Noname manuscript No.
(will be inserted by the editor)

Cycles in Collatz Sequences: A Data Science Approach

Narrowing the condition for the existance of non-trivial Collatz Cycles

Christian Koch · **Eldar Sultanow**

Received: date / Accepted: date

Abstract The Collatz conjecture is an unsolved number theory problem. We approach the question by examining cycles in Collatz sequences. A cycle occurs when a number in a sequence is equal to the starting value. Every cycle, except the trivial variant, starting with the number one, would falsify the conjecture. Our research focuses on the original form of the Collatz problem, 3x+1, as well as the generalised variant kx+1. Aside from classical mathematical methods, we use techniques of data science. Based on the analysis of 250,000 sequences we provide a list of known cycles and show that there are clear restrictions for their occurence. We prove that cycles are are only possible for certain sequence-lengths. Future investigations benefit from this knowledge, as they will be able to concentrate on a significantly narrowed problem space in search for a comprehensive proof of the conjecture.

Keywords First keyword · Second keyword · More

1 Introduction

The Collatz conjecture is one of the unsolved "Million Buck Problems" [1]. When Lothar Collatz began his professorship in Hamburg in 1952, he mentioned this problem to his colleague Helmut Hasse. From 1976 to 1980, Collatz wrote several letters but missed referencing that he first proposed the problem

C. Koch
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

E. Sultanow
second address

in 1937. He introduced a function $g : \mathbb{N} \to \mathbb{N}$ as follows:

$$g(x) = \begin{cases} 3x+1 & 2 \nmid x \\ x/2 & \text{otherwise} \end{cases} \tag{1}$$

This function is surjective, but it is not injective (for example $g(3) = g(20)$) and thus is not reversible. The Collatz conjecture states that for each start number $x_1 > 0$ the sequence $x_1, x_2 = g(x_1), x_3 = g(x_2), \dots$ will at some point enter the so called trivial cycle $1, 4, 2$. One example is the sequence $17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1$ starting at $x_1 = 17$. The assumption has not yet been proven. If the conjecture were wrong, then for a starting number x_1 the sequence either would diverge indefinitely or enter a cycle different from the trivial one (a so called non-trivial cycle). Subject of our investigation are the non-trivial cycles and the question if these cycles are possible, under which condition they might occur and which cycles can be considered to be impossible. In order to specify compressed Collatz sequences containing only the odd members, Bruckman [3] for instance used the more convenient function that opts out all even integers:

$$f(x) = (3x+1) \cdot 2^{-\alpha(x)}, \text{where} \quad 2^{\alpha(x)} \parallel (3x+1) \tag{2}$$

Note that $\alpha(x)$ is the largest possible exponent for which $2^{\alpha(x)}$ exactly divides $3x + 1$. Especially for prime powers, one often says p^α *divides* the integer x *exactly*, denoted as $p^\alpha \parallel x$, if p^α is the greatest power of the prime p that divides x.

We follow this approach. A compressed Collatz sequence, which for example starts at $v_1 = 17$, we denote as $v_1, v_2, v_3, v_4 = 17, 13, 5, 1$ and $\alpha_1, \alpha_2, \alpha_3 = 2, 3, 4$ are the exponents indicating the divisions by two between two Collatz sequence members. For a given sequence $v_1, \dots, v_{n+1}$ the sum of these exponents we simply call $\alpha = \sum_{i=1}^{n} \alpha_i$.

2 Related Research

Hercher [2] dealed with conditions for a cycle and showed that for the (odd) Collatz sequence $v_1, \dots, v_n$ forming a cycle the following condition holds:

$$2^\alpha = \prod_{i=1}^{n} \left(3 + \frac{1}{v_i}\right)$$

Also here, the exponent α is the number of divisions by two within this sequence.

3 (Template) Defining the conditions for cycles

Text with citations [4] and [**?**].

3.1 Subsection title

as required. Don't forget to give each section and subsection a unique label (see Sect. 3).

Paragraph headings Use paragraph headings as needed.

$$a^2 + b^2 = c^2 \tag{3}$$

References

1. S. W. Williams, Million Buck Problems, National Association of Mathematicians Newsletter, 31(2), 1-3 (2000)
2. C. Hercher, Über die Länge nicht-trivialer Collatz-Zyklen, Die Wurzel, 6-7, 1-13 (2018)
3. P. S. Bruckman, RETRACTED ARTICLE: A proof of the Collatz conjecture, International Journal of Mathematical Education in Science and Technology, 39(3), 403–407 (2008)
4. Author, Book title, page numbers. Publisher, place (year)