
British Journal of Mathematics & Computer Science
4(21): 3023-3027, 2014

**SCIENCEDOMAIN** *international*
www.sciencedomain.org

SCIENCEDOMAIN

# A Short Elementary Proof of the Unprovability of the Collatz Conjecture

**Robin Nag**$^{*1}$

[1] *School of Digital Arts and Engineering, University of Kent, Canterbury, United Kingdom.*

***Original Research Article***

*Received: 04 July 2014*
*Accepted: 03 August 2014*
*Published: 19 August 2014*

## Abstract

Consider any positive integer $n$. If $n$ is even, halve it. If $n$ is odd, multiply it by $3$ and add $1$. This algorithm is then repeated indefinitely. It has been conjectured by Collatz that this process, which is also known as Hasse's algorithm, eventually reaches $1$. A new perspective on this problem is offered by considering Hasse's algorithm in binary representation. Some important consequences are used to establish that no proof of the Collatz conjecture exists.

*Keywords: Collatz; binary; hailstone*

2010 Mathematics Subject Classification: 11A07

## 1 Introduction

Take any positive integer $n$ and perform the following arithmetic operations. If $n$ is even, halve it to get $\frac{n}{2}$. If $n$ is odd, multiply it by $3$ and add $1$. It has been conjectured (see [1, Guy]) by Collatz and others that if this procedure (which is also known as Hasse's algorithm) is repeated indefinitely, then the number $1$ will eventually be reached. An informative background on this topic has been provided by [2, Lagarias]. We begin by considering this algorithm in its binary representation which we then use to obtain a proof that the Collatz conjecture cannot be proved to be true.

## 2 Formulation of Hasse's Algorithm

**Definition 2.1.** Without loss of generality, we assume that $n$ is an odd integer. Suppose that $n_0 = n$. Define the $k^{th}$ *step*, where $1 \le k \le N$ for some positive integer $N$, to be given by

$$n_k = \frac{3n_{k-1}+1}{2^{j_k}}, \tag{2.1}$$

where $n_k$ is an integer and $2^{j_k+1} \nmid (3n_{k-1}+1)$ for some positive integer $j_k$, provided that $n_{k-1} \neq 1$. If $n_k = 1$ is reached then we make this particular value of $k$ equal to $N$, so that the final step is given by

$$n_N = 1. \tag{2.2}$$

*Corresponding author: E-mail: rn73@kent.ac.uk*

By considering $n$ in binary, we convert this procedure into an algorithm in binary. Note that any explicit value of $n_j$, where $0 \leq j \leq N$, will be taken in its binary representation.

**Example 2.1.** *Suppose we apply Hasse's algorithm to $n = 21 = 2^4 + 2^2 + 2^0$. Now we rewrite $n$ in binary, so that $n = 10101$. We also perform all subsequent calculations for this particular example in binary. Then by considering $2n = 101010$, it follows that $3n+1 = 010101+101010+1 = 111111+1 = 1000000$. The operation of removing an even binary number's final digit which is necessarily zero is equivalent to halving the given even binary number. The last six binary digits of $3n + 1$ are all zeros. It follows immediately that if $3n + 1$ is divided by $2^6$, the resulting number is $1$. Therefore $n_1 = 1$.*

**Example 2.2.** *Suppose we apply Hasse's algorithm to $n = 5 = 2^2 + 2^0$. Now we rewrite $n$ in binary, so that $n = 101$. As before, all subsequent calculations for this particular example in binary. Then by considering $2n = 1010$, it follows that $3n + 1 = 0101 + 1010 + 1 = 1111 + 1 = 10000$. The operation of removing an even binary number's final digit which is necessarily zero is equivalent to halving the given even binary number. The last four binary digits of $3n + 1$ are all zeros. It follows immediately that if $3n + 1$ is divided by $2^4$, the resulting number is $1$. Therefore $n_1 = 1$.*

# 3 Deductions from Binary Arithmetic

In this section, we show how binary arithmetic can be used to determine some important results involving Hasse's algorithm. We make use of the following definitions.

**Definition 3.1.** Define the number of digits of any given binary number to be equivalent to the total number of digits of that binary number such that its first digit (counting from left to right) is a nonzero digit.

**Definition 3.2.** Define a *string* to be any collection of one or more consecutive adjacent binary digits which form either part or the whole of a given binary number. For example, $10$ is a string inside $1011$.

**Definition 3.3.** Define $11 \ldots 1$ to represent a string in which every digit is nonzero, so that there is at least one such digit. Define $011 \ldots 1$ to represent a string in which every digit except the first is nonzero, so that there is at least one nonzero digit.

**Definition 3.4.** Define the *end* digits of any binary number with $a$ digits to be its last $b$ digits, such that $a > b$ and that the $t^{th}$ end digit (where $1 \leq t \leq b$) is the $(b-t)^{th}$ digit from the final digit of the original binary number. We say that the original binary number with $a$ digits *ends with* its last $b$ digits, where the $t^{th}$ end digit appears before (i.e. to the left of) the $t - 1^{th}$ end digit. For instance, in Example 1.1, we say that $n$ ends with $101$, or alternatively, that its last three digits are given by $101$.

**Definition 3.5.** Define an *endstring* to represent a string which is part of a binary number, such that all digits in this string appear at the end of the afore-mentioned binary number. For example, we may select $0111$ and $111$ as possible endstrings of $n = 100111$.

The following lemmas provide us with some fundamental insights which are used to initiate our approach.

**Lemma 3.1.** *Suppose that $n \neq 1$, and that we reach $n_N = 1$. Then $n_{N-1}$ is given by $10 \ldots 1$, in which each digit differs in value from any other digit adjacent to it.*

*Proof.* Since $n_N = 1$, it is evident that $3n_{(N-1)} = 11 \ldots 1$. The desired result follows easily. ☐

**Lemma 3.2.** *Every odd positive integer $n$ can be expressed as a binary number with an endstring given by $011\dots1$, which contains $r$ nonzero digits such that $r \geq 1$. Then, if $r > 1$ we have*

$$n_{r-\kappa} = (3n_{r-(\kappa+1)} + 1)/2, \tag{3.1}$$

*where the positive integer $\kappa$ satisfies $1 \leq \kappa \leq r-1$, and if $r \geq 1$ we have*

$$n_r \leq (3n_{r-1} + 1)/4, \tag{3.2}$$

*so that a division by a number greater than or equal to $4$ is not made until the $r^{th}$ step is performed.*

*Proof.* Without loss of generality, it may be assumed that the last two binary digits of the odd integer $n$ are given by either $01$ or $11$. If the latter instance applies, then the last three binary digits of $n$ are given by either $011$ or $111$. This provides us with three different cases, for which we apply the formulation of Hasse's algorithm given in Definition 2.1.

(i) Suppose that $n$ ends with $01$. Then $2n$ contains exactly the same digits as $n$ as well as an extra zero digit at the end. It follows that the last two binary digits of $2n$ are given by $10$. Then, by considering that $3n = n + 2n$, the last two digits of $3n$ are given by $11$, and so the last two digits of $3n + 1$ are given by $00$. This means that if $n$ ends with $01$, then $n_1 \leq (3n + 1)/4$.

(ii) Suppose that the last three binary digits of $n$ are given by $011$. Then, because $2n$ contains exactly the same digits as $n$ as well as an extra zero digit at the end, the last four binary digits of $2n$ are given by $0110$. Then, by considering that $3n = n + 2n$, the last three digits of $3n$ are given by $001$, and so the last three digits of $3n + 1$ are given by $010$. This means that if $n$ ends with $011$, then the binary representation of $n_1$ ends with $01$. Therefore, by considering the preceding case, it follows immediately that $n_2 \leq (3n_1 + 1)/4$.

(iii) Suppose that the last three binary digits of $n$ are given by $111$. Then, because $2n$ contains exactly the same digits as $n$ as well as an extra zero digit at the end, the last four binary digits of $2n$ are given by $1110$. Then, by considering that $3n = n + 2n$, the last three digits of $3n$ are given by $101$, and so the last three digits of $3n + 1$ are given by $110$. This means that if $n$ ends with $111$, then the binary representation of $n_1$ ends with $11$. Therefore, it follows immediately that $n_1 = (3n + 1)/2$. Because the last three binary digits of $n$ are given by $111$, it can be assumed without loss of generality that $n$ ends with $r + 1$ digits given by $011\dots1$ (i.e. a zero digit followed by exactly $r$ digits, each of which are nonzero), such that $r \geq 3$. Note that if $n = 11\dots1$ then we may write $n = 011\dots1$. By considering that $3n = n + 2n$, the binary representation of $3n$ contains a zero digit which is derived by adding the afore-mentioned zero digit in $n$ to the corresponding nonzero digit from $2n$ and the nonzero carry digit from adding the afore-mentioned $r - 1$ pairs of nonzero digits. This zero digit in the binary representation of $3n$ is the first of $r + 1$ end digits (of $3n$) which are given by $011\dots101$, and so the binary representation of $3n + 1$ ends with $110$. It follows that the binary representation of $n_1$ ends with $r$ digits which are given by $011\dots1$. Consider the $\kappa^{th}$ step, such that $1 \leq \kappa \leq r - 1$. It follows by an easy inductive argument that the binary representation of $n_\kappa$ ends with $r - (\kappa - 1)$ digits which are given by $011\dots1$. At the $(r - 2)^{th}$ step, we have reduced this case to that of ii). Therefore, we eventually reach the $(r - 1)^{th}$ step for which the binary representation of $n_{r-1}$ ends with $01$, and so (3.1) holds. It follows that we have reduced this case to that of i), so that $n_r \leq (3n_{r-1} + 1)/4$.

The statement of the lemma now follows easily. □

It was conjectured by Catalan and proved in [3, Mihailescu (2004)] that $8$ and $9$ are the only two positive perfect powers which are also consecutive integers. We make use of this famous result in the following theorem.

**Theorem 3.3.** *Let the odd binary number $n = 11\dots1$ contain $d$ digits, such that $d > 2$. Then $n_i \neq 1$, where $i$ is any integer such that $0 \leq i \leq d$. Moreover, if $N$ exists then we have $N > d$.*

*Proof.* Since $n$ is odd, it is evident from Definition 1.1 that $n_1 \nrightarrow 1$ as $d \to \infty$. By applying Lemma 3.2, we have $n_{d-\kappa} = (3n_{d-(\kappa+1)} + 1)/2 > n_{d-(\kappa+1)}$, where the positive integer $\kappa$ satisfies $1 \leq \kappa \leq d-1$. It follows that $n_{d-1}$ is given by an integer which ends with $01$, such that $n_{d-\kappa}$ cannot exceed $n_{d-1}$ for any permissible $\kappa$.

When performing Hasse's algorithm on $n$, as executed in Definition 1.1, we find that $n_1 = 1011\ldots1$, in which the endstring $11\ldots1$ consists of $d-1$ digits which are each nonzero. By continuing with this algorithm, we find that $n_2 = 100011\ldots1$, in which the endstring $11\ldots1$ consists of $d-2$ nonzero digits.

Now if we remove the largest possible endstring $11\ldots1$ (which, by considering Lemma 3.2, consists of $d-(d-\kappa) = \kappa$ nonzero end digits) from the binary number $n_{d-\kappa}$, we assert that the remaining digits constitute the binary representation of a decimal number which is given by $3^{d-\kappa}-1$. We prove this last statement by easy induction as follows. If $\kappa = d-1$, then removing the endstring $11\ldots1$ from $n_1$ leaves behind $10$ which is the binary representation of $2 = 3^{d-(d-1)}-1$. Therefore the assertion holds for the case when $\kappa = d-1$. Let us assume as our inductive hypothesis that if we remove the endstring $11\ldots1$ consisting solely of $d-(d-\kappa) = \kappa$ nonzero digits (for which $\kappa > 1$), from the binary number $n_{d-\kappa}$, then the remaining digits constitute the binary representation of a number which is equal to $3^{d-\kappa}-1$. When performing the next step of the algorithm in Definition 1.1, so that $n_{d-(\kappa-1)} = 3n_{d-\kappa}+1$, we are multiplying the previous step's afore-mentioned remaining digits (which, by themselves alone, constitute $3^{d-\kappa}-1$) by $3$ and adding $1$ (since $2n_{d-\kappa}$ is obtained by shifting the digits of the largest possible endstring $11\ldots1$ of $n_{d-k}$) plus a carry digit obtained by adding the pair of strings given by $11\ldots1$ consisting of $d-(d-\kappa)-1 = \kappa-1$ nonzero digits. It follows that after removing the endstring $11\ldots1$ consisting of $d-(d-(\kappa-1)) = \kappa-1$ nonzero digits from $n_{d-(\kappa-1)}$, the remaining digits in this step constitute the binary representation of a decimal number given by $3(3^{d-\kappa}-1)+1+1 = 3^{d-(\kappa-1)}-1$. This completes the proof of the earlier assertion concerning the remaining digits by induction.

Suppose that $\iota$ is any integer satisfying $0 \leq \iota < d-1$. We recall that $n_\iota$ ends with $11$, and that both $n_0$ and $n_{d-\kappa}$ cannot exceed $n_{d-1}$, for any integer $\kappa$ satisfying $1 \leq \kappa \leq d-1$. Then $n_0 > 1$ and $n_\kappa > 1$, and so we must have $N > d-1$. Suppose that $n_d = 1$. It follows from Lemma 3.1 that $n_{d-1}$ must be given by $1010\ldots101$. We have already established that if we remove the largest possible endstring (which consists of $d-(d-1) = 1$ nonzero end digit) from the binary number $n_{d-1}$, then the remaining digits constitute the binary representation of a decimal number which is given by $3^{d-1}-1$. It follows that $3^{d-1}-1$ must have a binary representation given by $1010\ldots10$, i.e.

$$3^{d-1}-1 = \sum_{k=0}^{m-1} 2^{2k+1}, \tag{3.3}$$

for some positive integer $m$. By considering Lemma 3.1 or otherwise, it is easy to see that

$$\frac{3}{2}\sum_{k=0}^{m-1} 2^{2k+1} = \sum_{k=0}^{m-1}(2^{2k+1}+2^{2k}) = 2^{2m}-1. \tag{3.4}$$

It follows from (3.3) and (3.4) that

$$3^{d-1}-1 = \frac{2}{3}(2^{2m}-1), \tag{3.5}$$

i.e.

$$3^d - 2^{2m+1} = 1, \tag{3.6}$$

where $m$ and $d$ are positive integers, and $d > 2$. It is well-known that this diophantine equation has no solution, because it is a special case of Catalan's conjecture, which was proved in [3, Mihailescu (2004)]. We therefore have a contradiction, and so our assumption that $n_d = 1$ must be false. The statement of the theorem now follows immediately. □

**Corollary 3.4.** *Suppose that* $n = 11\ldots1$ *contains an unbounded number of binary digits. Then* $n_i \neq 1$ *for any nonnegative integer* $i$.

*Proof.* We apply the preceding theorem in the case as $d \to \infty$. The desired result follows. □

# 4 Conclusion

Suppose that $u$ is any positive integer. We remark that we have not discounted the possibility that the Collatz conjecture is true for any finite positive integer $n$. By taking the specific case where $n = 2^u - 1$ and by considering the previous statement about not having found any finite counterexample for the Collatz conjecture, we conclude from Corollary 3.4 that it is impossible to prove that the number $1$ can be reached if Hasse's algorithm is applied to $n$. In other words, the Collatz conjecture cannot be proved to be true. We have demonstrated how a direct elementary algorithmic approach involving binary numbers with concrete inductive arguments can be used to explain exactly why this major unsolved conjecture has resisted attempts at its proof for several decades.

# Competing Interests

The author declares that no competing interests exist.

# References

[1] Guy RK. Unsolved Problems in Number Theory. 2004;336-337. Springer, New York.

[2] Lagarias JC. The Ultimate Challenge, the 3x+1 problem. American Mathematical Society; 2010.

[3] Mihailescu P. Primary Cyclotomic Units and a Proof of Catalan's Conjecture, J. Reine angew. Math. 2004;572(572):167-195.



***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)*
*www.sciencedomain.org/review-history.php?iid=623&id=6&aid=5783*