MATHEMATICS OF COMPUTATION, VOLUME 32, NUMBER 144
OCTOBER 1978, PAGES 1281–1292

On the "3x+1" Problem

By R. E. Crandall

Abstract. It is an open conjecture that for any positive odd integer m the function

$$C(m) = (3m + 1)/2^{e(m)},$$

where $e(m)$ is chosen so that $C(m)$ is again an odd integer, satisfies $C^h(m) = 1$ for some h. Here we show that the number of $m \leqslant x$ which satisfy the conjecture is at least x^c for a positive constant c. A connection between the validity of the conjecture and the diophantine equation $2^x - 3^y = p$ is established. It is shown that if the conjecture fails due to an occurrence $m = C^k(m)$, then k is greater than 17985. Finally, an analogous "$qx + r$" problem is settled for certain pairs $(q, r) \neq (3, 1)$.

1. Introduction. The "$3x + 1$" problem enjoys that appealing property, often attributed to celebrated number-theoretic questions, of being quite easy to state but difficult to answer. The problem can be expressed as follows: define, for odd positive integers x, the function:

$$C(x) = (3x + 1)/2^{e(x)}, \tag{1.1}$$

where $2^{e(x)}$ is the highest power of two dividing $3x + 1$. Since $C(x)$ is again an odd integer, one can iterate the C function any number of times. The problem: for any initial odd positive x is some iterate $C^k(x)$ equal to one?

This "$3x + 1$" problem has found a certain niche in modern mathematical folklore, without, however, having been extensively discussed in the literature. The function C defined in equation (1.1) is essentially Collatz' function [1] but the true origin of the problem seems obscure. The algorithm defined by successive iteration of C has been called the "Syracuse algorithm" [2]. Some authors [3], [8] have defined functions equivalent to C and proclaimed the conjecture, that $C^k(x) = 1$ for some $k(x)$, a long-standing one. Some partial results concerning the conjecture are known, but for the most part the behavior of the C function remains shrouded in mystery. It is hoped that the partial results contained in the next sections will shed some light on the problem.

2. Preliminary Observations. Let Z^+ denote the positive integers and let D^+ denote the odd elements of Z^+. Some elementary properties of the function C; $D^+ \longrightarrow D^+$ as defined in (1.1) will now be discussed.

DEFINITION. For $m \in D^+$ the *trajectory* of m is the sequence $T_m = \{C(m), C^2(m), \ldots\}$, where it is understood that the sequence terminates upon the first occurrence of $C^k(m) = 1$, $k \in Z^+$. If there is no such k, then T_m is an infinite sequence.

Received March 28, 1977; revised March 20, 1978.
AMS (MOS) subject classifications (1970). Primary 10A25.
Key words and phrases. Algorithm, diophantine equation.

DEFINITION. For $m \in D^+$ the *height* of m, denoted $h(m)$, is the cardinality of the trajectory T_m. In the case that T_m is a finite sequence, $h(m)$ will be the least number of iterations of C required to reach 1.

DEFINITION. For $m \in D^+$, we denote by inf T_m the least positive integer in the sequence T_m. Further, if T_m is bounded, we denote by sup T_m the greatest integer in the sequence T_m. If T_m is unbounded, we say that sup T_m is infinite.

The following table should serve as an example for the previous notation:

m	T_m	$h(m)$	sup T_m
1	$\{1\}$	1	1
7	$\{11, 17, 13, 5, 1\}$	5	17
27	$\{41, \ldots, 1\}$	41	3077
$2^{1000} - 1$	$\{?\}$	4316	$> 10^{476}$
$2^{1000} + 1$	$\{?\}$	2417	$< 10^{301}$
$2^{4096} - 1$	$\{?\}$	19794	?

It is partly the erratic behavior of the height function h that gives interest to the "$3x + 1$" problem. The main conjecture is:

CONJECTURE (2.1). *For every $m \in D^+$, $h(m)$ is finite.*

This unsolved conjecture has been verified for $m < 10^9$ [1], [5]. One of the few partial results concerning the problem is that of Everett [3]:

THEOREM (2.1) (EVERETT). *For almost all $m \in D^+$*, inf $T_m < m$.

This theorem is interesting because if inf $T_m < m$ for all $1 < m \in D^+$, then the main conjecture (2.1) is clearly true. It should be noted, however, that while Theorem (2.1) reveals a definite tendency for trajectories T_m to descend below their generating integers m, the theorem has little to say concerning Conjecture (2.1). In fact, it is not even known whether a positive density of odd integers m satisfy the conjecture.

The next observation gives further indication of computational difficulties encountered in the "$3x + 1$" problem.

THEOREM (2.2). *As $m \in D^+$ increases,* (sup T_m)/m *is unbounded.*

Proof. Define, for $k \in Z^+$, the number $m_k = 2^k - 1$. It is readily verified that $C(m_k) = 3 \cdot 2^{k-1} - 1$ if $k > 1$, and in general $C^j(m_k) = 3^j 2^{k-j} - 1$ if $k > j$. Thus, for $k > 1$ the number

$$C^{k-1}(m_k) = 3^{k-1}2 - 1$$

is a member of the trajectory of m_k. Therefore, for $k > 1$:

$$\frac{\sup T_{m_k}}{m_k} \geqslant \frac{3^{k-1}2 - 1}{2^k - 1} > (3/2)^{k-1}$$

and the right-hand side grows without bound as k runs through the positive integers.

It is evident that numerical calculations of the heights of various numbers by computer methods must necessarily involve the storage of trajectory elements which are, in proportion to the starting numbers m, arbitrarily large; unless, of course, some theoretical method is discovered to simplify such computations. It is natural to ask whether $(\sup T_m)/m^a$ is unbounded for various powers a. The set of integers used in Theorem (2.2) is sufficient to show unboundedness only for $a < \log_2 3$. The irrational number $t = \log_2 3 = \log 3/\log 2$ will arise in a natural way in later results.

3. A Random-Walk Argument. An heuristic argument that lends credibility to the main conjecture (2.1) can be stated as follows. Assume that for odd integers m sufficiently large the real number $\log(C(m)/m)$ is a "random variable" with a distribution determined by the behavior of the function $e(m)$ as defined in Eq. (1.1). One "expects" that $e(m) = k$ with probability 2^{-k}. Thus, $\log(C(m)/m)$ is approximately $\log 3 - k \log 2$ with probability 2^{-k}. Therefore, one "expects" the number $\log(C(m)/m)$ to be:

$$\sum_{k \in Z^+} 2^{-k} \log(3/2^k) = -\log(4/3),$$

indicating a tendency for $C(m)$ to be less than m. If one then imagines that iteration of the C function induces a random walk beginning at $\log m$ on the real number line, an heuristic estimate for the height function $h(m)$ might be:

$$h(m) \sim \frac{\log m}{\log(4/3)}. \tag{3.1}$$

What is notable about this argument, aside from its lack of precision, is that estimate (3.1) is not too far from the mark in a certain sense. Since we expect the h function to behave erratically, we define a smoother function called the average order of h:

$$H(x) = \frac{2}{x} \sum_{\substack{m \leqslant x \\ m \in D^+}} h(m). \tag{3.2}$$

The heuristic estimate (3.1) translates into an estimate for $H(x)$:

$$H(x) \sim 2(\log(16/9))^{-1} \log x = (3.476\ldots) \log x. \tag{3.3}$$

This simple statistical argument seems to be partially supported by the following data:

x	$H(x)/\log x$
11	1.440 . . .
101	2.546 . . .
1001	3.206 . . .
10001	3.330 . . .
100001	3.298 . . .

It would be of interest to compute $H(x)/\log x$ for some much larger value of x, say

10^{10}. In the absence of such knowledge, we simply conjecture:

CONJECTURE (3.1). $H(x) \sim 2 \log x/\log(16/9)$.

This conjecture is stronger than the main conjecture (2.1) in the sense that if there be even one $m \in D^+$ with infinite height, then (3.1) is false.

4. Uniqueness Theorem. We shall presently establish a certain uniqueness theorem which is useful in obtaining partial results concerning the main conjecture (2.1). Define, for $a \in Z^+$ and n rational the function:

$$B_a(n) = (2^a n - 1)/3. \tag{4.1}$$

In general, $B_a(n)$ is not an integer. Further define, for (backwards-ordered) sequences

$$\{a_i | i = k, k-1, \ldots, 2, 1; a_i, k \in Z^+\},$$

the functions

$$B_{a_k \cdots a_1}(n) = B_{a_k}(B_{a_{k-1} \cdots a_1}(n)) = (2^{a_k} B_{a_{k-1} \cdots a_1}(n) - 1)/3. \tag{4.2}$$

Every function $B_{\{a_i\}}(n)$ is thus rational by construction for any sequence $\{a_j, a_{j-1}, \ldots, a_2, a_1\} = \{a_i\}$.

LEMMA (4.1). *If $n \in D^+$ and $B_{a_j \dots a_1}(n)$ is an integer, then for $1 \leqslant i < j$ all numbers $B_{a_i \dots a_1}(n)$ are odd integers; and further*

$$C(B_{a_{i+1} \cdots a_1}(n)) = B_{a_i \cdots a_1}(n), \qquad C(B_{a_1}(n)) = n.$$

Proof. Assume for some k with $1 < k \leqslant j$ that $B_{a_k \dots a_1}(n) \in Z^+$. Then from (4.2) we conclude $2^{a_k} B_{a_{k-1} \dots a_1}(n)$ is also an integer. But by construction, $B_{a_{k-1} \dots a_1}(n) = y/3^{k-1}$ for some integer y since n is an integer. Since 2 and 3 are coprime it follows that $B_{a_{k-1} \dots a_1}(n)$ is an integer itself. Thus, as $B_{a_j \dots a_1}(n)$ is an integer it follows by induction that all $B_{a_i \dots a_1}(n)$ for $1 \leqslant i < j$ are integers. Further, from (4.2) we conclude that $3B_{a_i \dots a_1}(n) + 1$ is even so each $B_{a_i \dots a_1}(n)$ must be odd. Finally,

$$C(B_{a_{i+1} \cdots a_1}(n)) = B_{a_i \cdots a_1}(n) 2^{a_{i+1} - e} = B_{a_i \cdots a_1}(n)$$

from (4.2), (1.1), and the fact that application of the C function on an odd integer yields another odd integer. That $C(B_{a_1}(n)) = n$ follows from (4.1).

LEMMA (4.2). *If an integer $m = B_{a_j \dots a_1}(1)$ and $a_1 > 2$, then the trajectory of m is*

$$T_m = \{B_{a_{j-1} \cdots a_1}(1), B_{a_{j-2} \cdots a_1}(1), \ldots, B_{a_1}(1), 1\}.$$

Proof. From Lemma (4.1) the assumption $m = B_{a_j \dots a_1}(1)$ is an integer yields the correct iterations for the trajectory. It remains to show that none of the $B_{a_i \dots a_1}(1)$ is equal to 1. This follows from the fact that $C(1) = 1$, so that we only need show $B_{a_1}(1) \neq 1$. But this follows from the assumption $a_1 > 2$.

DEFINITION. Denote by G the set of finite sequences $\{a_j, a_{j-1}, \ldots, a_1\}$, where each $a_i \in Z^+$; $a_1 > 2$; and the following congruences are satisfied:

$$\begin{aligned} 2^{a_1} &\equiv 1 \pmod 3 \\ &\not\equiv 1 \pmod 9, \\ 2^{a_i}B_{a_{i-1}\cdots a_1}(1) &\equiv 1 \pmod 3 \\ &\not\equiv 1 \pmod 9 \quad \text{for } 2 \leqslant i \leqslant j-1, \\ 2^{a_j}B_{a_{j-1}\cdots a_1}(1) &\equiv 1 \pmod 3 . \end{aligned}$$

LEMMA (4.3). *Let* $\{a_i\} = \{a_j, a_{j-1}, \ldots, a_1\}$. *Then* $B_{\{a_i\}}(1)$ *is an integer of height* j *if and only if* $\{a_i\} \in G$.

Proof. Assume $B = B_{a_j\cdots a_1}(1)$ is an integer of height j. Then from Lemma (4.2) each $B_{a_i\cdots a_1}(1)$ is in the trajectory of B, and thus $B_{a_1}(1) \neq 1$, so $a_1 > 2$. The congruences (mod 3) follow from Eq. (4.2) and the congruences (mod 9) follow from (4.2) applied twice. Thus, the sequence $\{a_i | i = j, j-1, \ldots, 1\}$ is in G.

Now assume $\{a_i\}$ is in G. Then the last congruence implies $B_{a_j\cdots a_1}(1)$ is an integer, and Lemma (4.2) shows this integer has height j.

The previous lemmas and definition of the set of sequences G enable us to prove the uniqueness theorem:

THEOREM (4.1). *Consider the set of integers which satisfy the main conjecture* (2.1):

$$M = \{m \in D^+,\ m > 1 \,|\, h(m) \textit{ finite}\}.$$

Then, for each $m \in M$ *there is a unique sequence* $\{a_i\} \in G$ *such that*

$$m = B_{a_{h(m)}\cdots a_1}(1),$$

and conversely, for each $\{a_i\} \in G$, $B_{\{a_i\}}(1) \in M$.

Proof. If $m \in M$, then define the integers

$$a_i = e(C^{h(m)-i}(m)) \quad \text{for } 1 \leqslant i \leqslant h(m),$$

where the e function is as defined in (1.1). Then by inspection we have $m = B_{a_{h(m)}\cdots a_1}(1)$ and by Lemma (4.3) we have $\{a_i\} \in G$. To show uniqueness, assume for some $\{b_i | i = k, k-1, \ldots, 1\}$ we have $m = B_{b_k\cdots b_1}(1)$. Then from Lemma (4.2) $k = h(m)$ and, since the C function is well defined, the exponents b_i must agree termwise with the a_i so the sequences $\{a_i\}$ and $\{b_i\}$ are equal. For the converse, we note that $\{a_i\} \in G$ implies $B_{\{a_i\}}(1)$ has finite height by Lemma (4.3)

5. Numbers with Given Height. The set M of integers $m > 1$, $m \in D^+$ of finite height is naturally partitioned by the height function h. There are infinitely-many numbers m with $h(m) = 1$, in fact these m are just numbers of the form $(4^k - 1)/3$. It is natural to ask whether there are always numbers of any given height. This question can be answered in the affirmative. We shall give an estimate of the relative density of the odd integers with given height.

LEMMA (5.1). $B_{a_j\cdots a_1}(1) < 2^{a_1+a_2+\cdots+a_j}/3^j$.

Proof. Since $B_{a_1}(1) = (2^{a_1} - 1)/3$, the assertion is true for $j = 1$. From Eq. (4.2) the result follows by induction.

We shall also need a lower bound for the number of solutions to the congruences that define the set G:

LEMMA (5.2). *For a real number $z > 0$ the number of sequences of length j in the set G with $a_1 + \cdots + a_j \leqslant z$ is greater than or equal to* $(2[(z-2)/6j])^j$.

Proof. Solutions can be restricted by the inequalities

$$a_1 - 2 \leqslant (z-2)/j; \quad a_i \leqslant (z-2)/j \quad \text{for } 1 < i \leqslant j,$$

which together force the sum $a_2 + \cdots + a_j$ to be $\leqslant z - a_1$. The number of ways of choosing the quantity $a_1 - 2 \leqslant (z-2)/j$ is at least $2[(z-2)/6j]$; since only $2^{a_1} \equiv 4$ or $7 \pmod 9$ is required, and out of every six consecutive integers at least two must satisfy the congruence. Similarly, once a_1 is chosen there must be at least $2[(z-2)/6j]$ choices for a_2 since the congruence $2^{a_2}B_{a_1}(1) \equiv 4$ or $7 \pmod 9$ has at least two solutions a_2 out of every set of six consecutive integers. In this way the number of solutions to $a_1 + \cdots + a_j \leqslant z$ is seen to be at least the jth power of the quantity $2 \cdot [(z-2)/6j]$.

These lemmas can be used to estimate the number of $m \in D^+$, $m \leqslant x$ which have a given height h.

THEOREM (5.1). *Let $\pi_h(x)$ be the number of $m \in M$ with $h(m) = h$ and $m \leqslant x$. Then there exist real positive constants r, x_0 independent of h such that for* $x > \max(x_0, 2^{h/r})$,

$$\pi_h(x) > (\log_2^h(x^r))/h!.$$

Remark. This theorem shows that for any positive integer h there are infinitely many $m \in D^+$ with height h.

Proof. From the uniqueness theorem (4.1) and Lemma (5.1) it follows that $\pi_h(x)$ is at least as large as the number of sequences $\{a_i\} \in G$, of length h, with $a_1 + \cdots + a_h \leqslant \log_2(3^h x)$. By Lemma (5.2) this implies

$$\pi_h(x) \geqslant \left(\frac{\log_2(3^h x/4)}{3h} - 2\right)^h,$$

obtained by bounding the square-brackets from below. Now we choose x_0 such that $x > x_0$ implies $x/4 > x^{1/2}$; and choose a real number r, $0 < r < 1$, such that

$$r \log_2(3/64) + 1/2 > 3er > 0.$$

Then, noting from Stirling's formula that $h!e^h > h^h$, we can transform our last inequality for $\pi_h(x)$, in the case that x is greater than both x_0 and $2^{h/r}$, to

$$\pi_h(x) \geqslant h^{-h}\left(\frac{\log_2 x^{r\log_2(3/64)+1/2}}{3}\right)^h > \frac{\log_2^h(x^r)}{h!}.$$

6. An Estimate for $\pi(x)$. Let $\pi(x)$ be the number of integers $m \leqslant x$ which belong to the set M of Theorem (4.1). Conjecture (2.1) is equivalent to the statement that $\pi(x)$ is precisely the number of odd integers greater than one but not greater than x. We establish here a lower bound for the function $\pi(x)$.

LEMMA (6.1).

$$\lim_{t \to \infty} e^{-t} \sum_{\substack{u \in Z^+ \\ u \leq [t]}} \frac{t^u}{u!} = \frac{1}{2}.$$

Remark. This lemma can be proved using asymptotic expansions of certain error functions. An exposition of various formulas pertaining to the lemma can be found in [4].

Using the results of the last section, we now prove

THEOREM (6.1). *There exists a positive constant c such that for sufficiently large x*

$$\pi(x) > x^c.$$

Proof. Using Theorem (5.1), we can find $r > 0$ and $x_0 > 0$ such that $x > x_0$ implies

$$\pi(x) = \sum_{h \in Z^+} \pi_h(x) \geq \sum_{h=1}^{[r \log_2 x]} \frac{\log_2^h(x^r)}{h!}.$$

From Lemma (6.1) we deduce that for any $d > 0$ there is an $x_1 > x_0$ such that $x > x_1$ implies

$$\pi(x) > (1/2 - d)e^{r \log_2 x},$$

which is enough to obtain the lower bound x^c.

7. **Cycles.** Assume that all trajectories T_m for $m \in D^+$ are bounded, and assume that no $m > 1$ appears in its own trajectory. Then the main conjecture (2.1) is true, for under the two assumptions, the iterates $C(m), C^2(m), \ldots, C^{\sup T_m}(m)$ are distinct except for possible ones; and since each iterate is less than sup T_m, the number 1 must in fact appear in the list of iterates. It is not known whether either of the two assumptions is true. We shall show that numbers $m > 1$ which appear in their own trajectories, if they exist at all, are necessarily difficult to uncover. More precisely, we shall establish a lower bound for the period of an infinite cyclic trajectory in terms of its smallest member.

Let $m = B_{b_k \ldots b_1}(n)$ for some positive integer sequence $\{b_i\}$. Define

$$A_i = \sum_{j=k-i+1}^{k} b_j \quad \text{for } 1 \leq i \leq k; \tag{7.1}$$

$$A_0 = 0.$$

Then the B function can be expanded to give the identity:

$$2^{A_k} n - 3^k m = \sum_{j=0}^{k-1} 2^{A_j} 3^{k-1-j}. \tag{7.2}$$

The identity always follows from the statement that $n \in T_m$, for if n appears at the kth position of T_m then the assignment $b_j = e(C^{k-j}(m))$ gives $m = B_{b_k \ldots b_1}(n)$ and finally (7.2). Conversely, if (7.2) holds for some monotone sequence of increasing integers $0 = A_0 < A_1 < \cdots < A_k$, then the assignment $d_i = A_{k-i+1} - A_{k-1}$ gives $m = B_{d_k \ldots d_1}(n)$, so by Lemma (4.1) $C^k(m) = n$, implying either $n = 1$ or n appears in the kth position of T_m. In either case $n \in T_m$ and we have demonstrated

THEOREM (7.1). *Let $m, n \in D^+$. Then $n \in T_m$ if and only if there exists a sequence of integers*

$$0 = A_0 < A_1 < A_2 < \cdots < A_k$$

such that Eq. (7.2) *holds. Further, if such a sequence $\{A_i\}$ exists, then $n = 1$ or the kth element of T_m is n.*

COROLLARY (7.1). *If $1 < m \in D^+$ and $m \in T_m$ and the period of T_m is k, then there exists a sequence of integers $0 = A_0 < A_1 < \cdots < A_k$ such that*

$$m(2^{A_k} - 3^k) = \sum_{j=0}^{k-1} 2^{A_j} 3^{k-1-j}. \tag{7.3}$$

Conversely, if Eq. (7.3) *holds for such a monotone sequence, then $m \in T_m$ and either $m = 1$ or m is the kth element of T_m.*

It is of interest that if m appears in its own trajectory, then in Eq. (7.3) the factor $2^{A_k} - 3^k$ must divide the right-hand side. If Conjecture (2.1) is true, the diophantine equation (7.3) must have no solutions for $m > 1$, subject to the constraint of monotonicity on the $\{A_i\}$. It is known that the equation $2^x - 3^y = z$ has only finitely-many solutions in integers x, y for each value of z [9] and, in fact, can have at most one solution for sufficiently large z [10]. But sharp results for special prime values of z can be obtained on the assumption that the main conjecture (2.1) is true.

If Conjecture (2.1) is true, then the impossibility of solutions to the diophantine equation (7.3) can be shown to imply that if 2 is a primitive root of a prime p, any solution in integers x, y to $2^x - 3^y = p$ must have $y < p/(t-1)$, where $t = \log_2 3$.

It is of interest that if a pair of integers (a, b) can be found such that $b > 1$ and

$$0 < 2^a - 3^b \quad \text{divides} \quad 2^{a-b} - 1,$$

then the main conjecture (2.1) is false. Indeed, if we write

$$2^{a-b} - 1 = d(2^a - 3^b),$$

then the number $m = 2^b d - 1$ satisfies:

$$m(2^a - 3^b) = 3^b - 2^b = \sum_{j=0}^{b-1} 2^j 3^{b-j-1},$$

and since $m > 1$, Corollary (7.1) implies $h(m)$ is infinite.

Corollary (7.1) shows that if the main conjecture is true then powers of two and three tend to be poor approximations of each other. The number $t = \log_2 3$ must be accordingly difficult to approximate with rational numbers. This notion will be made precise shortly. For the moment we display the first 50 elements of the continued fraction for $t = \log_2 3$

$$
\begin{aligned}
(7.5)\qquad t &= [1, 1, 1, 2, 2, 3, 1, 5, 2, 23, 2, 2, 1, 1, 55, 1, 4, 3, 1, 1, 15,\\
&\qquad 1, 9, 2, 5, 7, 1, 1, 4, 8, 1, 11, 1, 20, 2, 1, 10, 1, 4, 1, 1, 1\\
&\qquad\qquad 1, 1, 37, 4, 55, 1, 1, 49, \ldots]\\
&= [a_0, a_1, a_2, \ldots, a_{49}].
\end{aligned}
$$

The convergents to this fraction are determined by recurrences [6], [7]

$$
(7.6)\qquad
\begin{aligned}
p_0 &= a_0; \quad p_{-1} = 1;\\
q_0 &= 1; \quad q_{-1} = 0;\\
p_n &= a_n p_{n-1} + p_{n-2} \quad \text{for } n \in Z^+;\\
q_n &= a_n q_{n-1} + q_{n-2} \quad \text{for } n \in Z^+.
\end{aligned}
$$

The ratio p_n/q_n is called the nth convergent to t.

The next three lemmas stem from the theory of rational approximation. Relevant material can be found in [6] and [7].

LEMMA (7.1). *Let p_n/q_n denote the nth convergent to $t = \log_2 3$. Then for any pair of integers (x, y) with $y < q_n$,*

$$|p_n - q_n t| < |x - yt|.$$

LEMMA (7.2). *For p_n/q_n convergents to t,*

$$|p_n - q_n t| > (q_n + q_{n+1})^{-1}.$$

LEMMA (7.3). *For p_n/q_n convergents to t, let $y < q_m$. Then*

$$|2^x - 3^y| > 3^y \log 2 |p_m - q_m t|.$$

Proof.

$$
\begin{aligned}
|2^x - 3^y| &= 3^y |\exp(x \log 2 - y \log 3) - 1|\\
&> 3^y \log 2 |x - yt|.
\end{aligned}
$$

When $y < q_m$, Lemma (7.1) *implies the desired inequality.*

We now focus our attention on numbers $m > 1$ for which $m \in T_m$. It is clear that every infinite cyclic trajectory contains such an m.

LEMMA (7.4). *If $1 < m = \inf T_m$, then for the A_j as defined in Eq.* (7.1),

$$2^{A_j} \leqslant (3 + 1/m)^j.$$

Proof. From $C(x) = (3x + 1)/2^{e(x)}$ we infer, since $m \leqslant C^j(m)$ for each j, that

$$C^{j+1}(m) \leqslant C^j(m)(3 + 1/m)/2^{e(C^j(m))}.$$

But this implies

$$m \leqslant C^j(m) \leqslant m(3 + 1/m)^j/2^{A_j}$$

giving the desired inequality.

LEMMA (7.5). *Let $1 < m = \inf T_m$ and let k be the period of the trajectory T_m. Then*

$$m < k(3 + 1/m)^{k-1}/(2^{A_k} - 3^k).$$

Proof. From Corollary (7.1) and Lemma (7.4) we have

$$m(2^{Ak} - 3^k) < \sum_{j=0}^{k-1} (3 + 1/m)^{k-1}$$

and the desired inequality follows from direct evaluation of the sum.

These lemmas enable us to establish a connection between the size of the period of a possible cyclic trajectory and the continued fraction for $t = \log_2 3$.

THEOREM (7.2). *Let p_n/q_n be convergents to t. Let $1 < m = \inf T_m$ and let k be the period of T_m. Then for $n \geqslant 4$:*

$$k > \min(q_n, 2m/(q_n + q_{n+1})).$$

Proof. If $k > q_n$, the result is trivial so assume $k \leqslant q_n$. Then from Lemmas (7.3) and (7.5) we have

$$m < \frac{k(3 + 1/m)^{k-1}}{3^k \log 2 |p_n - q_n t|},$$

from which we see that

$$k > m(\log 2)(3 + 1/m)|p_n - q_n t|(1 - k/3m).$$

But $n \geqslant 4$ implies $q_n + q_{n+1} > 20$ using Eq. (7.5). Thus, if $k \geqslant m/10$, the result of the theorem is trivially true. So we assume $k < m/10$. But this implies $(1 - k/3m) > 29/30$ and from Lemma (7.2),

$$k > \frac{(3m + 1) \log 2}{q_n + q_{n+1}} (29/30) > \frac{2m}{q_n + q_{n+1}}.$$

COROLLARY (7.2). *For any given k there are finitely many cyclic trajectories with period k.*

Proof. This corollary follows directly from the observation that $\min(q_n, (2 \inf T_m)/(q_n + q_{n+1}))$ is unbounded for any infinite set of trajectories as n increases.

The known result that no $1 < m < 10^9$ appears in its own trajectory can be used to establish a lower bound for any period k.

THEOREM (7.3). *Assume $m > 1$ and $C^k(m) = m$. Then $k > 17985$.*

Proof. For $1 < m = C^k(m)$ the trajectory T_m is infinite cyclic, and there is thus an $m_0 = \inf T_m$ with $m_0 > 10^9$. Since $m_0 = C^k(m_0)$, the number k is greater than or equal to the period of T_{m_0}. Therefore, from Theorem (7.2) we infer that

$$k > \min(q_n, 2 \cdot 10^9/(q_n + q_{n+1}))$$

for any $n \geqslant 4$. From Eq. (7.5) we compute:

$$q_{10} = 31867; \qquad q_{11} = 79335;$$

so either $k > 31867$ or $k > 2 \cdot 10^9/111202$. The latter bound is greater than 17985.

8. **The "$qx + r$" Problem.** Define for $q, r \in D^+$; $q > 1$ the function

$$C_{qr}(m) = (qm + r)/2^{e_{qr}(m)} \tag{8.1}$$

valid for $m \in D^+$ with $e_{qr}(m)$ always chosen to force $C_{qr}(m) \in D^+$. The generalized "$qx + r$" problem can be stated thus: given (q, r), for every $m \in D^+$ is there a k such that $C^k_{qr}(m) = 1$? What is remarkable about this problem is that strong evidence points to the following rather curious conjecture:

CONJECTURE (8.1). *In the "$qx + r$" problem, with $q, r \in D^+$ and $q > 1$, some $m \in D^+$ fails to satisfy an equation $C^k_{qr}(m) = 1$; except in the case $(q, r) = (3, 1)$.*

This conjecture says in a word that unless $q = 3$ and $r = 1$, the generalized height function $h_{qr}(m)$ will be infinite for some $m \in D^+$. We shall presently attempt to argue the plausibility of the conjecture.

It is easy to prove Conjecture (8.1) in the special case $r > 1$. Indeed for such an r choose any $m \equiv 0 \pmod r$. Then $C_{qr}(m) \cdot 2^{e_{qr}(m)} = qm + r \equiv 0 \pmod r$ and, as r is odd, $C_{qr}(m) \equiv 0 \pmod r$. But this means all iterates of m are destined to be divisible by r, hence $C^k_{qr}(m) = 1$ is impossible for $r > 1$ and $m \equiv 0 \pmod r$.

The status of Conjecture (8.1) is unsettled for most of the remaining "$qx + 1$" problems. It is known that the conjecture is true for $q = 5$, 181, and 1093; but all other cases remain elusive.

The case $q = 5$ is settled by the observation that $m = 13$ gives rise to a cyclic trajectory $\{33, 83, 13, \ldots\}$, whose occurrence can be traced back to the diophantine equation $2^7 - 5^3 = 3$; and an equation similar to (7.3) exists with 3 replaced by 5 and k set equal to 3.

The case $q = 181$ is resolved by the observation that $m = 27$ gives rise to a cyclic trajectory $\{611, 27, \ldots\}$; arising from the diophantine equation $2^{15} - 181^2 = 7$.

The third solvable case, $q = 1093$, is rather peculiar. In this problem, a number m has height one if and only if it is of the form

$$m = (2^{364p} - 1)/1093;$$

but, as it turns out, all other m have infinite height. This follows from the fact that in the "$1093x + 1$" problem, no m can have height 2, since Eq. (7.2), with 3 replaced by 1093, $n = 1$, and k set equal to two, is easily seen to be impossible. This case is the only one for which it is known that almost all m have infinite height.

The examples $q = 5$, 181, 1093 being the only known q for which Conjecture (8.1) is true (with $r = 1$), why is the conjecture plausible? The answer is, simply, it is likely that for any $q > 3$, some m has a diverging trajectory in the "$qx + 1$" problem. In fact, the heuristic arguments of Section 3 indicate that $\log(C_{q1}(m)/m)$ should be about $\log(q/4)$, which is positive for odd $q > 3$, so that numbers should tend to be "pushed upward" by application of the C_{q1} function.

In spite of the above considerations, it is unknown whether even a single m in a single "$qx + 1$" problem gives rise to an unbounded trajectory. The outstanding unsolved case is the "$7x + 1$" problem, for which there may be no infinite cyclic trajectories. What makes the "$7x + 1$" problem all the more interesting is the empirical observation that the number $m = 3$ gives rise to a trajectory reaching to 10^{2000} and beyond, with no apparent tendency to return.

Acknowledgements. The author wishes to acknowledge the hospitality of the administration of Reed College, at whose Computer Center the calculations for this paper were performed. The author also wishes to thank Joe Roberts and J. P. Buhler for their valuable suggestions concerning the manuscript.

2234 S. E. 20th Street
Portland, Oregon 97214

1. J. H. CONWAY, "Unpredictable iterations," *Proc.* 1972 *Number Theory Conference*, Univ. of Colorado, Boulder, Colorado, 1972, pp. 45–52.

2. I. N. HERSTEIN & I. KAPLANSKY, *Matters Mathematical*, 1974.

3. C. J. EVERETT, "Iteration of the number-theoretic function $f(2n) = n$, $f(2n + 1) = 3n + 2$," *Advances in Math.*, v. 25, 1977, pp.42–45.

4. M. ABRAMOWITZ & I. STEGUN (Editors), *Handbook of Mathematical Functions*, 9th printing, Dover, New York, 1965.

5. *Zentralblatt für Mathematik*, Band 233, p. 10041.

6. J. ROBERTS, *Elementary Number Theory– A Problem Oriented Approach*, M.I.T. Press, Cambridge, Mass., 1977.

7. Y. A. KHINCHIN, *Continued Fractions*, Univ. of Chicago, Chicago, Ill., 1964.

8. *Pi Mu Epsilon Journal*, v. 5, 1972, pp. 338, 463.

9. S. S. PILLAI, *J. Indian Math. Soc.*, v. 19, 1931, pp. 1–11.

10. A. HERSCHEFELD, *Bull. Amer. Math. Soc.*, v. 42, 1936, pp. 231–234.