PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 1, May 1981

ON THE COLLATZ $3n + 1$ ALGORITHM

LYNN E. GARNER

ABSTRACT. The number theoretic function $s(n) = \frac{1}{2}n$ if n is even, $s(n) = 3n + 1$ if n is odd, generates for each n a Collatz sequence $\{s^k(n)\}_{k=0}^{\infty}$, $s^0(n) = n$, $s^k(n) = s(s^{k-1}(n))$. It is shown that if a Collatz sequence enters a cycle other than the 4, 2, 1, 4, ... cycle, then the cycle must have many thousands of terms.

1. Introduction. The *Collatz* $3n + 1$ *algorithm* is defined as a function s: $N \to N$ on the set of positive integers by

$$s(n) = \begin{cases} n/2 & \text{if } n \text{ is even,} \\ 3n + 1 & \text{if } n \text{ is odd.} \end{cases}$$

Let $s^0(n) = n$ and $s^k(n) = s(s^{k-1}(n))$ for $k \in N$. The *Collatz sequence* for n is

$$C(n) = \{s^k(n)\}_{k=0}^{\infty}.$$

For example, $C(17) = \{17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, \ldots\}$.

The original problem of Collatz concerns the existence of cycles in Collatz sequences. It is conjectured that every Collatz sequence ends in the cycle 4, 2, 1, 4, ... So many people have worked on this problem in the nearly fifty years of its existence that it is almost part of mathematical folklore. Martin Gardner reports [2] that the conjecture has been verified for all $n < 60{,}000{,}000$; Riho Terras says that the conjecture has been verified for all $n < 2{,}000{,}000{,}000$.

The only published results to date [1], [3] are probabilistic in nature, and tend to strengthen belief in the conjecture.

This paper proves that there are no other "short" cycles: if a cycle exists which does not contain 1, then it has many thousands of terms.

2. Stopping time. Collatz's conjecture is equivalent to the conjecture that for each $n \in N$, $n > 1$, there exists $k \in N$ such that $s^k(n) < n$. The least $k \in N$ such that $s^k(n) < n$ is called the *stopping time* of n, which we will denote by $\sigma(n)$.

It is not hard to verify that

$\sigma(n) = 1$ if n is even,

$\sigma(n) = 3$ if $n \equiv 1 \pmod 4$,

$\sigma(n) = 6$ if $n \equiv 3 \pmod{16}$,

$\sigma(n) = 8$ if $n \equiv 11$ or $23 \pmod{32}$,

$\sigma(n) = 11$ if $n \equiv 7, 15$, or $59 \pmod{128}$,

$\sigma(n) = 13$ if $n \equiv 39, 79, 95, 123, 175, 199$, or $219 \pmod{256}$,

and so forth. Everett [1] proves that almost all $n \in N$ have finite stopping time, and Terras [3] gives a probability distribution function for stopping times. Most positive

Received by the editors April 25, 1980; presented to the Society, January 6, 1980.
1980 *Mathematics Subject Classification.* Primary 10L10; Secondary 10A40.

integers have small stopping times; the above list accounts for $237/256 \approx 93\%$ of them. However, stopping times can be arbitrarily large, for $\sigma(2^n - 1) > 2n$. Some interesting cases of larger stopping times are $\sigma(27) = 96$, $\sigma(703) = 132$, $\sigma(35{,}655) = 220$, $\sigma(270{,}271) = 311$, and $\sigma(1{,}027{,}341) = 347$. In a computation of stopping times of integers up to 1,065,000, the largest observed stopping time was 347.

3. A term formula. Let $C^k(n)$ consist of the first k terms of the Collatz sequence for n. Let m (which depends on n and k) be the number of odd terms in $C^k(n)$, and let d_i be the number of consecutive even terms immediately following the ith odd term. Let d_0 be the number of even terms preceding the first odd term. Then the next term in the Collatz sequence for n is

$$s^k(n) = \frac{3^m}{2^{k-m}} n + \sum_{i=1}^{m} \frac{3^{m-i}}{2^{d_i + \cdots + d_m}}. \tag{1}$$

Note that $k - m = d_0 + d_1 + \cdots + d_m$.

4. Coefficient stopping time. By the *coefficient of* $s^k(n)$ is meant the coefficient of n in (1), namely $3^m/2^{k-m}$. The *coefficient stopping time* of n is the least $k \in N$ such that the coefficient of $s^k(n)$ is less than 1, and is denoted by $\kappa(n)$. Thus $\kappa(n)$ is the least k such that $3^m < 2^{k-m}$.

It is clear that if $s^k(n) < n$, then the coefficient of $s^k(n)$ is less than 1; thus $\kappa(n) \leqslant \sigma(n)$ for all $n \in N$, $n > 1$. We conjecture that $\kappa(n) = \sigma(n)$, and have verified it for all $n \leqslant 1{,}150{,}000$.

For $m = 0$ or $m \in N$, let $p(m) = [m \log_2 3]$; then $2^{p(m)} \leqslant 3^m < 2^{1+p(m)}$. Thus if $\kappa(n) = k$, then $k - m = 1 + p(m)$, or $k = m + 1 + p(m)$. These, then, are the possible coefficent stopping times.

We can also identify those $n \in N$ with a given coefficient stopping time. Clearly, $\kappa(n) = 1$ if and only if n is even. If $k = m + 1 + p(m) > 1$ is a possible coefficient stopping time, let $d_0 = 0$ and $d_1, d_2, \ldots, d_m \in N$ be such that $d_1 + \cdots + d_i \leqslant p(i)$ for $i = 1, 2, \ldots, m-1$, and $d_1 + \cdots + d_m = p(m) + 1$. Let x and y be integers such that $3^m x + 2^{1+p(m)} y = 1$. Then $\kappa(n) = k$ if and only if

$$n \equiv -x\left(\sum_{i=1}^{m} 3^{m-i} 2^{d_0 + \cdots + d_{i-1}}\right) \qquad (\mathrm{mod}\ 2^{1+p(m)}).$$

We are able to prove that $\kappa(n) = \sigma(n)$ under certain bound conditions which arise from a study of the powers of 2 and the powers of 3.

5. Powers of 2 and 3. The behavior of a Collatz sequence is clearly related to the way in which the powers of 2 are distributed among the powers of 3. We were surprised to find that the powers of 2 appear to be bounded away from the powers of 3 by an amount which grows almost as rapidly as the power of 3.

To be specific, we choose $M \in N$, and let

$$b(M) = \max_{j \leqslant M} \{-\log_3(1 - 2^{p(j)} 3^{-j})\}$$

and

$$B(M) = \max_{j \leqslant M} \{-\log_3(2^{1+p(j)} 3^{-j} - 1)\}.$$

Then it follows that

$$3^m - 2^{p(m)} > 3^{m-b(M)} \tag{2}$$

and

$$2^{1+p(m)} - 3^m > 3^{m-B(M)} \tag{3}$$

for all $m < M$. The values of M at which $b(M)$ and $B(M)$ increase are given in Table 1, with the corresponding values of b and B. (These calculations were carried out on a Hewlett-Packard HP-19C programmable calculator.)

M	b(M)	B(M)	M	b(M)	B(M)	M	b(M)	B(M)
1	1	1	306		6.267	8286		6.9
2	2		359	6.23		8951		7.0
3		1.535	665	9.14		9616		7.1
5		2.665	971		6.31	10281		7.2
7	2.508		1636		6.35	10946		7.3
12	3.921		2301		6.39	11611		7.4
17		2.946	2966		6.44	12276		7.6
29		3.346	3631		6.49	12941		7.8
41		4.062	4296		6.53	13606		8.0
53	5.618		4961		6.59	14271		8.4
94		4.246	5626		6.65	14936		8.9
147		4.477	6291		6.71	15601		10.2
200		4.785	6956		6.77	16266	9.4	
253		5.253	7621		6.8	31867	9.9	

TABLE 1. Values at which $b(M)$ and $B(M)$ increase

6. The main theorem. Now we can prove that $\kappa(n) = \sigma(n)$ if the number of odd terms encountered in the Collatz sequence is not too great. The following theorem makes the bound condition precise.

THEOREM. *In the notation of* (1), *if* $\kappa(n) = k$ *and*

$$m < \min\left\{M, (n/2)3^{1-B(M)}(1 - 3^{-b(M)})^{-1}\right\},$$

then $\sigma(n) = \kappa(n)$.

PROOF. If $\kappa(n) = k$, then $3^m/2^{k-m} < 1$ and

$$\frac{3^{m-i}}{2^{d_i + \cdots + d_m}} < \frac{2^{d_0 + \cdots + d_{i-1}}}{3^i} < \frac{2^{p(i-1)}}{3^i} < \frac{1}{3}(1 - 3^{-b(M)})$$

by (2). Thus

$$s^k(n) < \frac{3^m}{2^{k-m}} n + \frac{m}{3}(1 - 3^{-b(M)}).$$

Now if $s^k(n) \geqslant n$, then $n < (3^m/2^{k-m})n + (m/3)(1 - 3^{-b(M)})$, and therefore

$$n < \frac{2}{3} m(1 - 3^{-b(M)}) \frac{2^{p(m)}}{3^m} \cdot 3^{B(M)}$$

by (3). This leads by the hypothesis to $n < n$, a contradiction. Hence $s^k(n) < n$, and $\sigma(n) \leqslant k$. But $\kappa(n) \leqslant \sigma(n)$, so $\sigma(n) = \kappa(n)$.

It is conjectured that the bound condition in the theorem is unnecessary, or is automatically satisfied so that $\sigma(n) = \kappa(n)$ in all cases.

7. Application to cycles. Suppose a Collatz sequence enters a cycle which does not contain 1. Let n be the least term of the cycle; then $s^k(n) = n$ for some $k \in N$. Hence $\kappa(n) \leqslant k$, so that if $\sigma(n) = \kappa(n)$, then n is not the least term of the cycle after all. Thus if stopping time and coefficient stopping time are always the same, then the only cycle is the 4, 2, 1, . . . cycle.

The same contradiction arises if the number of odd terms in the cycle satisfies the bound in the theorem. Thus if there is a cycle not containing 1, the number m of odd terms in the cycle must satisfy $m \geqslant \min\{M, (n/2)3^{1-B(M)}(1 - 3^{-b(M)})^{-1}\}$, where n is the least term of the cycle.

Using $n > 60{,}000{,}000$, $M = 14{,}000$, $B(M) = 8.0$, and $b(M) = 9.14$, we find $m \geqslant 13{,}700$, and hence $k \geqslant 35{,}400$. Thus any cycle not containing 1 must have at least 35,400 terms.

If the number 2 billion turns out to be the lower bound for n, as alleged, then $M = 41{,}000$, $B(M) = 10.2$, and $b(M) = 9.9$ yield $m \geqslant 40{,}700$ and $k \geqslant 105{,}000$.

In any event, there is only one "short" cycle, the known one.

References

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

2. M. Gardner, *Mathematical games*, Sci. Amer. **226** (1972), 115.

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

Department of Mathematics, Brigham Young University, Provo, Utah 84602