Publ. Math. Debrecen
45 / 1-2 (1994), 35–39

Cycles in Collatz sequences

By BUSISO P. CHISALA (Harare, Zimbabwe)

1. Introduction

The Collatz, or $3n+1$-problem has enjoyed a wide interest since its origin in the 1950's. It is this: starting with any positive integer a_1, generate the sequence (a_n) by the algorithm

$$a_{n+1} = \begin{cases} \dfrac{a_n}{2}, & \text{if } a_n \text{ is even;} \\ 3a_n + 1, & \text{if } a_n \text{ is odd.} \end{cases}$$

The classical formulation of the problem is: does every initial a_1 eventually arrive at the "cycle" 1, 4, 2, 1,…?

To date, this is apparently unsettled, and the most recent computer check has shown this "convergence" for all integers $a_1 < 2^{40}$ ([1], [5]). Our contribution will be to the issue of the possible existence of other cycles in the natural numbers. Previous work on this includes articles by GARNER [4], TERRAS [7] and CRANDALL [3]. We can define the same algorithm for negative integers, with the result that there are an additional 3 cycles. The general conjecture due to several authors is that there are only finitely many cycles. LAGARIAS' article [5] contains references to this and related conjectures.

2. Extensions

We will only consider the subsequence of *odd* terms, in other words, the iteration defined by the map on odd integers: $\mathcal{C}(n) = \dfrac{3n+1}{2^d}$, where d is the highest power of 2 dividing $3n+1$, or the 2-adic ordinal of this integer.

Now 1 is a fixed point of $\mathcal{C}$, and the problem is whether $\mathcal{C}^k(n) = 1$ for some k. Again, this map extends to the integers $\mathbb{Z}$, and then to the rationals $\mathbb{Q}$ in the form

$$\mathcal{C}(x) = (3x+1)|3x+1|_2,$$

with $|\,.\,|_2$ the normalised 2-adic norm for which $|2^k|_2 = \dfrac{1}{2^k}$. Clearly, we can even go up to the 2-adic completion $\mathbb{Q}_2$, but for our present purposes, we will take $\mathbb{Q}$ as the domain of $\mathcal{C}$. A cycle of length m is now a sequence x, $\mathcal{C}(x)$, $\mathcal{C}^2(x)$, $\ldots$, $\mathcal{C}^{m-1}(x)$, with $\mathcal{C}^m(x) = x$.

To each x, we associate the sequence of the ordinals $d_i = \mathrm{Ord}_2(3\mathcal{C}^{i-1}(x)+1)$ for $i = 1, 2, \cdots$. Then each $d_i \geq 1$, and if for every $m \geq 1$, we set $n = d_1 + \cdots + d_m$, we have

$$2^n\mathcal{C}^m(x) = 3^m x + \mathcal{G}_m(d_1, d_2, \ldots, d_{m-1}, d_m),$$

where the function $\mathcal{G}_m$ is given by

$$\mathcal{G}_m(d_1, \ldots, d_m) = \\ = 3^{m-1} + 3^{m-2}2^{d_1} + \cdots + 3 \cdot 2^{d_1+d_2+\cdots+d_{m-2}} + 2^{d_1+d_2+\cdots+d_{m-1}}.$$

It follows that x lies on an m-cycle if and only if

$$(2^{\Sigma d_i} - 3^m)x = \mathcal{G}_m(d_1, \ldots, d_m),$$

with the integers d_i determined by x. The surprising thing about extending to the rationals is the abundance of cycles. For any sequence of integers $(d_1, d_2, \ldots, d_m)$ with $d_i \geq 1$, the rational number $\dfrac{\mathcal{G}_m(d_1, \ldots, d_m)}{2^{\Sigma d_i} - 3^m}$ evidently belongs to an m-cycle, which is the unique one in $\mathbb{Q}$ associated to the sequence. For example, the fixed points of $\mathcal{C}$ are given by the sequences (d), $d \geq 1$, so that the sequences (1), (2) yield the only integral fixed points $x = -1, 1$.

3. Bounding Rational Cycles

Fixing $m \geq 1$, we would like to get an upper bound on the *least* member of any rational m-cycle. Since we wish to apply the results to the natural numbers, we will consider cycles with sequences $(d_1, \ldots, d_m)$ for which $2^n > 3^m$, where $n = \sum d_i$. The key lemma is the following, whose proof is due to Gary Nelson:

Lemma 3.1. *Let* $(d_1, \ldots, d_m)$ *be any sequence of real numbers. Given a sequence of weights* $(w_1, \ldots, w_m)$*, let* $A = \sum_{i=1}^m d_i w_i / \sum_{i=1}^m w_i$ *be the weighted average. Then up to a cyclic permutation, we can renumber the elements so that for* $1 \le k \le m$*, all the partial weighted averages* $\sum_{i=1}^k d_i w_i / \sum_{i=1}^k w_i$ *are bounded above by* A*.*

Proof. If all the d_i's are equal to A, we are done. Otherwise, let $d_{i_1}, d_{i_2}, \ldots, d_{i_s}$, be the elements d_j satisfying $d_j \le A$, but $d_{j-1} > A$ (take $j-1 = m$ if $j = 1$). This defines s blocks: the j^{th} block consisting of the elements from d_{i_j} up to but not including $d_{i_{j+1}}$.

Now make a new sequence, replacing each block by a single element whose value is the weighted average of the block, and whose weight is the sum of the weights of the members of the block. It is easy to check that if the lemma holds for the new (smaller) sequence, it holds for the old. We are done by induction. □

This proof gives an effective method for determining the "starting point" for the rearrangement. When all the weights are 1, this says that for a sequence $(d_1, \ldots, d_m)$, and average n, we can, after a cyclic permutation, assume that

$$d_1 \le A,\ d_1 + d_2 \le 2A, \ldots, d_1 + d_2 + \cdots + d_{m-1} \le (m-1)A.$$

It then follows that

$$\begin{aligned} \mathcal{G}_m(d_1, \ldots, d_m) &\le \mathcal{G}_m(A, \ldots, A) = \\ &= 3^{m-1} + 3^{m-2}2^A + \cdots + 3 \cdot 2^{(m-2)A} + 2^{(m-1)A} = \\ &= \frac{2^{mA} - 3^m}{2^A - 3}. \end{aligned}$$

Since $mA = n$, we conclude that in *any* m-cycle, setting $n = \sum d_i$ there is an element $\mathcal{C}^i(x) = \dfrac{\mathcal{G}(d_i, d_{i+1}, \ldots, d_{i-1})}{2^n - 3^m}$ with $\mathcal{C}^i(x) \le \dfrac{1}{2^{n/m} - 3}$. Since the denominator is minimised when $n = \lceil m \log_2 3 \rceil$, the least integer greater than $m \log_2 3$, we have

Proposition 3.2. *For any* m*-cycle of positive rationals, the least element is at most*

$$\frac{1}{2^{(\lceil m \log_2 3 \rceil / m)} - 3}.$$

In the next section, we say more about where this bound is attained.

4. Intermediate convergents

For an irrational number ξ, we denote its continued fraction by $[a_0; a_1, \ldots]$, so that $C_k = [a_0; a_1, \cdots, a_k]$ is its k^{th} convergent.

These may also be written as $\dfrac{p_k}{q_k}$ in reduced from, with the p_k, q_k given by the recurrence formulae in terms of the a_k's: $p_{k+2} = a_{k+2}p_{k+1} + p_k$, and $q_{k+2} = a_{k+2}q_{k+1} + q_k$. In particular, the denominators increase with k. The k^{th} convergent is the "best rational approximation" to ξ with denominator less than q_k, with the odd ones being particularly interesting here since they are all *greater* than ξ. Between C_k and C_{k+2} lie the so-called intermediate or quasi-convergents, which we shall write as C_k^i for $0 \le i \le a_{k+2}$, and define by:

$$p_k^i = ip_{k+1} + p_k, \; q_k^i = iq_{k+1} + q_k, \; C_k^i = \frac{p_k^i}{q_k^i}.$$

In particular, $C_k^0 = C_k$ and $C_k^{a_{k+2}} = C_{k+2} = C_{k+2}^0$. It is easily checked that if $i + 1 \le a_{k+2}$ and k is odd, then $C_k^i > C_k^{i+1} > \xi$. Thus the intermediate convergents form a strictly decreasing sequence lying above ξ, with successively larger denominators.

For any integer $m \ge 1$, the rational number $\dfrac{\lceil m\xi \rceil}{m}$ is greater than ξ, and is the best such approximation to ξ with denominator m. The following result, which appears in various (albeit disguised) forms in the literature (viz. [6] ch. 7, problem 5, and [2] ch. XXXII, §§15), extends the sense in which the odd convergents are closest to ξ:

Proposition 4.1. *The numbers C_k^i, p_k^i, q_k^i for odd k satisfy*

(1) $p_k^i = \lceil q_k^i \xi \rceil$, *so that* $C_k^i = \dfrac{\lceil q_k^i \xi \rceil}{q_k^i}$.

(2) $\dfrac{\lceil m\xi \rceil}{m} \ge C_k^i$ *for any m such that* $1 \le m < q_k^{i+1}$.

Of course, and analogous result holds for the lower intermediate convergents based on the even k's, with (1) replaced by $p_k^i = \lfloor q_k^i \xi \rfloor$, and inequality $\frac{\lfloor m\xi \rfloor}{m} \le C_k^i$ in the statement of (2).

5. Integral cycles

Suppose that it is known that no positive integers less than N lie on a cycle. Using these facts we are ready to prove our main result. Taking $\xi = \log_2 3$ and using the notation of section 4:

Theorem. *Suppose that $\mathcal{C}^t(n) = 1$ for all positive integers $n < N$, and let i and the odd integer $k \geq 1$ be defined by*

$$\frac{1}{2^{C_k^i} - 3} < N < \frac{1}{2^{C_k^{i+1}} - 3},$$

then there are no integral cycles with fewer than q_k^{i+1} terms.

Proof. For $m < q_k^{i+1}$, we have $\dfrac{\lceil m \log_2 3 \rceil}{m} \geq C_k^i$, by Prop. 4.1 (2). From proposition 3.2, the least element of any m-cycle is less than

$$\frac{1}{2^{(\lceil m \log_2 3 \rceil / m)} - 3} < \frac{1}{2^{C_k^i} - 3},$$

which is less than N. This least element is not an integer by assumption, so there are no integral m-cycles. $\square$

For instance, with the presently best known bound of $N = 2^{40}$, we find that $k = 13$, $i = 0$, with $a_{15} = 1$. So the upper bound on cycles is $q_{15} = 1.07813 \times 10^7$ — any integral cycle, with today's data on the Collatz problem, would have to have at least 10 million *odd* terms!

References

[1] Shiro Ando, Letter to J.C. Lagarias, Feb. 18, 1983, Reports that Prof. Nabuo Yoneda (Dept. of Information Science, Tokyo Univ.) has verified the $3x+1$ conjecture for all $n < 2^{40} \approx 1.2 \times 10^{12}$.

[2] G. Chrystal, Algebra, an elementary text-book Part II, *Chelsea*, 1952.

[3] R. E. Crandall, On the "$3x + 1$" problem, *Math. Comp.* **32** (1978), 1281–1292.

[4] L. E. Garner, On the Collatz $3n + 1$ algorithm, *Proc. A.M.S* **82** (1981), 19–22.

[5] J. C. Lagarias, The $3x + 1$ problem and its generalizations, *American Math. Monthly* **92** (1985), 3–23.

[6] I. Niven and H. Zuckermann, An introduction to the theory of numbers, *Wiley & Sons*, 1968.

[7] R. Terras, A stoppings time problem on positive integers, *Acta Arith.* **30** (1976), 241–252.

BUSISO P. CHISALA
MATHEMATICS DEPARTMENT
UNIVERSITY OF ZIMBABWE

(Received January 25, 1993)