Statistics and Probability Letters 109 (2016) 39–44

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A randomized version of the Collatz $3x + 1$ problem

Aristides V. Doumas *, Vassilis G. Papanicolaou

Department of Mathematics, National Technical University of Athens, Zografou Campus, 157 80 Athens, Greece

ARTICLE INFO

Article history:
Received 1 September 2015
Received in revised form 21 October 2015
Accepted 30 October 2015
Available online 11 November 2015

MSC:
60J10

Keywords:
Collatz $3x + 1$ problem
Markov chain

ABSTRACT

We consider a Markov chain on the positive odd integers, which can be viewed as a stochastic version of the Collatz $3x+1$ Problem. We show that, no matter its initial value, the chain visits 1 infinitely often. Its values, however, are unbounded.

1. The Collatz $3x + 1$ problem

The classical Collatz $3x + 1$ Problem can be formulated as follows. Let x be a positive odd integer. Consider the sequence

$$x_0 = x, \qquad x_n = \frac{3x_{n-1} + 1}{2^{d_n}}, \quad n \geq 1, \tag{1.1}$$

where 2^{d_n} is the highest power of 2 dividing $3x_{n-1} + 1$. Hence $\{x_n\}_{n=0}^{\infty}$ is a sequence of positive odd integers (if, e.g., $x = 1$, then $x_n = 1$ for all n). Notice that $d_n \geq 1$ for all $n \geq 1$.

Suppose $L := \liminf x_n < \infty$. Then, since x_n takes only positive integral values, we must have that $x_n = L$ for infinitely many values of n. In particular, $x_k = x_{k+b} = L$ for some $k \geq 0$, $b \geq 1$. But, then, it follows from (1.1) that $x_{k+n} = x_{k+b+n}$ for all integers $n \geq 0$. Therefore, either

$$\lim_n x_n = \infty, \tag{1.2}$$

or the sequence $\{x_n\}_{n=0}^{\infty}$ is eventually periodic, namely there is a $b \geq 1$ and an $n_0 \geq 0$ such that

$$x_{n+b} = x_n \quad \text{for all } n \geq n_0. \tag{1.3}$$

Notice that, if $b = 1$, i.e. if there is a n_0 such that $x_{n+1} = x_n$ for all $n \geq n_0$, then (1.1) implies that $(2^{d_{n+1}} - 3)x_n = 1$, which forces $x_n = 1$ for all $n \geq n_0$.

The Collatz $3x + 1$ Problem pertains to the behavior of the sequence $\{x_n\}_{n=0}^{\infty}$ as $n \to \infty$. One famous and longstanding open question is whether there exists some initial value x for which $\lim_n x_n = \infty$, while another open question is whether it is possible to have an eventually periodic behavior with a (minimal) period $b > 1$.

The ultimate Collatz Conjecture is that, no matter what the initial value x is, we always have that $x_n = 1$ for all n sufficiently large. Needless to say that the conjecture has been verified for a huge set of initial values x. More information

* Corresponding author.
E-mail addresses: aris.doumas@hotmail.com (A.V. Doumas), papanico@math.ntua.gr (V.G. Papanicolaou).

http://dx.doi.org/10.1016/j.spl.2015.10.017

about the Collatz problem and its ramifications can be found in Applegate and Lagarias (2003), Applegate and Lagarias (2006), Borovkov and Pfeifer (2000), Crandall (1978), Franco and Pomerance (1995), Lagarias (1985), Lagarias and Weiss (1992), Volkov (2006) and the references therein.

2. A randomized version of the problem

Let x be a positive odd integer. We consider the sequence

$$X_0 = x, \qquad X_n = \frac{3X_{n-1} + \xi_n}{2^{d_n}}, \quad n \geq 1, \tag{2.1}$$

where $\{\xi_n\}_{n=1}^{\infty}$ is a sequence of independent and identically distributed (i.i.d.) random variables taking odd integral values ≥ -1, and 2^{d_n} is the highest power of 2 dividing $3X_{n-1} + \xi_n$ (notice that we again have $d_n \geq 1$ for all $n \geq 1$). Thus, $\{X_n\}_{n=0}^{\infty}$ is now a random sequence of positive odd integers.

Let us introduce the filtration

$$\mathcal{F}_0 := \{\emptyset, \Omega\}, \qquad \mathcal{F}_n := \sigma(\xi_1, \ldots, \xi_n), \quad n \geq 1, \tag{2.2}$$

where $(\Omega, \mathcal{F}, P)$ is the underlying probability space. Clearly, by (2.1) we have that the random variables X_n and d_n are $\mathcal{F}_n$-measurable for all $n \geq 1$. Notice that

$$\mathcal{F}_n^X := \sigma(X_1, \ldots, X_n) \subset \mathcal{F}_n \quad \text{and} \quad \mathcal{F}_n^d := \sigma(d_1, \ldots, d_n) \subset \mathcal{F}_n, \quad n \geq 1. \tag{2.3}$$

Of course,

$$\mathcal{F}_n^X \vee \mathcal{F}_n^d = \mathcal{F}_n, \quad n \geq 1, \tag{2.4}$$

where $\mathcal{F}_n^X \vee \mathcal{F}_n^d$ denotes the σ-algebra generated by $\mathcal{F}_n^X$ and $\mathcal{F}_n^d$.

Formula (2.1) implies that $\{X_n\}_{n=0}^{\infty}$ is a Markov chain with respect to $\mathcal{F}_n$, whose state space is the set $\mathbb{N}_{\text{odd}}$ of positive odd integers. Actually, the two-dimensional process $\{(X_n, d_n)\}_{n=0}^{\infty}$ can be also viewed as a Markov chain with respect to $\mathcal{F}_n$ (the value of d_0 is irrelevant; furthermore, conditioning on d_n is irrelevant for (X_{n+1}, d_{n+1})).

The most natural case to examine first seems to be the choice $P\{\xi_n = -1\} = P\{\xi_n = 1\} = 1/2$ (or $P\{\xi_n = 1\} = P\{\xi_n = 3\} = 1/2$). Here, however, we will consider the rather easier case

$$P\{\xi_n = 1\} = P\{\xi_n = 3\} = P\{\xi_n = 5\} = P\{\xi_n = 7\} = \frac{1}{4}. \tag{2.5}$$

To begin our analysis, let us observe that for any positive odd integral value of X_{n-1} we have

$$\{3X_{n-1} + 1,\ 3X_{n-1} + 3,\ 3X_{n-1} + 5,\ 3X_{n-1} + 7\} \equiv \{0, 2, 4, 6\} \mod 8. \tag{2.6}$$

Therefore, due to (2.5) and the independence of X_{n-1} and ξ_n we have

$$P\{3X_{n-1} + \xi_n \equiv k \mod 8\} = \frac{1}{4} \quad \text{for } k = 0, 2, 4, 6. \tag{2.7}$$

The above formula motivates us to set

$$m_n := 3 \wedge d_n, \quad n \geq 1 \tag{2.8}$$

(as usual, $a \wedge b$ denotes the minimum of a and b), where d_n is the random exponent appearing in (2.1). Then, (2.1), (2.7), the Markov property of (X_n, d_n), and the independence of X_{n-1} and ξ_n imply

$$P\{m_n = 1 \mid \mathcal{F}_{n-1}\} = P\{m_n = 1 \mid X_{n-1}\} = P\{3X_{n-1} + \xi_n \equiv 2 \mod 4 \mid X_{n-1}\} = \frac{1}{2} \tag{2.9}$$

for all $n \geq 1$. Likewise,

$$P\{m_n = 2 \mid \mathcal{F}_{n-1}\} = P\{m_n = 3 \mid \mathcal{F}_{n-1}\} = \frac{1}{4}, \quad n \geq 1. \tag{2.10}$$

Formulas (2.9) and (2.10) tell us that m_n and $\mathcal{F}_{n-1}$ are independent for every $n \geq 1$. In particular (since m_n is $\mathcal{F}_n$-measurable for all $n \geq 1$) we have that $\{m_n\}_{n=1}^{\infty}$ is a sequence of i.i.d. random variables with

$$P\{m_n = 1\} = \frac{1}{2}, \qquad P\{m_n = 2\} = P\{m_n = 3\} = \frac{1}{4}. \tag{2.11}$$