Journal of Advances in Mathematics and Computer Science

Volume 39, Issue 4, Page 24-36, 2024; Article no.JAMCS.114167
ISSN: 2456-9968
(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

Block Format Solves the Collatz Conjecture

Khaled Moftah [a*]

[a] *Independent Researcher, 3319 Kenilworth Drive, Windsor, Ontario, N9E 4R3, Canada.*

Author's contribution

The sole author designed, analyzed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/JAMCS/2024/v39i41880

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc. are available here: https://www.sdiarticle5.com/review-history/114167

Original Research Article

Received: 04/01/2024
Accepted: 10/03/2024
Published: 23/03/2024

Abstract

Blocks are unit convergence between two consecutive odd numbers formed according to the three *x* plus one conjecture rules. The left odd number is the left hook, *L*, and the right odd number is the right hook, *R*. They include even numbers between their boundaries. They are divided into families ($F1$ = 5, 11, 17, … & $F2$ = 1, 7, 11, … & $F3$ = 3, 9, 15, …) and groups based on their group length (The number of the middle-even numbers between the two hooks). Blocks are taken individually and placed beside each other, similar to the domino tiles play, which, by their formulation, satisfies the conjecture rules. Formed chains reach number one in the convergence mode or continue generating odd positive numbers infinitely according to the generation mode. The final convergence to number one is reached because these blocks have all the positive integers included as left hooks ($L1$, $L2$, $L3$), and all the $F1$ and $F2$ odd positive numbers are included as right hooks ($R1$ and $R2$). Block rules mandate that a single left hook produces only one right hook. Accordingly, no looping or entanglement (Joining and consequent splitting) between chain branches would occur. Statistics show that *R* cannot increase infinitely. Repeated oscillation up and down without reaching number one would also violate the statistics. Statistics reveal that blocks of various lengths have a strict occurrence and repetition sequence along the positive integer series. Block lengths can extend infinitely, and each block length repeats its occurrence infinitely. In the generation mode, blocks are attached in reverse order to the conjecture/convergence rules. According to the rules, all positive integers can be generated starting from number one. Multiple sequences and clusters of specific block lengths occur according to specific rules and cannot continue infinitely.

**Corresponding author: Email: km.moftah@gmail.com;*

Keywords: *Three x plus one conjecture; Collatz conjecture; Ulam conjecture; Kakutani's problem; block format; odd positive integer families; base block; root block; block extension; lead-follower blocks; block statistics; block clusters.*

1 Introduction

Since its introduction by Collatz and others more than eighty years ago, the three *x* plus one conjecture has attracted varying interest over the years. Recently, there has been a revival of interest in the problem. Researchers have tackled the problem from different angles. This mathematical problem has been known under various names, including the Ulam conjecture, Kakutani's problem, The Thwaites conjecture, Hasse's algorithm, and the Syracuse problem.

Watching a video on the three *x* plus one conjecture created by Dr. Derek Muller in his Veritasium video series about a year ago triggered my interest in the subject.

The inherent characteristic of the three *x* plus one iteration makes it difficult to map it while expanding on one part of their domain and contracting on another part of their domain. The general class of these kinds of problems is of definite importance and is currently of great interest as an area of mathematical (and physical) research. J. Lagarias wrote a comprehensive book on the problem [1]. In his review article [2] in which he cites 99 references, Lagarias acknowledges that the problem has a general connection with other areas of mathematics, including number theory, dynamical systems, theory of computation, scholastic processes, probability theory, and computer science. This made it a respectable topic for mathematical research. Several well-known mathematicians have contributed their thoughts to it, including Y. G. Sinai [3,4] and T. Tao [5], among many others.

The original work on the $3x + 1$ problem viewed it as a problem in number theory. This effort continues in parallel with other veins of research.

Lei Li stated that he reached a high-precision formula for the average number of multiplications and divisions for the Collatz problem based on an analysis of the average computational complexity of the recursive algorithm [6]. In his paper, he suggests a proof of the conjecture based on "Simple mathematical induction."

Masahi Furuta used the Theorem proving support system "Idris" and Egison's computer algebra system to prove the three *x* plus one conjecture [7].

Agelos Kratimenos [8] claims that he discovered a formula for a characteristic function, which describes the functionality of the paths taken for each number based on the Collatz Sequence. The last section of his article provides proof that every number will eventually reach one using the characteristic function.

In their structured approach, Ken Surendran and Dasarazu Krishna Babu [9] included observations about the occurrence of the "Starters" (The odd numbers divisible by 3). They recognized the importance of the expressions $4x + 1$, $4x + 3$, 6m + 1, and 6m + 5 along the convergence path. Based on the odd numbers categories, they defined other considerations and mapped different structures along the convergence path.

Jinqing Zhang and Xintong Zhang [10] performed statistical analysis on the slopes of the Collatz iterations and proved that they are less than one, which is a sufficient and necessary condition for Collatz iteration convergence.

Recently, Chin-Lon Wey [11] developed a Collatz graph consisting of Collatz nodes located at various levels of the graph based on specific rules and showed that the developed Collatz graph generates complete Collatz trajectories for all positive integers. The sequences converge to one, claiming completeness of the developed Collatz graph and proof of the Collatz conjecture. The reference list includes a link to the Veritasium footage and a general description of the problem [12, 13]. Other authors have recently contributed their new methods and findings [14-20]

This current study explores and defines patterns in which natural numbers behave with consideration of the three *x* plus one conjecture. Natural odd numbers are divided into families and groups. Patterns and behaviors of these families and groups are investigated through the light of the three-plus-one conjecture. A unique block format was used during the study, which helped develop general governing rules. Distinctive characteristics of these families and groups formed from them have been discovered. The study investigates these characteristics in the convergence (Conjecture) mode and the opposite generation (Bottom-up, revers-conjecture) mode and suggests a statistical solution proving the validity of the conjecture.

If we can move from number one to all the natural positive integer series by applying the generation (Reverse conjecture) rules, then we can move down in the opposite direction from any natural positive number to number one, which proves the conjecture. Along this line of thought, the possibilities of looping, ever-increasing, and plateau formation are tackled.

The convergence mode moves in one direction in one chain from the start to number 1 in a final converging step (If so happens). During this one-direction path, other chains starting from other positive integers join the path at specific points. This creates a look of coming down from many (infinite) branches of a tree or a bush to more and more common ones to final trunks above ground, which are also infinite. These trunks join only underground at one root, the number one. On the other hand, in the generation mode (Reverse-conjecture rules), the movement starts at the root and spreads into primary and secondary branches. Different branches are explored until the target number is reached. This procedure is tedious and exhausting if one specific number is the target. However, general rules applicable to the whole set of the natural positive integers can be obtained. The current study aims to develop relationships and formulas that prove that all natural numbers (Positive integers) can be generated by applying these rules. There has been work along the convergence mode in checking individual numbers to prove or disprove the final convergence to number one. Other than finding a non-converging number, this kind of work just adds to a stronger belief in the validity of the conjecture.

2 Method

The conjecture steps are

1. Start with any natural number, even or odd.
2. If the start or result numbers are odd, multiply them into three and add one.
3. If the start or resulting number is even, divide by two.
4. Continue dividing even numbers by two until reaching an odd number.
5. Repeat the operations. The chain of these operations should always converge to 4, 2, 1.

The mathematical recursive expression of these rules is:

$$C(x) = \{3x + 1 \text{ if } x \equiv 1 \pmod 2),\ x/2 \text{ if } x \equiv 0 \pmod 2)$$

The generation rules (Move in the bottom-up/reverse conjecture direction) are:

1. Start with number 1.
2. For this or any resulting odd number, multiply by 2.
3. Deduct one and divide the result by 3 for the resulting even numbers
4. If the result of the previous step has a remainder, continue multiplying into 2
5. Continue the operation until a whole number without a remainder is obtained.
6. If the result of the previous operation is a number divisible by 3.
7. The operation can be terminated here if the required number is found
8. Otherwise, multiply the even number into 2 to skip this number and continue the operation
9. Explore all branches and complete calculations until the target number is reached.

The generation mode (Opposite-Convergence, Bottom-up, Generation procedure) can be continued infinitely until all the infinite positive number series are generated.

The Block Format

Procedures and formulations here apply a format called the block format. In this format, the process, whether convergence or generation, divides the chain of positive natural numbers into blocks. Blocks are treated as individual entities. A block starts with an odd number, called the left hook (L), and ends with an odd number, called the right hook (R). Even numbers are sandwiched between these odd numbers and are called the middle-even numbers (*E*). In this sense, a block is a unit convergence between two successive odd numbers. Blocks can form unit generation when R and *L* switch places. The odd number at one end of one block is repeated at the other end of the following block. Odd numbers shall always be treated in these joint couples like the domino tiles when dealing with the three *x* plus one conjecture. Block format in the convergence mode is presented in square brackets as [*L*/*E*/*R*] and in the generation mode as [> *R*/*E*/*L*]. The > sign is placed ahead of the *R* in the generation mode for distinguishing purposes.

Examples of short converging chains:

[21/6/1] (One block)
[13/3/5] [5/4/1] (Two blocks)
[75/1/113] [113/2/85] [85/8/1] (Three blocks)

In the generation mode, these chains would be written as:

[> 1/6/21]
[>1/4/5] [> 5/3/13]
[>1/8/85] [>85/2/113] [>113/1/75]

Obviously, if the starter in a converging chain is an even number, it is divided by two iteratively until an odd number is reached, at which point the block format starts.

Relationships between the left hooks, right hooks, and middle-even numbers are formulated and produce specific criteria for the blocks they belong to. Formulations include basic formulas as well as parametric ones. The parametric formulas present the left hook (*L*) in the form $L = A + BK$ and the right hook in the form $R = C + DK$, where *A* and *C* are odd numbers, *B* and *D* are even numbers, and *K* is an index taking values: 0, 1, 2, … to infinity. Blocks are categorized primarily based on the number of the middle-even numbers (The length of the block), then according to the R family (*R*1 or *R*2)) and according to the *L* family (Called the *L* brand: *L*1, *L*2, *L*3). Each categorized group of blocks has specific characteristics in general and in relation to the conjecture rules. Each group of blocks extends vertically as the *K* index increases to infinity. These block groups are connected in chains, which results in the well-observed unique behavior of the convergence or generation chains. Statistics of the occurrence and repetition of specific block groups have been studied. Formulas and statistics for individual block types and block couples in a Lead-Follower arrangement and clusters of specific block groups have also been studied.

In the block format, analyses of the odd numbers are sufficient to generalize the conclusions to the whole positive integer series.

The series of positive natural odd numbers is divided into three main families. The series of each family continues infinitely. The sum of these three families together forms the infinite odd numbers series. These families are:

Family *F*1: 5, 11, 17, ..., ∞
Family *F*2: 1, 7, 11, ..., ∞
Family *F*3: 3, 9, 15, ..., ∞
Formulas:
Family *F*1: $O = 5 + 6K$
Family *F*2: $O = 1 + 6K$
Family *F*3: $O = 3 + 6K$
O: Odd number
$K = 1, 2, 3, \ldots$

Blocks start with a left hook belonging to any of the three main families, described as brands *L*1, *L*2, and *L*3. Blocks end with an *R*. *R* hook belonging to *F*1 are denoted as the R1 family, and the R2 family for R belongs to *F*2. All blocks of the *R*1 family have an odd number of the middle-even numbers, and all blocks of the *R*2 family have an even number of the middle-even numbers. There is no *R*3 family, as shall be described later.

3 Formulations and Results

Formulations and results are included in Boxes. Each Box contains a data set in one or more tables. Related definitions and formulas are included in the same or the following box. This is in addition to the Definitions and nomenclature Box (Box D). This facilitates reading the data and grasping the related concepts included.
Box B1 includes blocks with fixed middle-even numbers (Fixed E) in the convergence mode. Box B1-B includes formulas governing the relationships of the block entities included in Box B1. Box B2 shows the first occurrence and the repetitions of different block lengths (Block groups) along the odd positive integers' series. Box B3 includes a summary of the generated L values versus E and K. Box B3-B contains observations and calculation formulas for Data in Box B3. Box B4 summarizes statistics of the occurrence of different block lengths. Box B4-B includes more statistics for R and L from the data presented in Box B4.

Box F1 includes tables of two consecutive blocks with both the lead and follower blocks having an odd number of E. Box F2 tabulates similar data for odd E in the lead block and even E in the follower block. Box F3 includes tables of two consecutive blocks where the lead block has an even number of E and the follower block has an odd number of E. Box F4 tabulates similar data for even E in both the lead and the follower blocks. Data in these four Boxes is arranged in the generation mode. Boxes F5 and F6 summarize the frequency of occurrence of these combinations and conclude statistics about their R and L.

Box C1 includes tables of Consecutive Single Middle-even number Blocks (CSMB) starting with $L3$, while Box C2 includes similar tables for blocks starting with $L2$. Box C3 includes formulas governing data in Boxes C1 and C2. Box C4 includes a table and formulas of one cluster type of the Consecutive Double Middle-even numbers Blocks (CDMB) starting with $L3$. Box C4-B includes formulas governing data presented in Box C4. Box C5 includes examples of CSMB and CDMB having three blocks ($3M$) and starting with $L3$. These examples are shown in the parametric form.

Box EX1 includes examples of the $L3$ skipping in the convergence mode and compares it to the generation mode. Box EX2 includes the parametric form of the convergence of number 27 in a one-lead, one-follower combination. Box EX3 contains an example of CSMB clusters along with a figure. Box EX4 includes data for $20M$ CSMB (A cluster of 20 Consecutive Single Middle-even numbers Blocks) and the complete convergence of this chain to number one. Box EX5 includes data for two types of CDMB, starting with $L1$. Box EX6 contains an example of the CDMB cluster having twenty blocks ($20M$) starting with $L3$ and completing the chain to final convergence to 1. Box EX7 includes examples of stretches of similar sequences of varying block lengths. Box EX8 contains the application of the proposed stretch formulas to selected sequences of the convergence of number 27.

4 Discussion

According to the three x plus one conjecture (The conjecture), even positive numbers are divided by two until an odd number is reached. Then, the odd number is multiplied into three, and one is added. The resulting even number is divided by two, and the process is repeated consecutively. The final convergence should reach number one. Even numbers descend smoothly to an odd number. Therefore, studying the positive odd numbers is sufficient to prove or disprove the conjecture. The Definition Box D details important properties and formulas for the blocks.

The general formula relating R to L in the convergence mode is:

$$R = (3/2^E)\, L + (1/2^E) \tag{1}$$

Only one single E value (Number of the middle-even numbers) produces a specific whole odd R-value for any odd whole L value. That means each block is unique, starting with the selected L value producing a unique R-value. E is determined by the group L belongs to. Boxes B1 and B1-B show this and other L/E/R relationships. Blocks are tabulated in groups of block lengths (Fixed E). General governing formulas are also contained in the box.
The general formula for the generation mode is:

$$L = (2^E/3)\, R - (1/3) \tag{2}$$

Series of E values satisfy the formula for whole odd L and R values. These R series produce a corresponding series of L values. The L values in these series obey specific rules and can be predicted. Reference is made to Boxes B3 and B3-B for that property. The L series is unique to each R-value. One specific L can not be generated from two different R values.

The base block for the $R1$ family has $E = 1$, and the $R2$ family has $E = 2$. There is no limit for E, as shown from the E series in the generation mode. For one R-value, infinite E values produce infinite odd L values in extended blocks having $E = 3, 5, 7, \ldots$ for $R1$ Blocks and $E = 4, 6, 8, \ldots$ for the $R2$ family.

If L and R are known, E can be calculated by the formula:

$$E = (\mathrm{Log}(3L + 1) - \mathrm{Log}(R))/\mathrm{Log}(2) \tag{3}$$

The $F3$ family of positive integers has unique properties. Numbers of family $F3$ can start a chain in the convergence mode, where they are used as the first left hooks ($L3_0$) in the first block. $F3$ numbers can not occur as an R of a block and consequently do not occur in the middle of a converging chain. This is a consequence of the convergence rules.

In the generation mode, members of the $F3$ family can be generated as L, but only once, and the generation process stops there. This terminates the procedure of generating more blocks along this chain. In this mode, $L3$ can be skipped by extending the block length by adding more to the middle-even numbers (Two more even numbers at a time or multiplying the middle-even number by 2^2), as shown in Box EX1.

In the generation mode, it is possible to move from any $R1$ or $R2$ values to reach all the members of the $L3$ (3, 9, 15, …). Thus, moving back to the convergence mode starting from the $L3$ brands (Left hooks belonging to the $F3$ family) is sufficient for studying the conjecture, and there is no need to study $L1$ and $L2$ separately. After the starting block, L of the next block always belongs to $F1$ or $F2$ ($L1$ or $L2$), and R also belongs to $F1$ or $F2$ ($R1$ or $R2$).

The following Table (Table 1) shows the six elementary blocks starting with $L3$.

Table 1. *R*1 and *R*2 produced from *L3* – Convergence Mode

***L*3**	***E***	***R*1**	***L Series***	**R Series**
$3 + 12K$	1	$5 + 18K$	3, 15, 27, ...	5, 23, 41, …
$117 + 192K$	5	$11 + 18K$	117, 309, 501, ...	11, 29, 47, …
$45 + 48K$	3	$17 + 18K$	45, 93, 141, ...	17, 35, 53, …
***L*3**	***E***	***R*2**	***L Series***	**R Series**
$21 + 384K$	6	$1 + 18K$	21, 405, 789, ...	1, 19, 37, …
$9 + 24K$	2	$7 + 18K$	9, 33, 57, ...	7, 25, 43, …
$69 + 96K$	4	$13 + 18K$	69, 165, 261, ...	13, 31, 49, …

The six formulas in the $R1$ & $R2$ columns cover all members of the $F1$ and $F2$ families.
The parametric form of the six L and six R formulas can be written in a block format [$A + BK/E/C + DK$] as follows:

[$3 + 12K$)/1/5 + 18K]
[$117 + 192K/5/11 + 18K$]
[$45 + 48K/3/17 + 18K$]
[$21 + 384K/6/1 + 18K$]
[9 + 24K/2/7 + 18K]
[$69 + 96K/4/13 + 18K$]

The general parametric forms for L and R are presented as follows:

$$L = A + BK \tag{4}$$
$$R = C + DK \tag{5}$$

Table 2 shows the parametric forms of *L* and *R* as *E* steps up (Increases by two). Odd *E* is for *R*1, and even *E* is for *R*2.

Table .2 *L* and *R* Parametric Forms

E	*L* Formula	*R* Formula
1	3 + 4*K*	5 + 6*K*
3	13 + 16*K*	5 + 6*K*
5	53 + 64*K*	5 + 6*K*
7	213 + 256*K*	5 + 6*K*
9	853 + 1024*K*	5 + 6*K*
11	3413 + 4096*K*	5 + 6*K*
2	1 + 8*K*	1 + 6*K*
4	5 + 32*K*	1 + 6*K*
6	21 + 128*K*	1 + 6*K*
8	85 + 512*K*	1 + 6*K*
10	341 + 2048*K*	1 + 6*K*
12	1365 + 8192*K*	1 + 6*K*

The index *K* takes values along the 1, 2, 3, … series. Thus, the tabulated *L*-*R* here and in the boxes can extend infinitely as the *K* value increases infinitely. Specific formulas govern the *L*-*R* relationship within each group (Block length). The general formulas for *L, R, E,* and *K* are stated below.

For odd E, $R_0 = 5$

$$L_K = (5/3)*2^E - (1/3) + 2^{E+1}K \tag{6}$$

$$R = 5 + 6K \tag{7}$$

For even E, $R_0 = 1$

$$L_K = (1/3)*2^E - (1/3) + 2^{E+1}K \tag{8}$$

$$R = 1 + 6K \tag{9}$$

$$R_0 = R \text{ at } K = 0 \tag{10}$$

A complete list of *L*/*E*/*R* relationships is presented in Boxes B1 and B1-B.

Box B1 also shows that each *L* brand repeats sequentially when *E* or *K* increases. This adds more of the characteristics of the families of odd numbers. The *L* brands can be categorized together at jumping *E* (*E* increases by six). Categorization of the same *L* brands along the *K* is done by increasing it by three at a time.

Box B2 categorizes the first occurrence and subsequent repetition (Reoccurrence) of block groups (Blocks having the same number of middle-even numbers *E)* per the *L* brand (*L*3, *L*2, *L*1). When the three L brands are merged, the move along the series of the odd numbers is the stepwise addition of two (1, 3, 5, …). 1*E* blocks occur every two steps (Addition of four) starting from number three along the odd number series. The 2*E* blocks reoccur every four steps (Addition of eight) starting from 1. The 3*E* blocks reoccur every addition of sixteen, beginning at number thirteen, and so on. Linear increase of *E* decreases the occurrence of the block by doubling the *B* value (A multiplier of two on the *B*), as shown in Box B1. Furthermore, it has almost the same geometric effect on the ratio of R/L in opposite slopes. The first occurrence of *L* in a specific block length is its *A*, and the repetition sequence is by adding the *BK* value in the parametric form.
The place of the odd number in the odd number series (Its value) very well determines its L brand, the block length they have, and the R family they belong to. The resulting *R*, used as *L* in the subsequent block, has a specific block length, and so on.

Box B4 summarizes the statistics of all block lengths for the *R*1 and *R*2 families in the generation mode. The sum of the *L* series in the *R*1 family equals one-third of the positive integer series, and the sum of the L in the R2 family equals one-sixth of the positive integer series. The total of these *L* values equals one-half of the positive integer series or the whole series of the odd positive integers. Thus, no number along the odd positive number series can be missing in all possible generation chains.

Box B3 shows *L* values vs. *R*1 and *R*2 in the generation mode as *E* increases horizontally and as *K* increases vertically. Both *E* and *K* can increase infinitely.

The general formulas presented in Box B1-B are repeated here.

For even *E:*

$$L_{E,0} = 4^n + 4^{n-1} + 4^{n-2} + \ldots + 4^{n-n} \quad (11)$$
$$n = (E - 2)/2 \quad (12)$$

For odd *E:*

$$L_{E,0} = 3{*}4^n + 4^{n-1} + 4^{n-2} + \ldots + 4^{n-n} \quad (13)$$
$$n = (E - 1)/2 \quad (14)$$

For even and odd *E*:

$$L_{E,K} = L_{E,0} + 2^{E+1}K \quad (15)$$

These formulas generate one single *L* for each *E* and *K*. *R* values in the far-left column are themselves generated *L* in the body of the table except for number one, which is the root number for all generations. One can move from number 1 and generate any number in the odd positive number series by successive generations, connecting the generated *L* in one block as *R* in the next block. Then, moving back to the convergence mode can be traced on a reversed path.

A block can be extended to a huge number of middle-even numbers. The first occurrence of long blocks gets bigger, too, and its recurrence frequency diminishes, but its effect on its *L* (The *R* over *L* ratio) is bigger, too. The first occurrence of a block having *E* = 30 starting at *L*3 and having *R* = 1 is *L*3 = 357,913,941. This block statement is: [357,913,941/30/1]

The second occurrence of such a block with *E* = 30 is after adding 2,147,483,648 to the first *L*. The *L* brand of this block is *L*2, and it has *R* = 7. The *L*3 hook brand repeats occurrence after adding 6,442,450,944 to the first *L*3. This block statement is [6,800,364,885/30/19].

Statistically, this block length occurs $1.86{*}10^{-9}$ times less frequently than blocks with 1*E*. In other words, blocks having 1*E* occur $5.37{*}10^{8}$ more times than blocks having 30*E*.

The effect of a block (*f*) in a converging series is the ratio of its *R* over its *L* (The *R* of the previous block). The effect of a block (*e*) is specific to its R, but the change is very small in large numbers. The block's impact (i) is an approximation that neglects the absolute term from the *R* definition.

$$e = R/L = (3L/2^E + 1/2^E)/L \quad (16)$$
$$e = (3/2^E) + (1/2^E)/L \quad (17)$$
$$i = 3/2^E \quad (18)$$

The impact (*i*) of the 30*E* block = $3/2^{30} = 2.79{*}10^{-9}$

The impact (*i*) of 1*E* block = $3/2^1 = 1.5$ (19)

It is worth noting here that, for example, once the number 6,800,364,885 occurs along a converging chain, it surely converges to number 19 after 30 iterative smooth divisions by tow. Then, from number 19, the convergence proceeds in the block stretch, as shown below.

[19/1/29] [29/3/11] [11/1/17] [17/2/13] [13/3/5] [5/4/1]

Convergence from 6,800,364,885 to number one took seven blocks.

Converging chains include all sorts of skewing from the statistics.

Number 21 converges to number one in one block, and six middle even numbers, as follows:

[21/6/1]

Number 19 needs six blocks and fourteen middle even numbers for the final convergence to number one.

Number 27 requires 41 blocks and 70 middle-even numbers to converge to number one.

A strain of blocks would have the same R and jumping number of E (The Middle even numbers increase by six) Examples are:

[3/1/5], [213/7/5], [13653/13/5]

All the L values 3, 213, 13653, 873,813, … require two blocks only to converge to number one. The first block is the base block with E = 1; the other blocks are extended blocks formed by adding six to the middle even numbers.

Longer chains contain skewed parts (stretches) within the complete converging chain. This is evident in the cluster examples, which will be discussed later.

The block study is further extended to the sequence of couples of blocks as contained in Boxes F1 through F6. These boxes show that the sum of the occurrence of follower blocks of all lengths equals the odd natural numbers series. Moving through all generated branches generates the whole odd numbers series. Generation branches are infinite. This asserts what was found in Box B4 for the single-E block study.

In the convergence mode, Each *R* follows from the previous *L*. No *R* can be produced from two or more *L* values in two or more branches. After two branches join in a single *R* (A node), the formed trunks move along the convergence rules and may join other branches and trunks along the convergence path, but they do not split into new branches again.

One possibility of not reaching the final 4, 2, 1 convergence includes infinite oscillation and not reaching the number one. This may happen when the *R*-value oscillates around a plateau or many plateaus. Another possibility is a continuous increase of the *R*-value along the convergence of a chain without finally converging to number one.

The conjecture asserts the convergence to number one, which means that this sequence of ever-increasing *R*-value should stop at some point and not continue endlessly. This trend is seen in Boxes C1 through C3. Formulas for the CSMB cluster occurrence and the rule for starting *L* in such clusters are proposed and included in these boxes.

Proposed formulas conclude that the number of such blocks in those clusters is finite. After that, the chain continues in the typical observed up-and-down pattern until it finally converges into one. In these CSMBs, *R* reaches infinity only when the starting *L* is infinite. Clusters of different ranks (Number of the CSMB) occur along the convergence path of a chain and follow the governing formulas. The chain continues, including different block lengths and types of clusters, until it finally converges to 1.

This is shown in Box EX4, which includes a chain with 20 CSMB and the continuation to final convergence to number one in 164 blocks. With this relatively high number of blocks, longer (Extended) blocks (3*E*, 4*E*, …) are more probable, and their dampening effect on the *R*-value is more pronounced. The *R* goes high after one or more CSMB and then goes down in one or more following longer blocks (Having two or more *E*) and up and down again in the commonly found hailstone pattern. According to the statistics, the Single Middle, even number blocks (SMB) occur half the time of all block lengths. The sum of the dampening effect of the longer blocks on the *L*-value outweighs the shooting effect of the SMB, and a final convergence to number one occurs. It is only a matter of adding a finite number of blocks. The number of the blocks reaches infinity only if the starting *L* of a CSMB cluster is also infinite or if these SMB and CSMB reappear infinitely. Statistics prohibit this occurrence, too.

Boxes C4 and C4-B include tables and formulas for one type of Consecutive Double Middle-even number Blocks (CDMB) clusters.

The effect of this type of block is:

$$e = R/L = (3L/2^E + 1/2^E)/L \quad (20)$$
$$e = (3/2^2) + 1/2^2)/L \quad (21)$$
$$i = 3/4 \quad (22)$$

The effect and the impact of the CDMB is a moderate decrease in the L-value. As for the CSMB, the number of the CDMB increases to infinity as the starting L increases to infinity. Reference is made to Box EX6 to see the effect of a CDMB cluster having 20M and the complete convergence to number one, which shows the repeated but limited occurrence of SMB, CSMB, and other block lengths along the convergence path. Notably, 8E and 10E blocks occurred along its convergence path.

There are many other cluster patterns. Formulating the governing formula is similar in principle to those mentioned in this treatment.

Box EX5 shows examples of groups of CDMB clusters.

The root blocks are the blocks having $R_0 = 1$. Their block formula is:

[L/E/1] on the convergence mode or
[> 1/E/L] in the generation mode
There are three possibilities:

1) Blocks that have $L3$ following the form:

[> 1/E/$L3$] where E = 6, 12, 18, … and $L3$ = 21, 1365, 87381, …

2) Blocks that have $L2$ following the form:

[> 1/E/$L2$] where E = 2, 8, 14, …$L2$ = 1, 85, 5461, …

3) Blocks that have $L1$ following the form:

[> 1/E/$L1$] where E = 4, 10, 16, … and $L1$ = 5, 341, 21845, …

Each $L1$ and $L2$ shall be the R for the next block in the generation process, and the process continues. This process creates branches along the positive integers' series. Branches diverge and continue infinitely.

The L brand changes sequentially as the E steps up (Increases by two). An Example of the alternation from $L2$ to $L1$ to $L3$ for the same K value is shown in Table 3 below:

Table 3. Example of *L* Brand Alternation as E Increases

> *R*	***E***	***L* Value**	***L* type**
5	1	3	$L3$
5	3	13	$L2$
5	5	53	$L1$
5	7	213	$L3$
5	9	853	$L2$
5	11	3413	$L1$

L changes from $L3$ to $L2$ to $L1$, and the sequence repeats every addition of six to E.

The generation (Reverse-conjecture) formulas are:

$$L_E = (2^E/3)R - (1/3) \quad (23)$$

$$L_{E+2} = (2^{E+2}/3)R - (1/3) \quad (24)$$

Which leads to the formula:

$$L_{E+2} = 4L_E + 1 \tag{25}$$

Moving through the same *L* brand and jumping *E* (E increases by six at a time) follows the formula:

$$L_{E+6} = 64L_E + 21 \tag{26}$$

The block [> 1/2/1] is the only root and base block, but only its extensions are used in the generation process.

The statistics in Boxes F5 and F6 show the frequency of different follower block lengths for R1 and R2. The occurrence frequency of each block length is well defined by its length (Group or number of the middle-even numbers). These frequencies add to reach the whole odd positive integer's series. This adds weight to what was calculated in Box B4 for single blocks of different lengths.

Table 4 below shows the first occurrence, the sequence of block lengths, and the *R*-value moving along the positive odd number as *L* in the convergence mode. There is a very determined pattern, which is evident from the rearrangement shown in Box B1.

Table 4. *R-L* values

L	*E*	*R*	*L*	*E*	*R*	*L*	*E*	*R*
1	2	1	25	2	19	51	1	77
3	1	5	27	1	41	53	5	5
5	4	1	29	3	11	55	1	83
7	1	11	31	1	47	57	2	43
9	2	7	33	2	25	59	1	89
11	1	17	35	1	53	61	3	23
13	3	5	37	4	7	63	1	95
15	1	23	39	1	59	65	2	49
17	2	13	41	2	31	67	1	101
19	1	29	43	1	65	69	4	13
21	6	1	45	3	17	71	1	107
23	1	35	47	1	71	73	2	55
25	2	19	49	2	37	75	1	113

The complete convergence of number 7 is shown in Bx EX2 arranged in a one-lead-one-follower table using the parametric formulas. CSMB and CDMB clusters in this chain can be checked using the formulas stated in Boxes C3 and C4-B. For example, a 5M CSMB cluster occurs starting at $L2 = 319$.

Checking formula Q (In Box C3) for the calculation of $L2_0$ gives:

$$L2_0 = 2^{M+1} + 2^{M+3} - 1 = 319 \tag{27}$$

In the convergence mode, moving from the starting L_0, branches join at specific numbers (Nodes) to form "Thicker branches." This process continues until all branches meet at the root blocks.

Boxes EX7 and EX8 include parts (Stretches of converging chain with no apparent pattern of their block lengths. Proposed formulas are included for the reoccurrence of such a sequence of the block lengths in similar stretches. In addition to the examples studied here, many other cluster patterns and types occur. Expanding the study of the block followers, clusters, and stretches may make it possible to generalize their governing formulas with systematic derivations.

If these proposed formulas are universally validated, then, for example, a block length sequence similar to the convergence of number 27 would have the following values:

Total number of blocks, $M = 41$
Total number of middle-even numbers, $S = 70$
Distance between starting L of two consecutive sequences, $h_0 = 2^{71}$
Starting L of the first sequence, $L_{0,0} = 27$
R at the end of the first sequence, $R_{M-1,0} = 1$
Starting L of the second sequence, $L_{0,1} = 27 + 2^{71}$
L at the end of the second sequence, $L_{M+1} = 1 + 2*3^{41}$
Note: R_{M-1} will be L_M for stretches not ending at number one.

5 Concluding Remarks

This treatment used a unique arrangement of the positive numbers called the block format, which helped in the study of the Collatz (The three *x* plus one) conjecture and produced valuable statistics leading to the conclusion of its validity. The study explored the convergence path as well as the opposite generation path. The Probability of not reaching number one as a final convergence was investigated with given examples.

Supplementary Materials

Supplementary material is available in the following link: https://journaljamcs.com/media/Supplementary_2024_JAMCS_114167.pdf

Competing Interests

Author has declared that no competing interests exist.

References

[1] Jeffery C. Lagarias, "The Ultimate Challenge: The 3x+1 Problem, American Mathematical Society; 2010. ISBN: 978-1-4704-7289-4
Available: http://www.cecm.sfu.ca/organics/papers/lagarias/paper/html/node3.html

[2] Jeffrey C. Lagarias, "The 3x+1 Problem: An Overview", arXiv:2111.02635v1, November 2021

[3] Ya G. Sinai, "Statistical (3*X* + a)-Problem", Comm. Pure Appl. Math. 2003;56(7):1016-1028

[4] Ya G. Sinai, "Uniform Distribution in the (3*X* + a) Problem", Moscow Math. Journal. 2003;3(4):1429-1440

[5] Tao T. Almost all orbits of the Collatz map attain almost bounded values", ARxiv:1909.03562 *Submitted on 8 Sep 2019 (v1), last revised 15 Feb 2022 (this version, v5)]* Published online by Cambridge University Press: 2022

[6] Lei Li "A Proof for the Collatz Conjecture", arXiv: 2209.13541, 2022

[7] Masashi Furuta "Proof of Collatz Conjecture Using Division Sequence", Advances in Pure Mathematics. 2022;12(02):96-108.

[8] Agelos Kratimenos, "Proof of the Collatz Conjecture", arXiv.1911.02400v2, 2019

[9] Ken Surendran, Desarazu Krishna Babu, "Collatz Conjecture: Exposition and Proof Through A Structured Approach",
arxiv.org/abs/2008.02161, 2020

[10] Jinqing Zhang[1], Xintong Zhang, "Collatz Iterative Trajectories of All Odd Numbers Attain Bounded Values", Journal of Applied Mathematics and Physics. 2023;11:10.

[11] Chin-Long Wey, Proof of the Collatz Conjecture by Collatz Graph", arXiv:2309.09991v2, 2024

[12] Footage in the Veritasium video series by Dr. Derek Muller Available:https://www.youtube.com/watch?v=094y1Z2wpJg

[13] Syracuse problem. *Encyclopedia of Mathematics.* URL: Available:http://encyclopediaofmath.org/index.php?title=Syracuse_problem&oldid=48940

[14] Jinqing Zhang, Xintong Zhang "Iterative Trajectories of All Odd Numbers Attain Bounded Values", *Journal of Applied Mathematics and Physics*. 2023;11:10.

[15] Mike Remer, A Comparative Analysis of the New -3(-*n*) - 1 Remer Conjecture and a Proof of the 3*n* + 1 Collatz Conjecture" *Journal of Applied Mathematics and Physics*. 2023;11:8.

[16] Ke Li. On the Change Rule of 3*x* + 1 Problem", *Journal of Applied Mathematics and Physics*. 2022;10:3.

[17] David C. Kay, "A Solution to the 3x + 1 Problem", *American Journal of Computational Mathematics*. 2023;13:2.

[18] David C. kay, "Sequences and Characteristic Zero-One Strings: Progress on the 3*x* + 1 Problem", *American Journal of Computational Mathematics*. 11:3, 2021

[19] Zuojun Hu, "The Analysis of Convergence for the 3*X* + 1 Problem and Crandall Conjecture for the a*X* + 1 Problem" *Advances in Pure Mathematics*. 2021;11:5.

[20] Maoze Wang, Yongbao Yang, Zhenxiang He, Meiyi Wang, "The Proof of the 3X + 1 Conjecture", Advances in Pure Mathematics. 2022;12:1.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/114167