Why do Collatz series converge?

Kanber, B.
University of Leicester, Leicester, UK. Telephone: +44 116 258 5140, Fax: +44 116 258 60 70, e-mail: bk20@leicester.ac.uk.

Abstract

The aim of this study was to investigate why Collatz series, starting with positive whole numbers, always seem to converge (i.e. end in the 4,2,1 cycle), regardless of their initial behaviour. The results presented here appear to show that symmetrical, Gaussian-like distributions may be inherently involved in the generation of these sequences. The Collatz process appears to draw at random (although not proven or tested in this preliminary report) from this distribution, presumably resulting in its convergent behaviour. It is noted that this is not an attempt or claim to prove or disprove the Collatz conjecture, but only a collection of the results obtained in the aforementioned investigation.

Introduction

Collatz sequences are constructed by taking some arbitrary positive whole number and iterating it through the Collatz function. The latter divides a given number by 2 if it is even, multiplies it by 3 and adds 1 if it is odd:

$$C_i = C_{i-1}/2 \text{ if } C_{i-1} \text{ is even}$$

$$C_i = 3C_{i-1}+1 \text{ if } C_{i-1} \text{ is odd}$$

An alternative, but equivalent form of the Collatz function is one which includes a division by 2 for odd, as well as even numbers:

$$C_i = C_{i-1}/2 \text{ if } C_{i-1} \text{ is even}$$

$$C_i = (3C_{i-1}+1)/2 \text{ if } C_{i-1} \text{ is odd}$$

The Collatz conjecture is that all such sequences eventually reach unity, or more correctly result in the 4,2,1 cycle [1-7]. Some starting numbers exhibit partially divergent trajectories, for example, the starting number 27, only to eventually converge back towards unity.

Methods

For the purposes of this investigation, we write the Collatz process as follows:

Given an odd starting number:

$$2^n q-1 \text{ (q odd)}$$

(1) Carry out the operation $(3x+1)/2$, n times to obtain the even number:

$$2^m r \text{ (r odd)}$$

(2) Carry out the operation $x/2$, m times to obtain the odd number r.

(3) Repeat the whole process for the new starting number, until unity is reached.

Alternatively, given an even starting number:

$$2^m r \text{ (r odd)}$$

(1) Carry out the operation x/2, m times to obtain the odd number:

$$2^n q-1 \text{ (q odd)}$$

(2) Carry out the operation (3x+1)/2, n times to obtain the even number:

$$3^n q-1 \text{ (q odd)}$$

(3) Repeat the whole process for the new starting number, until the number 2 is reached.

This investigation studies the quantity m-n.

Results

Figures 1 and 2 show the histograms of all initial m-n values for all starting numbers from 1 to 2^z-1 for z=10 and z=18. In both cases, the distributions are approximately symmetrical, and Gaussian-like with an approximately zero mean. In the case of Figure 2, the distribution spans a greater range of m-n values in accordance with the larger range of starting numbers studied.

Figures 3 and 4 show the sequence of m-n values encountered for the complete Collatz sequences starting at an arbitrary, large, odd number, and an arbitrary, large, even number, and the corresponding histograms. The m-n values appear to be randomly distributed (although not proven or tested in this preliminary report) around zero, while the histograms also appear approximately symmetrical, and Gaussian-like, with approximately zero means.

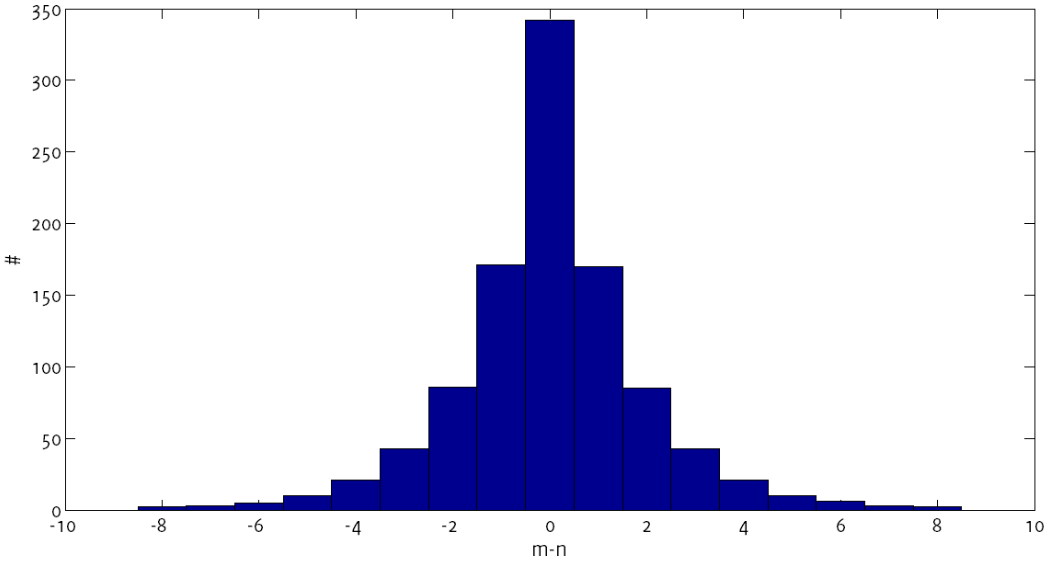


Figure 1 - The histogram of all initial m-n values for all starting numbers from 1 to 2^{10}-1.

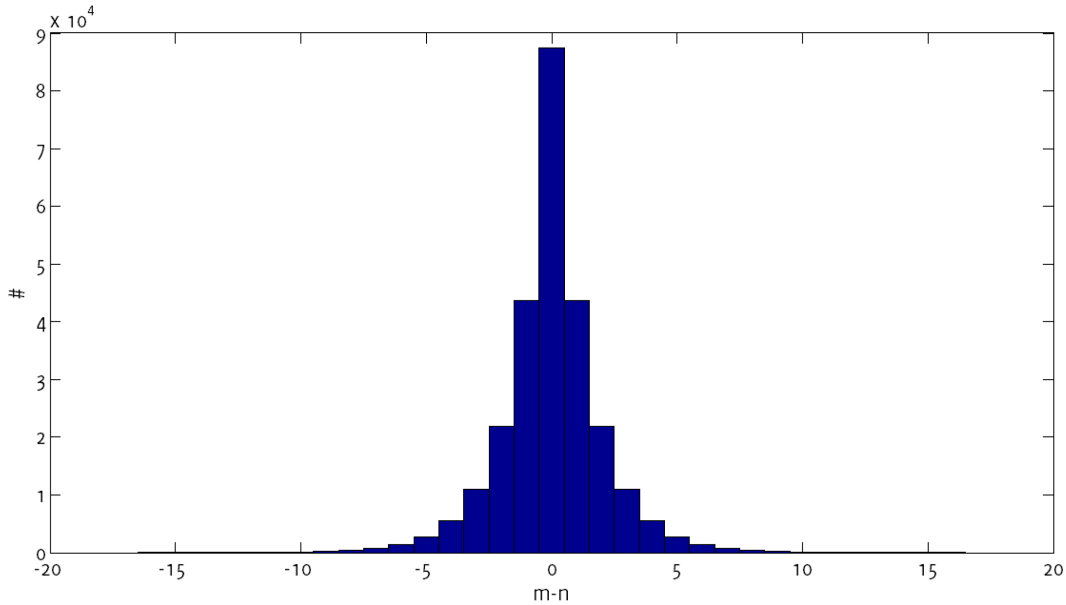


Figure 2 - The histogram of all initial m-n values for all starting numbers from 1 to 2^{18}-1.

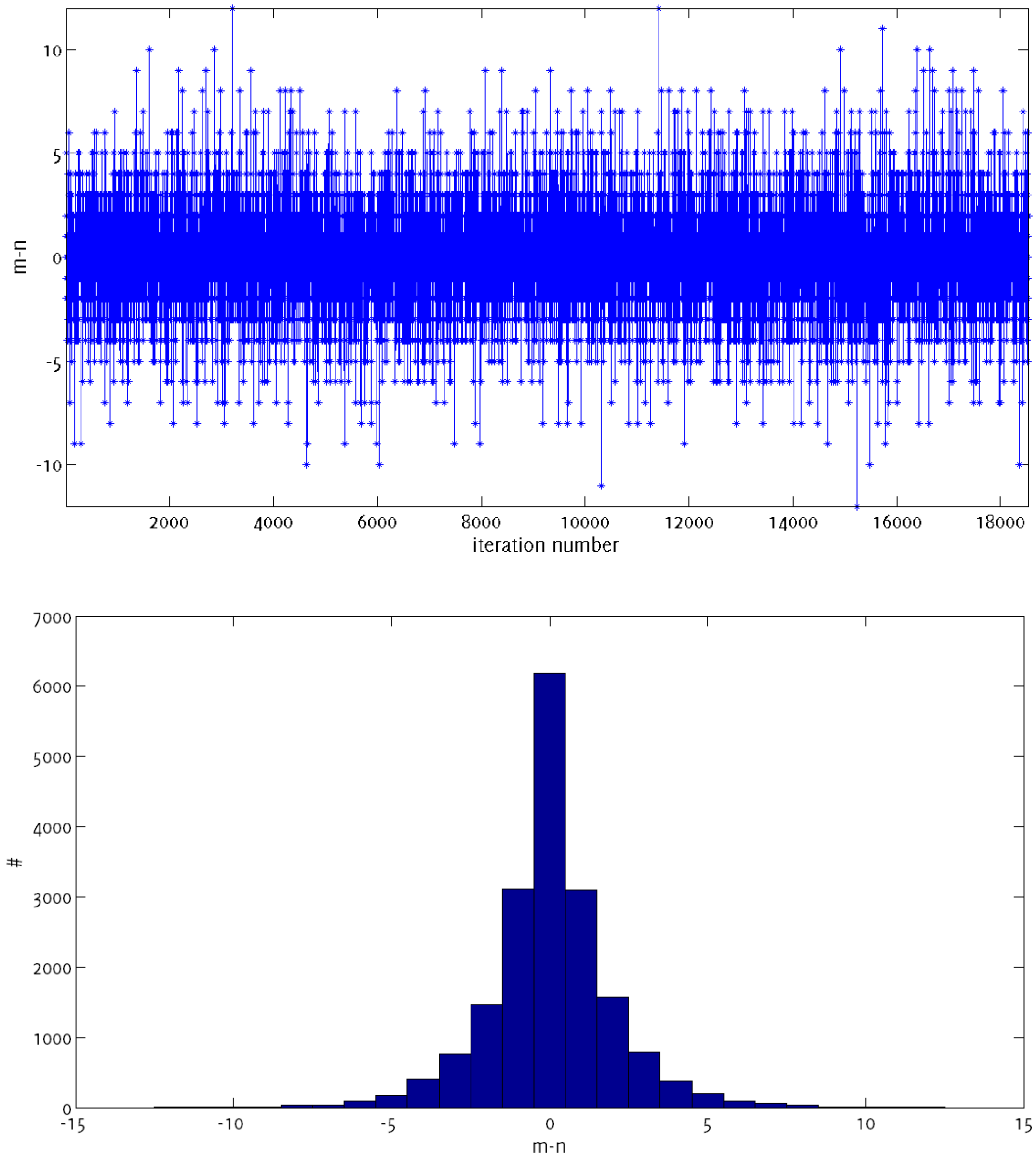
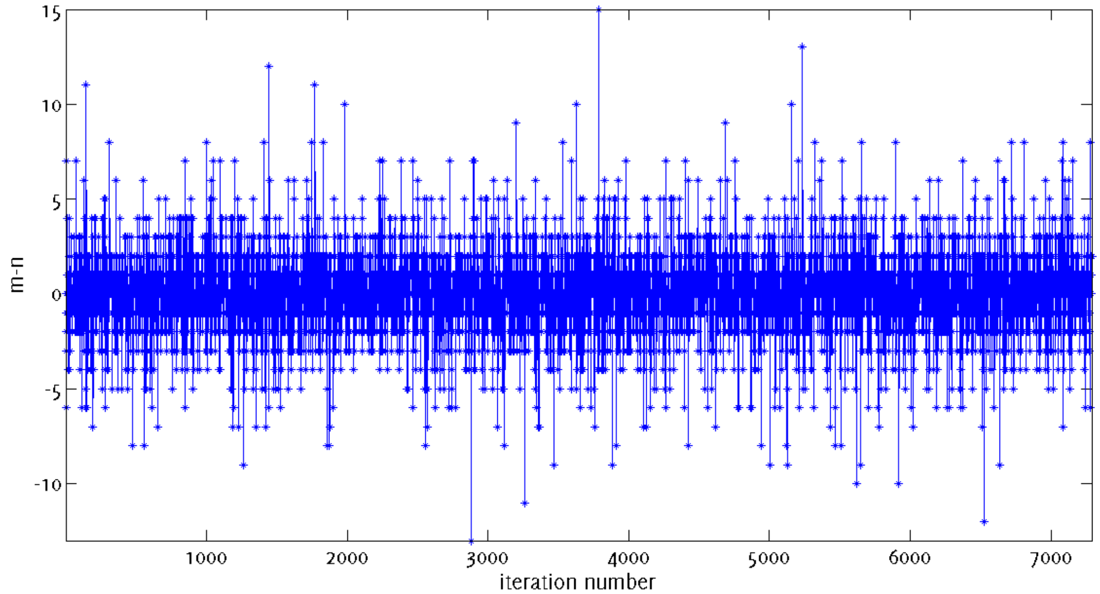


Figure 3 - The sequence of m-n values encountered for the Collatz sequence starting with an arbitrary, large, odd number 5^{6799} (top) and the corresponding histogram (bottom).



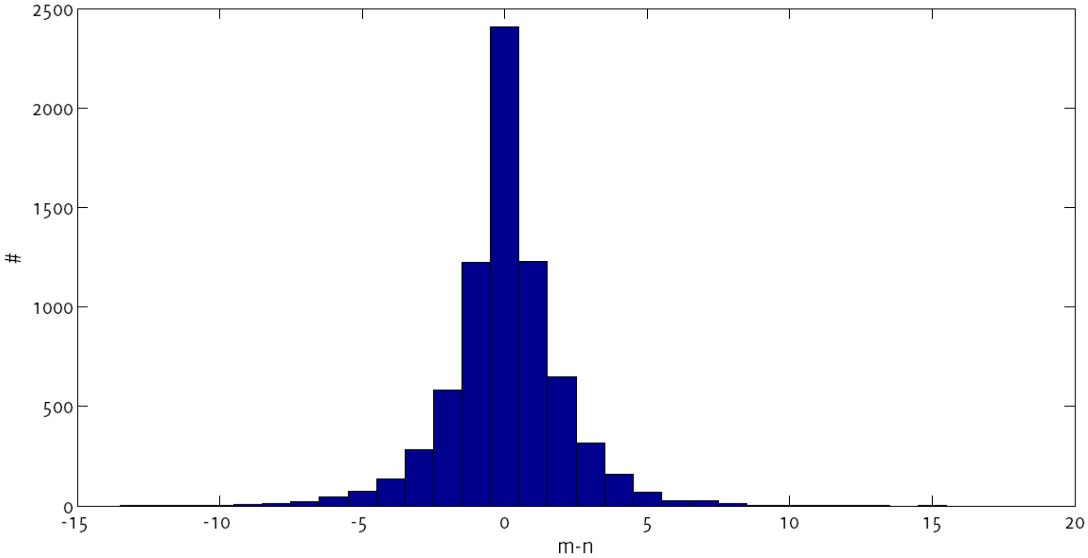


Figure 4 - The sequence of m-n values encountered for the Collatz sequence starting with an arbitrary, large, even number $77^{999}+1$ (top) and the corresponding histogram (bottom).

Discussion

This working paper highlighted that approximately symmetrical, Gaussian-like distributions may be inherently involved in the generation of Collatz sequences. Clearly, the relative increases in Collatz sequences represented by (3x+1)/2 are smaller or equal to the decrements by a factor of 2 represented by the x/2 operations, and for continued series growth, increments must overwhelm decrements. The numbers of these operations depend on the numbers of odd and even numbers encountered by the Collatz process, which the mathematical literature indicates, are difficult to analyze analytically. The new method of analysis presented in this paper shows that the difference of the numbers of the two operations may be following an approximately symmetrical, Gaussian-like distribution which has an approximately zero mean. This means that, despite the behaviour of Collatz sequences appearing without structure and random [1], there may be an inherent statistical order in the process. It also means that continued Collatz series growth may be not viable because these distributions will ensure that (3x+1)/2 and x/2 operations will become approximately equal in numbers as more iterations are carried out, thus not allowing the former to overwhelm the latter. Therefore, the Conjecture, perhaps, will always be true unless cycles other than the trivial 4, 2, 1 cycle exist.

References

1. Andrei S, Masalagiu C. About the Collatz conjecture. Acta Informatica. 1998.
2. Zöbelein C. About the proof of the Collatz conjecture. Arxiv Preprint. 2013.
3. Sayama H. An artificial life view of the Collatz problem. Artificial Life. 2011;17:137-140.
4. Bruschi M. Two Cellular Automata for the 3x+ 1 map. Arxiv Preprint. 2005.
5. Matthews KR, Leigh GM. A generalization of the Syracuse algorithm. Journal of Number Theory. 1987.
6. Lagarias JC. The 3x+ 1 problem and its generalizations. American Mathematical Monthly. 1985.
7. Crandall R. On the 3x+ 1 problem. Mathematics of Computation. 1978.