Ivan Korec

A density estimate for the $3x + 1$ problem

Mathematica Slovaca, Vol. 44 (1994), No. 1, 85--89

Persistent URL: http://dml.cz/dmlcz/133225

Mathematica Slovaca
©1994
Mathematical Institute
Slovak Academy of Sciences

Math. Slovaca, **44** (1994), No. 1, 85–89

A DENSITY ESTIMATE FOR THE $3x+1$ PROBLEM

IVAN KOREC[1]

(*Communicated by Štefan Porubský*)

ABSTRACT. The set of those initial values y for which a value less than $y^{0.7925}$ is eventually reached after several steps of the algorithm from the $3x+1$ problem (called also Syracuse problem, Collatz-Kakutani problem, etc.) has asymptotic density 1.

Let $\mathbb{N}$ denote the set of nonnegative integers, and define for $y \in \mathbb{N}$

$$T(y) = \frac{3y+1}{2} \quad \text{if } y \text{ is odd,} \qquad T(y) = \frac{y}{2} \quad \text{if } y \text{ is even.}$$

Further, denote $T^0(y) = y$ and $T^{n+1}(y) = T\big(T^n(y)\big)$ for every $n, y \in \mathbb{N}$. By a well-known hypothesis, for every positive integer y there is n such that $T^n(y) = 1$; for references see e.g. [5]. This hypothesis is equivalent to the statement that for every positive integer y there is n such that $T^n(y) < y$. C. J. Everett [3] and R. Terras [6] proved that the asymptotic density of

$$\big\{y \in \mathbb{N} \mid \ (\exists\, n)\, \big(T^n(y) < y\big)\big\}\,; \tag{1}$$

is equal to 1. Remember that the asymptotic density of a set $M \subseteq \mathbb{N}$ is defined as $\lim\limits_{x\to\infty} \dfrac{\operatorname{card}\{y \in M \mid \ y < x\}}{x}$. In the present paper, there will be proved a similar result in which $T^n(y) < y$ will be replaced by a stronger inequality $T^n(y) < y^{0.7925}$. More precisely, it will be proved:

THEOREM 1. *For every real $c > \log_4 3$ $(= 0.79248125\dots)$ the set*

$$M_c = \big\{y \in \mathbb{N} \mid \ (\exists\, n)\, \big(T^n(y) < y^c\big)\big\} \tag{2}$$

has asymptotic density 1.

AMS Subject Classification (1991): Primary 11B83. Secondary 11B37.
Key words: Syracuse problem.
[1] This work was supported by Grant 363 of Slovak Academy of Sciences .

Notice that n in (2) is not bounded (as we shall see from the proof, it could be bounded by $\log_2 y$, but it could not be bounded independently of y). As a referee informed me, a similar result is contained also as a special case in [1], with however a larger bound $\frac{3}{2} - \log_3 2 = 0.86907\ldots$ for c. A similar result (more general, but without explicitly given constants) is obtained also in [4]. A strengthening of [3] is contained in [2], where the inequality $T^i(y) < y$ is requested for k consecutive values of i. A further related result was obtained in R. Terras [6], [7]: let $D(k)$ denote the asymptotic density of the set $\{y \in \mathbb{N} \mid (\exists n \leq k)(T^n(y) < y)\}$; then $\lim_{k\to\infty} D(k) = 1$.

In the proof of Theorem 1, we shall need some notation and results from [6]. For $k, m, y \in \mathbb{N}$ and a real d we define:

$$\begin{aligned}
X_k(y) &= \begin{cases} 1 & \text{if } T^k(y) \text{ is odd,} \\ 0 & \text{if } T^k(y) \text{ is even,} \end{cases} \\
E_k(y) &= \big(X_0(y), X_1(y), \ldots, X_{k-1}(y)\big)\,, \\
S_k(y) &= X_0(y) + X_1(y) + \cdots + X_{k-1}(y)\,, \\
U(m,d) &= \operatorname{card}\{y \in \mathbb{N} \mid \ 0 \leq y < 2^m \text{ and } S_m(y) \leq md\}\,.
\end{aligned}$$

Lemma 1. *For every $x, y, m \in \mathbb{N}$*

$$E_m(x) = E_m(y) \quad \textit{if and only if} \quad x \equiv y \pmod{2^m}\,.$$

This is the Periodicity theorem 2.1 from [6] (contained also in [3] in a more general form). It shows that $y \mapsto E_m(y)$ is a bijection between any set of 2^m consecutive nonnegative integers and the set $\{0,1\}^m$. Further, it implies

$$\operatorname{card}\{y \in \mathbb{N} \mid \ b \leq y < b + 2^m \text{ and } S_m(y) \leq md\} = U(m,d) = \sum_{k=0}^{\lfloor md \rfloor} \binom{m}{k};$$

in particular, this cardinality does not depend on b. We shall also need the following easy consequence of the central limit theorem:

Lemma 2. *For any real $d > \dfrac{1}{2}$ there holds* $\displaystyle\lim_{m\to\infty} \frac{U(m,d)}{2^m} = 1$.

Proof of Theorem 1. Let $\varepsilon > 0$ and $c > \log_4 3$ be given. Without loss of generality, $\varepsilon < 1$ and $c < 1$ can be assumed.

Since obviously $\operatorname{card}\{y \in M_c \mid y < a\} \le a$, it suffices to prove

$$\operatorname{card}\{y \in M_c \mid y < a\} \ge (1-\varepsilon) \cdot a$$

for every sufficiently large a (i.e. for every integer $a \ge a_0$, where $a_0 = a_0(c, \varepsilon)$ will be fixed later). Consider such a and find the least positive integer m such that $a \le m^2 \cdot 2^m$. Now let us consider arbitrary y satisfying

$$m \cdot 2^m \le y < a \tag{3}$$

and let us look for a simple condition which implies $y \in M_c$.

Set $d = \frac{1}{2} \cdot \left(\frac{c}{\log_4 3} + \frac{1}{2}\right)$. Since $c > \log_4 3$, we have $\frac{1}{2} < d < \frac{c}{\log_2 3}$.

Claim. *There is $n_1 = n_1(c)$ such that, if $m \ge n_1$, then the inequalities* (3) *and $S_m(y) < md$ imply $y \in M_c$.*

Clearly, it suffices to prove $T^m(y) < y^c$. Let us denote $k = S_m(y)$ the number of ones in the sequence $E_m(y)$, i.e. the number of odd integers among

$$T^0(y), T^1(y), \ldots, T^{m-1}(y).$$

Since $T^p(y) \ge \frac{y}{2^p} > \frac{m.2^m}{2^m} = m$ for every $p < m$, we have

$$\begin{aligned} T^m(y) &= y \cdot \frac{T^1(y)}{T^0(y)} \cdot \frac{T^2(y)}{T^1(y)} \cdot \ldots \cdot \frac{T^m(y)}{T^{m-1}(y)} < y \cdot \left(\frac{3m+1}{2m}\right)^k \cdot \left(\frac{1}{2}\right)^{m-k} \\ &= y \cdot \frac{1}{2^m} \cdot 3^k \cdot \left(1 + \frac{1}{3m}\right)^k \le y \cdot \frac{3^k}{2^m} \cdot \left(1 + \frac{1}{3m}\right)^m < y \cdot \frac{3^k}{2^{m-1}}. \end{aligned}$$

Therefore $y \in M_c$ whenever $y \cdot \frac{3^k}{2^{m-1}} \le y^c$, i.e. $y^{1-c} \cdot 3^k \le 2^{m-1}$, and this holds whenever

$$(m^2 \cdot 2^m)^{1-c} \cdot 3^k \le 2^{m-1}.$$

The last inequality is equivalent to

$$\frac{k}{m} \le \frac{c}{\log_2 3} - \frac{1 + 2(1-c)\log_2 m}{m}. \tag{4}$$

Since $\lim_{m\to\infty} \dfrac{1+2(1-c)\log_2 m}{m} = 0$, there is $n_1 = n_1(c)$ such that for every $m \geq n_1$ the inequality $\dfrac{k}{m} < d$ implies (4), and hence also $T^m(y) < y^c$. So the claim is proved.

Let us divide the integers (3) into $L(a)$ pairwise disjoint sets each of which consists of 2^m consecutive integers (and, maybe, one smaller set). Since $a > (m-1)^2 \cdot 2^{m-1}$, the number of such sets is

$$L(a) = \left\lfloor \frac{a - m \cdot 2^m}{2^m} \right\rfloor \geq \frac{a}{2^m} - 1 - m = \left(1 - \frac{(m+1)\cdot 2^m}{a}\right) \cdot \frac{a}{2^m}$$
$$\geq \left(1 - \frac{(m+1)\cdot 2^m}{(m-1)^2 \cdot 2^{m-1}}\right) \cdot \frac{a}{2^m} = \left(1 - \frac{2m+2}{(m-1)^2}\right) \cdot \frac{a}{2^m} > \left(1 - \frac{\varepsilon}{2}\right) \cdot \frac{a}{2^m}$$

whenever $m \geq n_2$ for some $n_2 = n_2(\varepsilon)$.

We have $d > \dfrac{1}{2}$, and therefore, by Lemma 2, there is $n_3 = n_3(c, \varepsilon)$ such that for all $m \geq n_3$

$$U(m,d) \geq \left(1 - \frac{\varepsilon}{2}\right) \cdot 2^m .$$

Now we are able to choose a_0: let $n = \max(n_1, n_2, n_3)$ and $a_0 = n^2 \cdot 2^n$. For arbitrary integer $a \geq a_0$ we have $m \geq n$, and hence

$$\operatorname{card}\{y \in M_c \mid y < a\} \geq L(a)\cdot U(m,d) \geq \left(1-\frac{\varepsilon}{2}\right)\cdot\frac{a}{2^m}\cdot\left(1-\frac{\varepsilon}{2}\right)\cdot 2^m > (1-\varepsilon)\cdot a ,$$

which completes the proof. □

The following examples show that Theorem 1 cannot be immediately derived from $\lim_{k\to\infty} D(k) = 1$, and that diminishing the bound for c in Theorem 1 could be nontrivial.

E x a m p l e 1. Let the function $t\colon \mathbb{N} \to \mathbb{N}$ be defined by

$$t(y) = \begin{cases} y & \text{if } y \text{ is a square,} \\ y-1 & \text{otherwise.} \end{cases}$$

Let the iterations t^i of the function t be defined in the usual way (i.e. like T^i above). Then for every $k \geq 1$ the set

$$\{y \in \mathbb{N} \mid (\exists n \leq k)\,(t^n(y) < y)\}$$

has asymptotic density 1. However, for every $c < 1$ the set

$$\{y \in \mathbb{N} \mid (\exists n)\,(t^n(y) < y^c)\}$$

is finite, and hence its asymptotic density is 0.

Example 2. Let $0 < d < 1$, and let the function $t\colon \mathbb{N} \to \mathbb{N}$ be defined by

$$t(y) = \begin{cases} y & \text{if } y = 0 \text{ or } y \text{ is a power of } 2, \\ 2^{\lfloor d \cdot \log_2 y \rfloor} & \text{otherwise.} \end{cases}$$

Then the set $\{y \in \mathbb{N} \mid (\exists n)\,(t^n(y) < y^c)\}$ has asymptotic density 1 if $c > d$ and asymptotic density 0 if $c < d$.

REFERENCES

[1] ALLOUCHE, J.-P.: *Sur la conjecture de 'Syracuse-Kakutani-Collatz'*. In: Seminar de Théorie des Nombres. Exp. No. 9, CNRS Talence, 1978–1979.

[2] DOLAN, J. M.—GILMAN, A. F.—MANICKAM, S.: *A generalization of Everett's result on the Collatz $3x+1$ problem*, Adv. in Appl. Math. **8** (1987), 405–409.

[3] EVERETT, C. J.: *Iteration of the number theoretic function $f(2n)=n$, $f(2n+1)=3n+2$*, Adv. Math. **25** (1977), 42–45.

[4] HEPPNER, E.: *Eine Bemerkung zum Hasse-Syracuse-Algorithmus*, Arch. Math. (Basel) **31** (1978), 317–320.

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

[6] TERRAS, R.: *A stopping time on the positive integers*, Acta Arith. **XXX** (1976), 241–252.

[7] TERRAS, R.: *On the existence of a density*, Acta Arith. **XXXV** (1979), 101–102.

Received April 30, 1992

Revised January 27, 1993

Mathematical Institute
Slovak Academy of Sciences
Štefánikova 49
SK-814 73 Bratislava
Slovakia