$a_k + c_k + d_k = me_k$$
$$a_k + b_k + d_k = mf_k$$

We have six equations in several variables, so the $a_k \dots f_k$ may all be scaled by some constant factor, but regardless of this factor we always get $m^5 - m^4 - 2m^3 - 2m^2 - m - 1 = 0$, so m is a constant ≈ 2.335. $m < 6$, so we expect allowed matrices to be increasingly less frequent (as a fraction of all matrices) for greater lengths k. □

E Small Cycles Among the First 100 p_{n+1}

Zhang and Yang note in their paper that there exist non-trivial cycles in the general case, in contrast to the apparent (but yet unproven) lack of these in the Collatz mapping [8]. A simple computer program can be used to make a search and find many of these cycles. Our C++ code for this search follows. Essentially, it takes a number and iterates several thousand times. We make the assumption that if a trajectory is going to cycle it will enter the cycle by this time. Then we store the next fifty or so iterated values and check if any are equal to each other, which would imply a cyclic trajectory. During the initial iterations, we also make some checks to ensure that the trajectory has not entered a previously known cycle (including the trivial cycle), and when this happens we move on to a new number. We test the first hundred thousand possible x for cycles. A few illustrative cycles are listed below. This list is of course far from exhaustive.

p_{n+1}	Cycle
11	$17 \to 47 \to 37 \to 17 \to \cdots$
13	$19 \to 31 \to 101 \to 73 \to 19 \to \cdots$
17	$43 \to 61 \to 173 \to 1471 \to 521 \to 4429 \to 4183 \to 2963 \to$
	$257 \to 437 \to 743 \to 1579 \to 2237 \to 3803 \to 2309 \to 19627 \to$
	$5561 \to 47269 \to 14881 \to 3833 \to 32581 \to 263 \to 43 \to \cdots$
59	$73 \to 359 \to 89 \to 101 \to 149 \to 157 \to 193 \to 73 \to \cdots$
61	$97 \to 269 \to 547 \to 97 \to \cdots$
61	$199 \to 607 \to 9257 \to 10457 \to 2593 \to 79087 \to 1206077 \to 199 \to \cdots$
61	$26833 \to 818407 \to 290249 \to 590173 \to 947383 \to 14447591 \to 26833 \to \cdots$
113	$1181 \to 1259 \to 5081 \to 41011 \to 1181 \to \cdots$

The actual C++ code used to detect these cycles follows. It was compiled by g++.

```
#include <iostream.h>

// Function prototypes

// Returns the value of x after iteration.
int iterate(int x);

//Checks whether or not a value is already within a known cycle.
bool isinloop(int x);

const int maxfound=10; // Maximum number of cycles we can hold for each p_n+1.
int found[maxfound];   // Matrix holds one value from each known cycle.
int foundind=0;        // Index for found[], counts number of cycles found.

// List of primes used to test various cases.
int primes[100]={2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37 , 41, 43,
                 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103,
                 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163,
                 167, 173, 179, 181, 191, 193, 197, 199, 211, 223, 227,
                 229, 233, 239, 241, 251, 257, 263, 269, 271, 277, 281,
                 283, 293, 307, 311, 313, 317, 331, 337, 347, 349, 353,
                 359, 367, 373, 379, 383, 389, 397, 401, 409, 419, 421,
                 431, 433, 439, 443, 449, 457, 461, 463, 467, 479, 487,
                 491, 499, 503, 509, 521, 523, 541};


int n; // Holds the number of the prime we are currently testing.

int main()
{

  // The number of iterations before we check for periodicity.
  // The maximum length of a cycle that the program will detect.
  // The value up to which we check for trajectories.
  const int beginits=10000, maxlength=50, maxcheck=100000;

  cout << "Cycle Searcher" << endl;


  // x is the current value of our number, loop stores values when we
  // look for a cycle.
  int x; int loop[maxlength];


  // We use a for loop over the first 100 cases of p_n.
  for (n=1; n<100; n++){
    foundind=0; // Set number of cycles found for this p_n to 0.
    cout << endl << primes[n] << ":" << endl;

  // Now we iterate over all x_0 (called a) up to our maximum value.
  for(int a=2; a<maxcheck; a++)
```

```
  {
    x=a; // Set x to its initial value a
    // We use another loop to calculate the first beginits iterations.
    for(int b=0; b<beginits; b++)
    {
      // When x is not already in a detected cycle, we iterate it.
      if (!isinloop(x))
x=iterate(x);
      else
        // When in a previously detected cycle, we stop calculating the
        // first beginits iterations and move on to a new number.
        b=beginits;
    }

    // If we didn't break out of the last loop when we entered a known
    // cycle, we look to see if our current trajectory is cyclic.
    if (!isinloop(x))
      // To do this, we take the first maxlength iterates and check to
      // see if any of them are equal, which would imply a cycle.
      for(int b=0; b<maxlength; b++)
      {
loop[b]=x; // Stores the current iterate in loop.
x=iterate(x); // Then calculates the next one.
// Then we check if our current iterate already showed up in
// this trajectory.
for(int c=0; c<=b; c++)
  // If we are in a cycle, there is no need to continue
  // calculating iterates.  We break out of the loop and store
  // our new value in found.
  if (x==loop[c])
  {
    b=maxlength;
    found[foundind]=x;
    foundind++;
  }
      }
  }
  if (foundind==0) // If we failed to find a cycle, we output this.
    cout << "No cycles found." << endl;
  for(int a=0; a<foundind; a++)  // Otherwise, we display cycles we found.
  {
    x=found[a];
    // We take one value from cycles that we have found
    // and display the first ten iterates.
    for(int b=0; b<10; b++)
    {
      cout << x << ", ";
      x=iterate(x);
    }
    cout << char(8) << char(8) << "..."; // Cleans up the output.
    cout << endl;
```

```
  }
  }
  return 0;
}

// Returns the value of x after iteration.
int iterate(int x)
{
  x=primes[n]*x+1; // We multiply by our p_n+1 and add one
  for(int i=0; i<n; i++)
    while (x%primes[i]==0) // and then when our number is divisible by
      x=x/primes[i];       // small primes we divide out by them.

  return x;
}

// Checks whether or not a number is part of a known cycle.
bool isinloop(int x)
{
  // Sometimes the int type is to small to hold our iterates and they
  // wrap to negative values.  When this happens, we pretend that we
  // found a cycle to make the program ignore this nonsense.
  if (x<=0)
    return true;

  // Then we check to see if our value is the same as any value in found.
  for(int i=0; i<foundind; i++)
    if (found[i]==x)
      return true;

  // If we didn't return a value after the last part then it's not part
  // of a known loop, so we return this.
  return false;
}
```

F Some Small Cycles in $3n + K$

It has long been known that when K is allowed to take on values other than 1 non-trivial cycles exist in the $3n + K$ cycle. We have briefly discussed this case and here give a list of some of these cycles. This list was generated by code analogous to that used above with a small number of simple changes. The essential algorithm remains the same. We note that this list does not contain all known cycles for each K; for example, $K = 15$ must contain at least as many cycles as $K = 5$.

k	Cycle
5	$1 \to 1 \to \cdots$
5	$5 \to 5 \to \cdots$
5	$19 \to 31 \to 49 \to 19 \to \cdots$
5	$23 \to 37 \to 29 \to 23 \to \cdots$
5	$187 \to 283 \to 427 \to 643 \to 967 \to 1453 \to 1091 \to 1639 \to 2461 \to$
	$1847 \to 2773 \to 2081 \to 781 \to 587 \to 883 \to 1327 \to 1993 \to 187 \to \cdots$
5	$347 \to 523 \to 787 \to 1183 \to 1777 \to 667 \to 1003 \to 1507 \to \cdots$
	$2263 \to 3397 \to 2549 \to 1913 \to 359 \to 541 \to 407 \to 613 \to 461 \to 347 \to \cdots$
15	$57 \to 93 \to 147 \to 57 \to \cdots$
21	$15 \to 33 \to 15 \to \cdots$
29	$61 \to 53 \to 47 \to 85 \to 71 \to 121 \to 49 \to 11 \to 31 \to 61 \to \cdots$
3275	$7 \to 309 \to 7 \to \cdots$