The Remarkably Simple Structure of the 3*x* + 1 Function

by

Peter Schorer
(Hewlett-Packard Laboratories, Palo Alto, CA (ret.))
2538 Milvia St.
Berkeley, CA 94704-2611
Email: peteschorer@gmail.com
Phone: (510) 548-3827

Apr. 2, 2020

Note: The structure of the $3x + 1$ function set forth in this paper is much, *much* simpler than the structure of the function that prevails in the $3x + 1$ literature. That structure is the Collatz graph of the function. Readers can see parts of that graph in papers referenced in "Abstract" on page 2, and by papers retrieved by a Google search on "Collatz graph".

We have never received a claim of an error in this paper. A mathematician has written us stating his belief that the paper is correct, with the possible exception of some terminology.

Yet leading $3x + 1$ researchers refuse to help us get this paper published. Such help is necessary because the author is not an academic mathematician: his degree is in computer science and most of his career has been spent doing research in the computer industry. Journal editors so far refuse even to consider the paper, apparently for two reasons: the author's lack of appropriate credentials, and the great difficulty of the $3x + 1$ Problem.

Key words: $3x + 1$ Problem, $3n + 1$ Problem, Syracuse Problem, Ulam's Problem, Collatz Conjecture, computational number theory, proof of termination of programs, recursive function theory

American Mathematical Society Classification Numbers: 11Y16, 11Z05, 03D20, 68Q60

Abstract

The $3x + 1$ function is the basis for the $3x + 1$ Problem, which asks if repeated iterations of the function $C(x) = (3x + 1)/(2^a)$ always terminate in 1. Here x is an odd, positive integer, and $a = ord_2(3x + 1)$, the largest positive integer such that the denominator divides the numerator. The conjecture that the function always eventually terminates in 1 is the *3x + 1 Conjecture*.

In this paper we present a structure of the function that is very orderly despite the function's being, in a technical sense, chaotic. It is *much* simpler than the structures that appear in the literature, where they are sometimes called *directed subgraphs of the Collatz graph* or simply *Collatz graphs*. For example, see:

https://www.fq.math.ca/Scanned/40-1/andaloro.pdf,

http://go.helms-net.de/math/collatz/aboutloop/collatzgraphs.htm

The structure has two parts: one ("tuple-sets") for the function in the "forward" direction, the other ("y-trees") for the function in the "inverse" direction. The structure, along with the fact that calculations of all odd, positive integers less than at least 10^{18}, have been found, by computer test, to terminate in 1, deserves investigation as a basis for a proof of the $3x + 1$ Conjecture.

Introduction

Statement of Problem

For x an odd, positive integer, set

$$C(x) = \frac{3x+1}{2^{ord_2(3x+1)}}$$

where $ord_2(3x + 1)$ is the largest exponent of 2 such that the denominator divides the numerator. Thus, for example, C(17) = 13 ($ord_2(3(17) + 1) = 2$) , C(13) = 5 ($ord_2(3(13) + 1) = 3$), C(5) = 1 ($ord_2(3(5) + 1) = 4$). Each of these constitutes one iteration of the $3x + 1$ function. The $3x + 1$ Problem, also known as the $3n + 1$ Problem, the Syracuse Problem, Ulam's Problem, the Collatz Conjecture, Kakutani's Problem, and Hasse's Algorithm, asks if repeated iterations of C always terminate at 1. The conjecture that they do is hereafter called the *3x + 1 Conjecture*, or sometimes, in this paper, just *the Conjecture*. We call C the $3x + 1$ function; note that $C(x)$ is by definition odd.

An odd, positive integer such that repeated iterations of C terminate at 1, we call a *non-counterexample*. An odd, positive integer such that repeated iterations of C never terminate at 1, we call a *counterexample*.

Other equivalent formulations of the $3x + 1$ Problem are given in the literature; we base our formulation on the C function (following Crandall) because, as we shall see, it brings out certain structures that are not otherwise evident.

Summary of Research on the Problem

As stated in (Lagarias 1985), "The exact origin of the $3x + 1$ problem is obscure. It has circulated by word of mouth in the mathematical community for many years. The problem is traditionally credited to Lothar Collatz, at the University of Hamburg. In his student days in the 1930's, stimulated by the lectures of Edmund Landau, Oskar Petron, and Issai Schur, he became interested in number-theoretic functions. His interest in graph theory led him to the idea of representing such number-theoretic functions as directed graphs, and questions about the structure of such graphs are tied to the behavior of iterates of such functions. In the last ten years [that is, 1975-1985] the problem has forsaken its underground existence by appearing in various forms as a problem in books and journals..."

Lagarias has performed an invaluable service to the $3x + 1$ research community by publishing several annotated bibliographies relating to the Problem (these are accessible on the Internet) and by publishing his book, *The Ultimate Challenge: the 3x + 1 Problem*[1].

On the Structure of This Paper

To enhance readability, we have placed proofs of all lemmas in "Appendix A — Statement and Proof of Each Lemma" on page 16.

1. American Mathematical Society (AMS), 2010.

On Terminology

Some of our terminology differs from that in most of the $3x + 1$ literature. This is intentional on our part, to avoid confusion that might result from the different (though equivalent) definitions of the $3x + 1$ function in the literature and in this paper.

Thus, for example, according to the definition[1] of the function that is often used in the literature, the iterates of the function beginning with the argument 13 constitute the *trajectory*, or *forward orbit,* of 13, and is 13, 40, 20, 10, 5, 16, 8, 4, 2, 1. On the other hand, we represent the iterates of the definition of the function that we use (Crandall's, which we call C) by the *tuple* $<13, 5, 1>$.

In Memoriam

Several of the most important lemmas in this paper were originally conjectured by the author and then proved by the late Michael O'Neill. He made a major contribution to this research, and is sorely missed.

1. This is the one that accompanied the original statement of the $3x + 1$ Conjecture: here, each iteration is either 3x + 1 (if x is odd), or $x/2$ if x is even.

Tuple-sets: The Structure of the 3*x* + 1 Function in the "Forward" Direction

Brief Description of Tuple-sets

1. We use the definition of the 3*x* + 1 function in which all successive divisions by 2 are collapsed into a single exponent of 2 (see "Statement of Problem" on page 3). Thus, for example, the tuple <9, 7, 11> represents the fact that

9 maps to 7 in one iteration of the function, via the exponent 2, because $(3(9) + 1)/2^2 = 7$;
7 maps to 11 in one iteration of the function, via the exponent 1, because $(3(7) + 1)/2^1 = 11$.

2. We see that the sequence of exponents associated with the tuple <9, 7, 11> is {2, 1}.

3. A tuple-set T_A is the set of all finite tuples that are associated with the exponent sequence *A* (and "approximations" to *A*, but this is not important for our proofs of the 3x + 1 Conjecture). In our example, *A* = {2, 1}.

In addition to the tuple <9, 7, 11> , the tuple-set $T_A = T_{\{2, 1\}}$ contains the tuples <25, 19, 29> , <41, 31, 47>, and an infinity of others, each associated with the exponent sequence {2, 1}. (See "Lemma 1.0: the "Distance" Functions d(i, i) and d(1, i)" on page 9.)

4. Facts about tuple-sets:

An *i*-level tuple-set T_A, $i \geq 2$, contains (among other tuples, see previous step) all (*i* + 1)-element tuples that are associated with the exponent sequence *A*.

There is an infinity of tuples in each tuple-set.

The set of all tuple-sets contains tuples representing all finite iterations of the 3*x* + 1 function.

Full Description of Tuple-sets

Definitions

Iteration

An *iteration* takes an odd, positive integer, *x*, to an odd, positive integer, *y*, via one application of the 3*x* + 1 function, *C*. Thus, in one iteration *C* takes 17 to 13 because *C*(17) = 13.

Tuple

A (finite) *tuple* is a finite sequence of zero or more successive iterations of *C*, that is, $<x, C(x), C^2(x), ..., C^k(x)>$, where $k \geq 0$.

A finite tuple is the prefix of an infinite tuple. If *x* is a non-counterexample, then *x* is the first element of an infinite tuple <*x*, *y*, ..., 1, 1, 1, ... >. Of course, if *x* is a range element of *C*, then *x* can be an element other than the first in another non-counterexample tuple. If *x* is not a range element, it is a multiple-of-3 (“Lemma 4.0: Statement and Proof” on page 23), and can only be the first element of a tuple.

In the literature, a tuple (finite or infinite) is usually called a *trajectory* or an *orbit*.

If x is a counterexample, then x is the first element of an infinite tuple $<x, y, ... >$ which does not contain 1. Of course, if x is a range element of C, then x can be an element other than the first in another counterexample tuple.

A counterexample tuple must be one of two types: either there is an infinitely-repeated finite cycle of elements (none of which is 1) in the infinite tuple having the counterexample x as first element, or else there is no such cycle, but there is no 1 in the infinite tuple having the counterexample x as first element — in other words, there is no upper bound to the elements of the infinite tuple.

Exponent, Exponent Sequence

If $C(x) = y$, with $y = (3x + 1)/2^a$, we say that *a is the exponent associated with x*. In more formal language, this can be expressed as $ord_2(3x + 1) = a$. Sometimes we simply write $e(x) = a$. The sequence $A = \{a_2, a_3, ..., a_i\}$, where $a_2, a_3, ..., a_i$ are the exponents associated with x, $C(x)$, ..., $C^{(i - 1)}(x)$ respectively, is called an *exponent sequence*. We number exponents beginning with a_2 in order that the subscript corresponds to a level number in the corresponding tuple-set. See “Levels in Tuples and Tuple-sets” on page 7. For all $i \geq 2$, there are always $i – 1$ exponents in the exponent sequence associated with an i-level tuple-set.

We say that *x maps to y via* a_i if $C(x) = y$ and $ord_2(3x + 1) = a_i$. By extension, we say that x *maps to z* if z is the result of a finite sequence of iterations of C beginning with x, that is if the tuple $<x, y, ..., z>$ exists.

Tuple-set[1]

Let $A = \{a_2, a_3, ..., a_i\}$ be a finite sequence of exponents, where $i \geq 2$. The *tuple-set* T_A consists of all and only the tuples that are associated with all successive approximations to A. Thus T_A consists of all and only the following tuples. (*Note*: First elements x in different tuples are different odd, positive integers. No two tuples in a tuple-set have the same first element.)

all tuples $<x>$ such that x does not map to an odd, positive integer via a_2;

all tuples $<x, y>$ such that x maps to y via a_2 but y does not map to an odd, positive integer via a_3;

all tuples $<x, y, y'>$ such that x maps to y via a_2 and y maps to y' via a_3, but y' does not map to an odd, positive integer via a_4;

...

1. The literature contains several results that establish properties of the $3x + 1$ function that are equivalent to some of those for tuple-sets. However, the language is very different, and the definition of the $3x + 1$ function that is used is not ours, but the original one, in which each division by 2 is a separate node in the tree graph representing the function.

all tuples <$x, y, y', ..., y^{(i-3)}, y^{(i-2)}$> such that x maps to y via a_2 and y maps to y' via a_3 and ... and $y^{(i-3)}$ maps to $y^{(i-2)}$ via the exponent a_i. (The longest tuple in an i-level tuple-set has i elements.)

In other words, for each i-level exponent sequence A:

there are tuples <x> whose associated exponent sequence is a prefix of A for no exponent of A, and
there are other tuples <x, y> whose associated exponent sequence is a prefix of A for the first exponent of A, and
there are other tuples <x, y, y'> whose associated exponent sequence is a prefix of A for the first two exponents of A, and
...
there are other tuples <$x, y, z, ..., y^{(i-2)}$> whose associated exponent sequence is a prefix of A for all $i - 1$ exponents of A.

Tuples are ordered in the natural way by their first elements.

The set of first elements of all tuples in a tuple-set is the set of odd, positive integers (see proof under "The Structure of Tuple-sets" on page 9). Thus, there is a countable infinity of tuples in each tuple-set.

For each $i \geq 2$, tuple-sets are a *partition* of the set of all i-level tuples.

Levels in Tuples and Tuple-sets

• **In an *i*-level tuple-set, where $i \geq 2$, the longest tuples have *i* elements.**

Thus, <7, 11, 17, 13> is a longest tuple in a 4-level tuple-set. We say that it is a *4-level tuple*.

• **The *i*-level exponent sequence defining an *i*-level tuple-set, where $i \geq 2$, is denoted $\{a_2, a_3, ..., a_i\}$**[1],

where:

a_2 is the exponent that maps to each level 2-element in tuples having at least 2 elements;
a_3 is the exponent that maps to each level 3-element in tuples having at least 3 elements;
...
a_i is the exponent that maps to each level-i element in tuples having i-elements.

• **An *i*-level tuple-set is denoted T_A, where $A = \{a_2, a_3, ..., a_i\}$,**

Thus, our tuple <7, 11, 17, 13> is a 4-level tuple in the 4-level tuple-set denoted T_A, where $A = \{1, 1, 2\}$, because

7 maps to 11 via the exponent 1;
11 maps to 17 via the exponent 1;
17 maps to 13 via the exponent 2.

• **Let T_A be the tuple-set determined by A. Then, by definition of tuple-set, there exist *j*-level tuples in T_A, where $1 \leq j \leq i$,** that is, tuples t = <$x, y, ..., z$>, where x is the 1-level element of

1. No tuple-set has only one level, because that would mean it is associated with no exponent sequence.

t, y is the 2-level element of t, ..., and z is the j-level element of t. We sometimes speak of the set of j-level tuple-elements in T_A, where $1 \le j \le i$.

• **For $2 \le j \le i$, two tuples are said to be *consecutive at level j* if no j-level or higher-level tuple exists between them.**

Example of Tuple-set

As an example of (part of) a tuple-set: in Fig. 1, where $A = \{a_2, a_3, a_4\} = \{1, 1, 2\}$ and where we adopt the convention of orienting tuples vertically on the page, the tuple-set T_A includes:

the tuple <1>, because $e(1) = 2 \ne (a_2 = 1)$;
the tuple <3, 5> , because $e(3) = (a_2 = 1)$, but $e(5) = 4 \ne (a_3 = 1)$;
the tuple <5>, because $e(5) = 4 \ne (a_2 = 1)$;
the tuple <7, 11, 17, 13> because $e(7) = 1$ $(a_2 = 1)$ and $e(11) = 1$ $(a_3 = 1)$ and $e(17) = 2$ $(a_4 = 2)$;
etc.

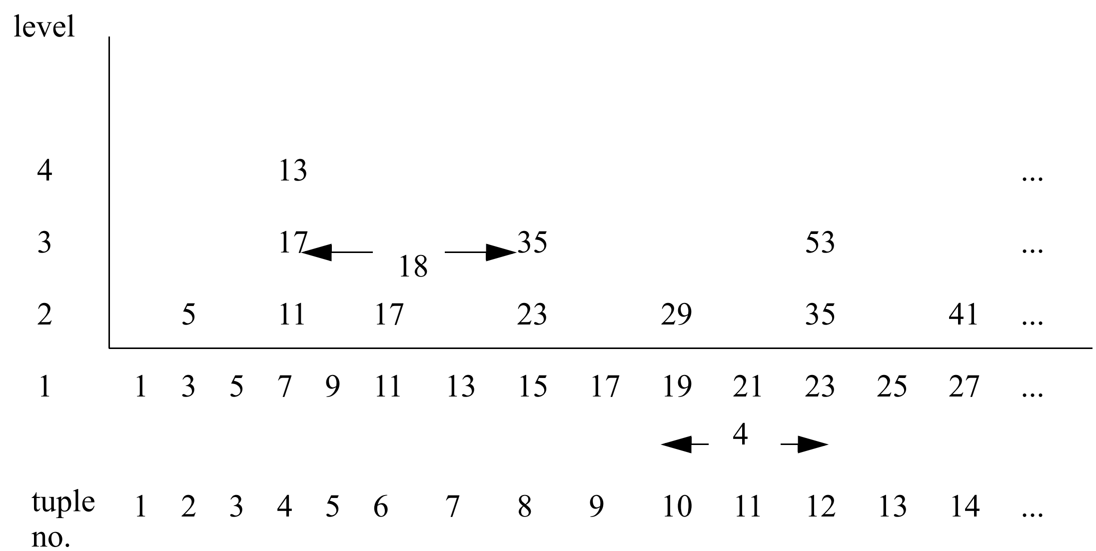

Fig. 1. Part of the tuple-set T_A associated with the sequence $A = \{1, 1, 2\}$

The number 18 between the arrows at level 3 and the number 4 between the arrows at level 1 are the values of the level 3 and level 1 distance functions, respectively, established by Lemma 1.0 (see “Lemma 1.0: the “Distance” Functions d(i, i) and d(1, i)” on page 9).

In each i-level tuple-set T_A, where $i \ge 2$, for each odd, positive integer x there exists a tuple whose first element is x. The tuple may be one-level (<x>), or 2-level (<x, y>), or ... or i-level (<$x, y, y', ..., y^{(i-3)}, y^{(i-2)}$>). Thus each tuple-set is non-empty.

Graphical Representation of the Set of All Tuple-sets

It is clear from the definition of *tuple-set* that the set of all tuple-sets can be represented by an infinitary tree in which each node is a tuple-set. We can imagine the tuple-set (which contains an infinity of tuples) extending into the page.

The Structure of Tuple-sets

It is important for the reader to understand that the structure of each tuple-set is unchanged by the presence or absence of counterexample tuples. Regardless if counterexample tuples exist or not, the set of first elements of all tuples in each tuple-set is always the same, namely, the set of odd, positive integers. *Proof:* Let x be any odd, positive integer and let $A = \{a_2, a_3, ..., a_i\}$, where $i \geq 2$, be any exponent sequence. Then there are exactly two possibilities:

(1) x maps to a y in a single iteration of the $3x + 1$ function, C, via the exponent a_2, or
(2) x does not map to a y in a single iteration of C via the exponent a_2.

But if (1) is true, then a tuple containing at least two elements, with x as the first, is in T_A; if (2) is true, then the tuple <x> is in T_{A}. There is no third possibility. □

For each tuple-set, the first element of the first tuple is 1, the first element of the second tuple is 3, the first element of the third tuple is 5, etc.

It can never be the case that, if counterexample tuples exist, then somehow there are "more" tuples in a tuple-set than if there are no counterexample tuples[1].

Furthermore, the distance functions defined in "Lemma 1.0: the "Distance" Functions d(i, i) and d(1, i)" on page 9 are the same regardless if counterexample tuples exist or not.

Extensions of Tuple-sets

Since there is a tuple-set for each finite sequence A of exponents, it follows that each tuple-set T_A has an extension via the exponent 1, and an extension via the exponent 2, and an extension via the exponent 3, ... In other words, if $A = \{a_2, a_3, ..., a_i\}$, then there is a tuple-set $T_{A'}$, where $A' = \{a_2, a_3, ..., a_i, 1\}$, and a tuple-set $T_{A''}$, where $A'' = \{a_2, a_3, ..., a_i, 2\}$, and a tuple-set $T_{A'''}$, where $A''' = \{a_2, a_3, ..., a_i, 3\}$, ...

All this is true whether or not the tuple-set T_A and/or any of its extensions contains counterexample tuples or not.

For further details on extensions of tuple-sets, see "How Tuple-sets 'Work'" and the proof that there exists an extension for each tuple-set ("Lemma 3.0 Statement and Proof") in our paper, "Are We Near a Solution to the $3x + 1$ Problem?" on occampress.com.

Lemma 1.0: the "Distance" Functions $d(i, i)$ and $d(1, i)$

(a) *Let* $A = \{a_2, a_3, ..., a_i\}$, *where* $i \geq 2$, *be a sequence of exponents, and let* $t_{(r)}, t_{(s)}$ *be tuples consecutive at level*[2] *i in* T_A. *Then d(i, i) is given by*:

1. To make this statement more precise: in no tuple-set does there ever exist a first element of a tuple, regardless how large that first element is, such that there are more tuples in that tuple-set having smaller first elements if counterexamples exist, than if counterexamples do not exist.
2. For $2 \leq j \leq i$, two tuples are consecutive at level j if no j-level or higher-level tuple exists between them (see "Levels in Tuples and Tuple-sets" on page 7.

$$d(i, i) = 2 \cdot 3^{(i-1)}$$

(b) *Let $t_{(r)}$, $t_{(s)}$ be tuples consecutive at level i in T_A. Then d(1, i) is given by:*

$$d(1, i) = 2 \cdot (2^{a_2})(2^{a_3})\ldots(2^{a_i})$$

Proof: see "Lemma 1.0: Statement and Proof" on page 16

It follows from part (a) of the Lemma that the set of all i-level elements of all i-level first tuples in all i-level tuple-sets is $\{z \mid 1 \le z < 2 \bullet 3^{i-1}\}$, where z is an odd, positive integer not divisible by 3.

Remark: Relationships similar to those described in parts (a) and (b) of the Lemma hold for successive j-level tuples, where $2 \le j < i$. The following table shows these relationships for $(i-j)$-level elements of tuples consecutive at level $(i-j)$ in an i-level tuple-set, where $0 \le j \le (i-1)$. The distances are easily proved using Lemma 1.0.

Table 1: Distances between elements of tuples consecutive at level *i*

Level	**Distance between $(i-j)$-level elements of tuples consecutive at level $(i-j)$, where $0 \le j \le (i-1)$**
i	$2 \cdot 3^{i-1}$
$i-1$	$2 \cdot 3^{i-2} \cdot 2^{a_i}$
$i-2$	$2 \cdot 3^{i-3} \cdot 2^{a_{i-1}}2^{a_i}$
$i-3$	$2 \cdot 3^{i-4} \cdot 2^{a_{i-2}}2^{a_{i-1}}2^{a_i}$
...	...
2	$2 \cdot 3 \cdot 2^{a_3}\ldots2^{a_{i-1}}2^{a_i}$
1	$2 \cdot 2^{a_2}2^{a_3}\ldots2^{a_{i-1}}\ 2^{a_i}$

Further details can be found in the section, "Remarks About the Distance Functions" in our paper, "Are We Near a Solution to the 3x + 1 Problem?", on occampress.com.

Lemma 2.0 Counterexample tuples in tuple-sets if counterexamples exist
Assume a counterexample exists. Then for all i ≥ 2, each i-level tuple-set contains an infinity of i-level counterexample tuples and an infinity of i-level non-counterexample tuples.

Proof: See "Lemma 2.0: Statement and Proof" on page 21.

y-Trees: The Structure of the 3x + 1 Function in the "Backward", or Inverse, Direction

Definition of the *y*-Tree

Let y be a range element of the $3x + 1$ function. Then y and the set of all odd, positive integers that map to y in one or more iterations of the $3x + 1$ function, is a tree called the *y-tree*.

Properties of the y-Tree

• Each root y of a y-tree is mapped to, in one iteration of the 1-tree, either by all even or by all odd exponents.

Proof:

1. Assume the root y is mapped to by all even exponents. Then for each exponent $2k$, there exists an x such that:

$$\frac{3x+1}{2^{2k}} = y$$

2. Multiply numerator and denominator by 2^2. Then we have

$$\frac{2^2}{2^2}\frac{(3x+1)}{2^{2k}} = y$$

or

$$\frac{(3(4x+1)+1)}{2^{2k+2}} = y$$

3. Our exponent increases by 2, yielding the next even exponent. A similar argument applies if we assume the root is mapped to by all odd exponents. □

Hence an infinity of odd, positive integers, which we call a "spiral", maps to y.

• Level 1 of a y-tree is the set of all odd, positive integers that map to y in one iteration of the $3x + 1$ function;
Level 2 of a y-tree is the set of all odd, positive integers that map to all elements of Level 1 in one iteration of the $3x + 1$ function;
Level 3 of a y-tree is the set of all odd, positive integers that map to all elements of Level 2 in one iteration of the 3x + 1 function;
etc.

• Let x be an element of a "spiral". Then the next larger element of the "spiral" is $4x + 1$.
Proof: See above proof regarding parity of exponents.

• The successive elements of a “spiral” are mapped to in accordance with a rule that can be expressed as ... 2, 1, 3, 2, 1, 3, ..., where “2” means “is mapped to by all even exponents”, “1” means “is mapped to by all odd exponents”, and “3” means “is not mapped to because element is a multiple-of-3, hence not a range element (“Lemma 4.0: Statement and Proof” on page 23”). The proof in brief is the following:

$$\frac{3w+1}{2^2} = x \rightarrow \frac{3(2w+1)+1}{2^1} = 4x+1$$

The reader can substitute the left-hand side of the left-hand equation for x in the right-hand side of the right-hand equation, and work through the algebra to see that the two equations in fact hold.□

The repetition of a multiple-of-3 every third successive element of a “spiral” can be seen from the following. Let $3m$ be a multiple-of-3. Then, by the $4x + 1$ rule described above in this list of properties, we have, for the third successive element after the $3m$ element:

$$4(4(4(3m)+1)+1)+1 = (64)(3)(m)+21$$

Each of the two terms on the right-hand side of the equation are multiples of 3, and so we have our result.□

Finally, we must prove that the next successive element of a “spiral” following a multiple-of-3 is an element that is mapped to by all even exponents.

Let $3m$ be a multiple-of-3. The next successive “spiral” element is $4(3m) + 1$. We ask if there exists a w such that

$$\frac{3w+1}{2^2} = 4(3m)+1$$

Multiplying through by 2^2 we get

$$3w = 2^2(4)(3m) + 2^2 - 1 = 3U + 3$$

Hence w exists. It is equal to $U + 1$.□

Because each “spiral” contains an infinity of range elements, each y-tree is infinitely deep.

• There is one and only one possible y-tree for each y.

Proof: Otherwise, it would be possible for a range element to be mapped to, in one iteration, by more than one “spiral”. But that is impossible by the arithmetic governing the $3x + 1$ function.□

Can y-Trees and Tuple-sets Be Merged?

The answer is yes. Pick any node x in a y-tree. Then the sequence of nodes mapped to y from x is a tuple in a tuple-set.

The 1-Tree

The 1-tree is a y-tree, where $y = 1$. The 1-tree contains all and only the odd, positive integers that map to 1. There is a possibility that the 1-tree can be the basis for a proof of the $3x + 1$ Conjecture. See "Lemma 3.0: Statement and Proof" on page 21 for more details on the 1-tree.

References

Lagarias, J., (1985), "The $3x + 1$ Problem and Its Generalizations", *American Mathematical Monthly*, **93**, 3-23.

Wirsching, Günther J.. *The Dynamical System Generated by the 3n + 1 Function*, Springer-Verlag, Berlin, Germany, 1998.

Appendix A — Statement and Proof of Each Lemma

Lemma 1.0: Statement and Proof

Definition: let T_A be an i-level tuple-set, where $i \geq 2$. Let $t(r)$, $t(s)$ denote tuples consecutive at level i, with $r < s$ in the natural ordering of tuples by first elements. Let $t(r)(h)$, $t(s)(h)$ denote the elements of $t(r)$, $t(s)$ at level h, where $1 \leq h \leq i$. Then we call $|t_{(s)(h)} - t_{(r)(h)}|$ the *distance* between $t(r)$ and $t(s)$ at level h. We denote this distance by $d(h, i)$ and call d the *distance functions* (one function for each h).

Lemma 1.0

(a) Let $A = \{a_2, a_3, ..., a_i\}$, where $i \geq 2$, *be a sequence of exponents, and let* $t_{(r)}$, $t_{(s)}$ *be tuples consecutive at level i in* T_A. *Then d(i, i) is given by*:

$$d(i, i) = 2 \cdot 3^{(i-1)}$$

(b) *Let* $t_{(r)}$, $t_{(s)}$ *be tuples consecutive at level i in* T_A. *Then d(1, i) is given by:*

$$d(1, i) = 2 \cdot (2^{a_2})(2^{a_3})\dots(2^{a_i})$$

Thus, in “Fig. 1. Part of the tuple-set TA associated with the sequence A = {1, 1, 2}” on page 8, the distance $d(3, 3)$ between $t_{8(3)} = 35$ and $t_{4(3)} = 17$ is $2 \cdot 3^{(3-1)} = 18$. The distance $d(1, 2)$ between $t_{12(1)} = 23$ and $t_{10(1)} = 19$ is $2 \cdot 2^1 = 4$.

Proof:

The proof is by induction.

Proof of Basis Step for Parts (a) and (b) of Lemma 1.0:

Let $t_{(r)}$ and $t_{(s)}$ be the first and second 2-level tuples, in the standard linear ordering of tuples based on their first elements, that are consecutive at level $i = 2$ in the 2-level tuple-set T_A, where $A = \{a_2\}$. (See Fig. 2 (1).)

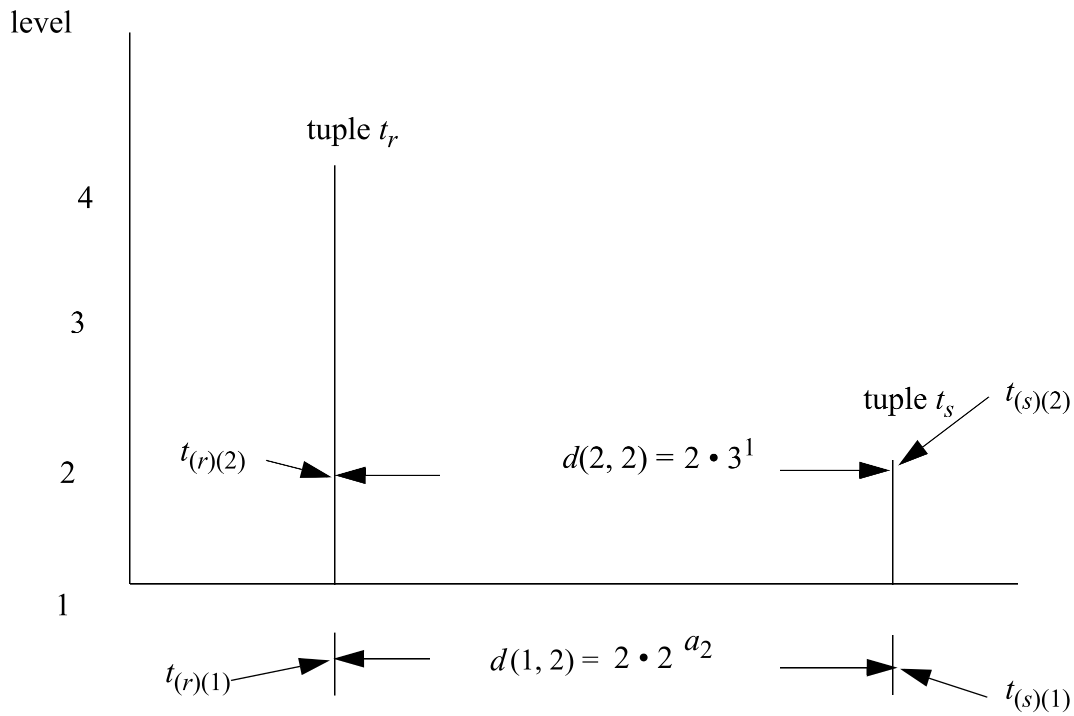

Fig. 2 (1). Illustration for proof of Basis Step of Lemma 1.0.

Then we have:

$$\frac{3t_{(r)(1)} + 1}{2^{a_2}} = t_{(r)(2)} \tag{1.1}$$

and since, by definition of $d(1, 2)$,

$$t_{(s)(1)} = t_{(r)(1)} + d(1, 2)$$

we have:

$$\frac{3(t_{(r)(1)} + d(1, 2)) + 1}{2^{a_2}} = t_{(s)(2)} \tag{1.2}$$

Therefore, since, by definition of $d(i, i)$,

$$t_{(r)(2)} + d(2, 2) = t_{(s)(2)}$$

we can write, from (1.1) and (1.2):

$$\frac{3t_{(r)(1)}+1}{2^{a_2}} + d(2, 2) = \frac{3(t_{(r)(1)} + d(1, 2)) + 1}{2^{a_2}}$$

By elementary algebra, this yields:

$$2^{a_2}d(2, 2) = 3 \cdot d(1, 2)$$

Now $d(2, 2)$ must be even, since it is the difference of two odd, positive integers, and furthermore, by definition of tuples consecutive at level i, it must be the smallest such even number, whence it follows that $d(2, 2)$ must = 3 • 2, and necessarily

$d(1, 2) = 2 \bullet 2^{a_2}$

A similar argument establishes that $d(2, 2)$ and $d(1,2)$ have the above values for every other pair of tuples consecutive at level 2.

Thus we have our proof of the Basis Step for parts (a) and (b) of Lemma 1.0.

Proof of Induction Step for Parts (a) and (b) of Lemma 1.0

Assume the Lemma is true for all levels j, $2 \le j \le i$ and that T_A is an i-level tuple-set, where A = $\{a_2, a_3, ..., a_i\}$.

Let $t_{(r)}$ and $t_{(s)}$ be tuples consecutive at level i, and let $t_{(r)}$ and $t_{(f)}$ be tuples consecutive at level i +1. (See Fig. 2 (2).)

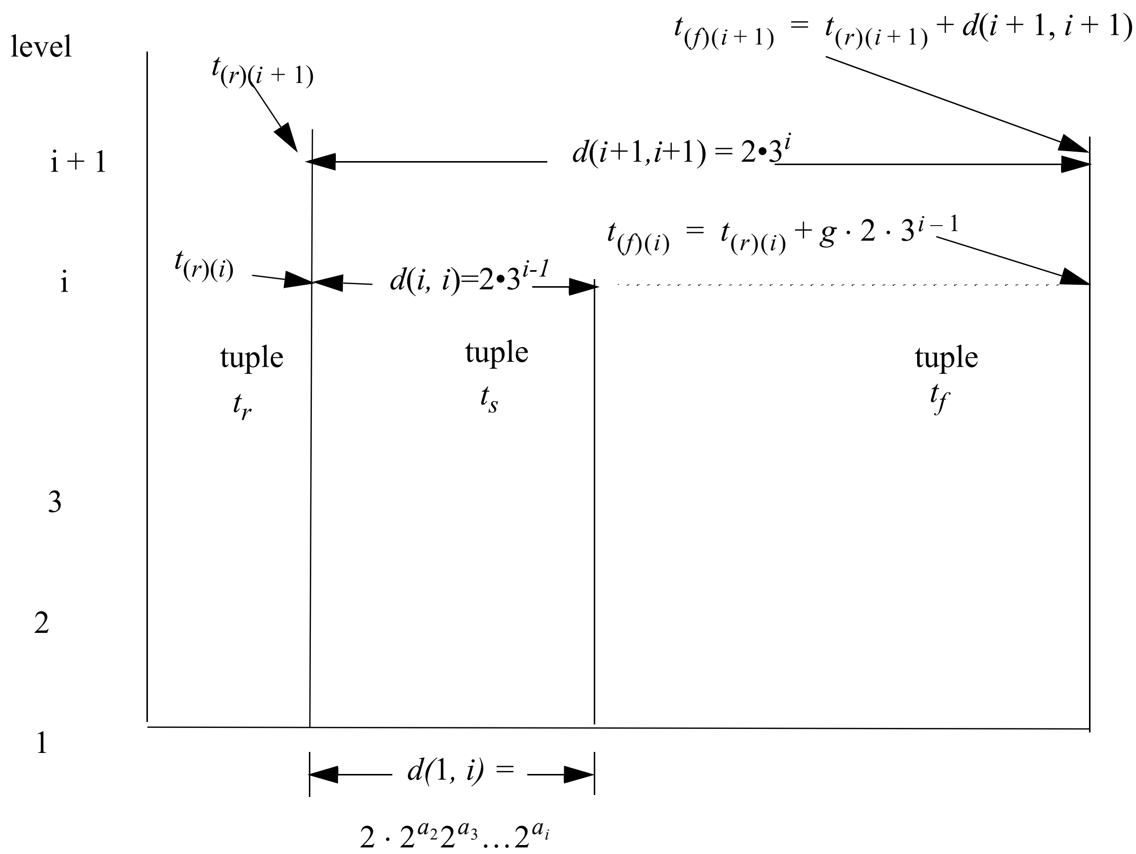

Fig. 2 (2). Illustration for proof of Induction Step of Lemma 1.0.

Then we have:

$$\frac{3t_{(r)(i)} + 1}{2^{a_{i+1}}} = t_{(r)(i+1)}$$

and since, by definition of $d(i, i)$,

$$t_{(f)(i)} = t_{(r)(i)} + g \cdot d(i, i)$$

for some $g \geq 1$, we have:

$$\frac{3(t_{(r)(i)} + g \cdot d(i, i)) + 1}{2^{a_{i+1}}} = t_{(f)(i+1)}$$

Thus, since

$$t_{(r)(i+1)} + d(i+1, i+1) = t_{(f)(i+1)}$$

we can write:

$$\frac{3t_{t_{(r)(i)}} + 1}{2^{a_{i+1}}} + d(i+1, i+1) = \frac{3(t_{t_{(r)(i)}} + gd(i, i)) + 1}{2^{a_{i+1}}}$$

This yields, by elementary algebra:

$$2^{a_{(i+1)}} d(i+1, i+1) = 3 \cdot gd(i, i)$$

As in the proof of the Basis Step, $d(i+1, i+1)$ must be even, since it is the difference of two odd, positive integers, and furthermore, by definition of tuples consecutive at level $i+1$, it must be the smallest such even number. Thus $d(i+1, i+1) = 3 \bullet d(i, i)$, and

$$g \cdot d(i, i) = 2^{a_{i+1}} d(i, i) \quad .$$

Hence

$$g = 2^{a_{i+1}}$$

Now g is the number of tuples consecutive at level i that must be "traversed" to get from $t_{(r)}$ to $t_{(f)}$. By inductive hypothesis, d(1, i) for *each* pair of these tuples is:

$$d(1, i) = 2 \cdot 2^{a_2} \cdot 2^{a_3} \cdot \ \ldots \cdot 2^{a_i}$$

hence, since

$$g = 2^{a_{i+1}}$$

we have

$$d(1, i+1) = d(1, i) \cdot 2^{a_{i+1}} .$$

A similar argument establishes that $d(i+1, i+1)$ and $d(1, i+1)$ have the above values for every pair of tuples consecutive at level $i+1$.

Thus we have our proof of the Induction Step for parts (a) and (b) of Lemma 1.0. The proof of Lemma 1.0 is completed. □

Lemma 2.0: Statement and Proof

Assume a counterexample exists. Then for all $i \geq 2$, each i-level tuple-set contains an infinity of i-level counterexample tuples and an infinity of i-level non-counterexample tuples.

Proof:

1. Assume a counterexample exists. Then:

There is a countable infinity of non-counterexample range elements.
Proof: Each non-counterexample maps to a range element, by definition of *range element*.
Each range element is mapped to by an infinity of elements ("Lemma 6.0: Statement and Proof" on page 25). A countable infinity of these are range elements (proof of "Lemma 7.0: Statement and Proof" on page 27).

There is a countable infinity of counterexample range elements.
Proof: same as for non-counterexample case.

2. For each finite exponent sequence A, and for each range element y, non-counterexample or counterexample, there is an x that maps to y via A possibly followed by a buffer exponent ("Lemma 7.0: Statement and Proof" on page 27). The presence of the buffer exponent does not change the fact that x is the first element of a tuple associated with the exponent sequence A. □

Lemma 3.0: Statement and Proof

There is one and only one possib1e 1-tree, whether or not counterexamples exist.

Short Proof:

1. "Once a non-counterexample, always a non-counterexample." *Proof*: the proof is a generalization of our canonical example, 13: "13 is a non-counterexample today; if the $3x + 1$ Conjecture is proved true tomorrow, it will be a non-counterexample; and if the Conjecture is proved *false* tomorrow it will *still* be a non-counterexample". Similar facts follow for all non-counterexamples by definition of the $3x + 1$ function (no odd, positive integer can map to two or more different values in one iteration of the function.) □

2. "Once a non-counterexample, always a non-counterexample" can be expressed as the statement of Lemma 3.0. □

Longer Proof:

Proof of "There is one and only one possible 1-tree..."

The 1-tree =

(Level 1 = {odd, positive integers y | y maps to 1 in one iteration of the $3x + 1$ function}[1]) ∪
(Level 2 = {odd, positive integers y | y maps to an element of Level 1 in one iteration of the $3x + 1$ function}) ∪
(Level 3 = ({odd, positive integers y | y maps to an element of Level 2 in one iteration of the $3x + 1$ function} ∪
...

Since 1 is a range element of the $3x + 1$ function, it is the root of a y-tree (in this case, the 1-tree). Each y-tree has several basic, well-defined properties. (For full details, and references to the elementary proofs, see "Properties of the y-Tree" on page 12 and "y-Trees: The Structure of the 3x + 1 Function in the "Backward", or Inverse, Direction" on page 12):

Each y is mapped to by an infinity of odd, positive integers in one iteration of the $3x + 1$ function. We call this infinity of odd, positive integers, a "spiral".

If x is an element of a "spiral", then $4x + 1$ is the next larger element.

Each "spiral" contains an infinity of range elements, and an infinity of multiples of 3, which are not range elements because they are not mapped to by any odd, positive integer.

Each "spiral" element maps to y (in one iteration of the $3x + 1$ function), by either all odd exponents, or by all even exponents.

The sequence of these types of "spiral" elements is given by a rule that can be expressed as ... 2, 1, 3, 2, 1, 3, ..., where "2" means "is mapped to by all even exponents", "1" means "is mapped to by all odd exponents", and "3" means "is not mapped to because element is a multiple-of-3, hence not a range element".

Because of the infinity of range elements in each "spiral", it is clear that the structure of each y-tree is the result of an infinitely recursive process. Thus each y-tree is infinitely deep.

Proof of "...whether or not counterexamples exist

If an odd, positive integer x maps to 1 (that is, if x is a non-counterexample, hence an element of the 1-tree), then it maps to 1 regardless if counterexamples exist or not. Informally, we say, "Once a non-counterexample, always a non-counterexample." Thus, for example,

13 maps to 1 today;
if the $3x + 1$ Conjecture is proved true tomorrow it will still map to 1;
if the $3x + 1$ Conjecture is proved *false* tomorrow it will *still* map to 1.

If it were not the case that "Once a non-counterexample, always a non-counterexample", some odd, positive integers could map to two different odd, positive integers, contrary to the definition of the $3x + 1$ function. □

Remark 1

Some readers claim that the Lemma is trivial, "unnecessary". But this claim is based on a false assumption. These readers assume (correctly) that 1 is mapped to by all exponents of only one parity, but they assume (incorrectly) that there are nevertheless two possibilities: (1) the 1-tree contains all odd, positive integers, or (2) the 1-tree contains only a proper subset of the odd, positive integers.

---

1. This set is $S = \{1, 5, 21, 85, 341, ...\}$.

However, that implies that a given range element, although it is mapped to by one and only one exponent (in one iteration of the $3x + 1$ function), nevertheless can be mapped to by two different odd, positive integers via that one exponent! But that is impossible, given the definition of the $3x + 1$ function.

In actuality, we know that 1 is mapped to by all even exponents. There is no possibility that it might be mapped to by any odd exponents. Furthermore, for each range element y (and 1 is a range element) and each (even) exponent, y is mapped to by one and only odd, positive integer in one iteration of the $3x + 1$ function.

Hence there is one and only one possible 1-tree, whether or not counterexamples exist.

Remark 2

The Lemma passes the $3x - 1$ Test. The reason is that the Lemma asserts that there is one and only one possible 1-tree, whether or not counterexamples exist. At the time of this writing, no counterexample to the 3x + 1 Conjecture is known, even though all consecutive odd, positive integers between 1 and at least $10^{18} - 1$ have been found, by computer test[1], to be non-counterexamples. But a counterexample to the $3x - 1$ Conjecture is known (the smallest is 5), and so it is emphatically not true that there is one and only one possible 1-tree for the $3x - 1$ function, whether or not counterexamples exist. If no counterexamples to the $3x - 1$ Conjecture existed, the 1-tree for the $3x - 1$ function would certainly be different than the existing one.

Remark 3

The Lemma statement is, of course, very counter-intuitive. Even we who first stated it, and then proved it, found ourselves spending time trying to understand how it could be true.

But it is true, as the reader can check by going over the proof.

Lemma 4.0: Statement and Proof

No multiple of 3 is a range element.

Proof :

If

$$\frac{3x + 1}{2^a} = 3m$$

then $1 \equiv 0 \bmod 3$, which is false. □

Lemma 5.0: Statement and Proof

Each odd, positive integer (except a multiple of 3) is mapped to by a multiple of 3 in one iteration of the 3x + 1 function.

---

1. See results of tests performed by Tomás Oliveira e Silva, www.ieeta.pt/~tos/3x+1/html. All consecutive odd, positive integers less than $20 \cdot 2^{58} \approx 5.76 \cdot 10^{18}$, which is greater than $3.33 \cdot 10^{16} \approx 2 \cdot 3^{(35 - 1)}$, have been tested and found to be non-counterexamples.

Proof:

Since the domain of the $3x + 1$ function is the odd, positive integers, the only relevant generators are $3(2k + 1)$, $k \ge 0$. We show that, for each odd, positive integer y not a multiple of 3, there exists a k and an a such that

$$y = \frac{(3(3(2k+1))+1)}{2^a} \quad , \tag{11.1}$$

where a is necessarily the largest such a, since y is assumed odd.

Rewriting (11.1), we have:

$$y2^{a-1} - 5 = 9k \ . \tag{11.2}$$

Without loss of generality, we can let $y \equiv r \bmod 18$, where r is one of 1, 5, 7, 11, 13, or 17 (since y is odd and not a multiple of 3, these values of r cover all possibilities mod 18). Or, in other words, for some q, r, $y = 18q + r$. Then, from (11.2) we can write:

$$18(2^{a-1})q + (2^{a-1})r - 5 = 9k \ . \tag{11.3}$$

Since the first term on the left-hand side is a multiple of 9, $(2^{a-1})r - 5$ must also be if the equation is to hold. We can thus construct the following table. (Certain larger a also serve equally well, but those given suffice for purposes of this proof.)

Table 2: Values of *r*, *a*, for Proof of Lemma

r	a	$(2^{a-1})r - 5$
1	6	27
5	1	0
7	2	9
11	5	171
13	4	99
17	3	63

Given q and r (hence y), we can use r to look up a in the table, and then solve (11.3) for integral k, thus producing the multiple of 3 that maps to y in one iteration of the $3x + 1$ function. □

Lemma 6.0: Statement and Proof

(a) Each range element y is mapped to, in one iteration of the 3x + 1 function, by every exponent of one parity only. Furthermore,

(b) For each of the two parities, there exists a range element that is mapped to by every exponent of that parity.

Proof of part (a):

Steps 1 and 2 are slightly edited versions of proofs by Jonathan Kilgallin and Alex Godofsky. Any errors are entirely ours. Step 3 is a slightly edited version of a proof by Michael Klipper. Any errors are entirely ours.

1. We first show that if y is mapped to by the exponent a, then y is mapped to by every exponent greater than a that is of the same parity as a.

Let y be a range element, and let x map to y via the exponent a. Then

$$y = \frac{3x + 1}{2^a}$$

We wish to show that there exists an x' such that x' maps to y via the exponent 2^{a+2}. That is, we wish to show that there exists an x' such that

$$y = \frac{3x' + 1}{2^{a+2}}$$

Rewriting, this gives

$$x' = \frac{2^{a+2}y - 1}{3}$$

Substituting for y yields

$$x' = \frac{2^{a+2}\left(\frac{3x+1}{2^a}\right) - 1}{3}$$

Simplifying, this gives $x' = 4x + 1$. Since x is an odd, positive integer, clearly x' is as well.

Thus, by induction, if y is mapped to via the exponent a , it is mapped to by every exponent greater than a of the same parity. □

2. Next we show that if y is mapped to by the exponent a which is greater than 2, then it is mapped to by every exponent less than a that is of the same parity as a.

Let y be a range element, and let x map to y via the exponent a where $a > 2$. Then

$$y = \frac{3x + 1}{2^a}$$

We wish to show that there exists an x' such that x' maps to y via the exponent 2^{a-2}. That is, we wish to show that there exists an x' such that

$$y = \frac{3x' + 1}{2^{a-2}}$$

Rewriting, this gives

$$x' = \frac{2^{a-2}y - 1}{3}$$

Substituting for y yields

$$x' = \frac{2^{a-2}\left(\frac{3x+1}{2^a}\right) - 1}{3}$$

Simplifying yields

$$x' = \frac{x - 1}{4}$$

3. We must now show that $x' = (x - 1)/4$ is an odd, positive integer. This means we must show that $(x - 1) = 4(2k + 1)$ for some $k \geq 0$, or that $(x - 1) = 8k + 4$, hence that $x = 8k + 5$. Thus, we must prove $x \equiv 5 \bmod 8$.

We know that x maps to y via a, where $a \geq 3$. Thus, $y = (3x + 1)/2^a$, so $2^a y = 3x + 1$. Because $a \geq 3$, $2^a y$ is a multiple of 8. Thus, $(3x + 1) \equiv 0 \bmod 8$, and $3x \equiv 7 \bmod 8$. This readily implies $x \equiv 5 \bmod 8$.

4. Thus, by induction, if y is mapped to via the exponent a, where $a > 2$, then it is mapped to by every exponent less than a of the same parity. □

Proof of part (b):

We now show that for each of the two parities there exists a range element that is mapped to by every exponent of that parity.

1. Fix a range element y, and suppose that x maps to y via the exponent a. Now a is either even or odd, hence $a = 2n + h$, where h is either 0 or 1. Since $y = (3x + 1)/2^a$, it follows that $(2^a)y = 3x+1$. Reduce the equation mod 3, and we get $(2^h)y \equiv 1 \bmod 3$, by the following reasoning: $(2^a)y \equiv 1 \bmod 3$ implies $(2^{2n+h})y \equiv 1 \bmod 3$ implies $2^{2n}\, 2^h y \equiv 1 \bmod 3$ implies $2^h y \equiv 1 \pmod 3$ because $2^{2n} = 4^n \equiv 1 \bmod 3$.

2. Since y is fixed, either $y \equiv 1$ or $y \equiv 2 \bmod 3$. (We know that y, a range element, is not a multiple of 3 by “Lemma 4.0: Statement and Proof” on page 23). If $y \equiv 1 \bmod 3$, then we have $2^h(1) \equiv 1 \bmod 3$, which implies that h must be 0. If $y \equiv 2 \bmod 3$, then we have $(2^h)(2) \equiv 1 \bmod 3$, implying that h must be 1. □

Lemma 7.0: Statement and Proof

Let y be a range element of the 3x + 1 function. Then for each finite exponent sequence A, there exists an x that maps to y via A possibly followed by a “buffer” exponent. (For example, if y is mapped to by even exponents, and our exponent sequence A ends with an odd exponent, then there must be an even “buffer” exponent following A, and similarly if y is mapped to by odd exponents and A ends with an even exponent. However, there are other cases in which a “buffer” exponent is required.)

Proof:

1. Each range element y is mapped to by all exponents of one parity (“Lemma 6.0: Statement and Proof” on page 25).

2. Each range element y is mapped to by a multiple of 3 (“Lemma 5.0: Statement and Proof” on page 23).

Each range element is mapped to by an infinity of range elements (“Lemma 5.0: Statement and Proof” on page 23).

3. Let y be a range element and let $S = \{s_1, s_2, s_3, ... \}$ be the set of all odd, positive integers that map to y in one iteration of the $3x + 1$ function. In other words, S is the set of all elements in a "spiral". Furthermore, let the s_i be in increasing order of magnitude. It is easily shown that $s_{i+1} = 4s_i + 1$.

(In Fig. 18, $y = 13$, $S = \{17, 69, 277, 1109, ... \}$

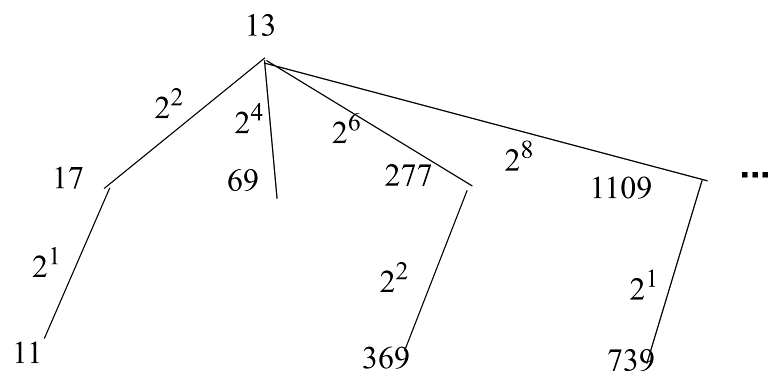

Fig. 18

(*Note*: for a graphical representation of part of the tree having 1 as its root instead of 13, see "Recursive "Spiral"s: The Structure of the 3x + 1 Function in the "Backward", or Inverse, Direction" in our paper, "Are We Near a Solution to the $3x$ + 1 Problem?", on occampress.com.)

4. If s_i is a multiple of 3, then $4s_i + 1$ is mapped to, in one iteration of the $3x + 1$ function, by all exponents of even parity.

To prove this, we need only show that x is an integer in the equation

$$4(3u) + 1 = \frac{3x + 1}{2^2}$$

Multiplying through by 2^2 and collecting terms we get

$$(48u) + 4 = 3x + 1$$

and clearly x is an integer.

5. If s_j is mapped to by all even exponents, then $4s_j + 1$ is mapped to, in one iteration of the $3x + 1$ function, by all exponents of odd parity.

(The proof is by an algebraic argument similar to that in step 4.)

6. If s_k is mapped to by all odd exponents, then $4s_k + 1$ is a multiple of 3.

(The proof is by an algebraic argument similar to that in step 4.)

7. The Lemma follows by an inductive argument that we now describe.

Let y be a range element. It is mapped to by all exponents of one parity. Thus it is mapped to by an infinite sequence of odd, positive integers. As a consequence of steps 1 through 6, we can represent an infinite sub-sequence of the sequence by

...3, 2, 1, 3, 2, 1, ...

where

"3" means "this odd, positive integer is a multiple of 3 and therefore is not mapped to by any odd, positive integer";

"2" means "this odd, positive integer is mapped to by all even exponents";

"1" means "this odd, positive integer is mapped to by all odd exponents".

Each type "2" and type "1" odd, positive integer is mapped to by all exponents of one parity. Thus it is mapped to by an infinite sequence of odd, positive integers. We can represent an infinite sub-sequence of the sequence by

...3, 2, 1, 3, 2, 1, ...

where each integer has the same meaning as above.

Temporarily ignoring the case in which a buffer exponent is needed, it should now be clear that, for each range element y, and for each finite sequence of exponents B, we can find a finite path down through the infinitary tree we have just established, starting at the root y. The path will end in an odd, positive integer x. Let A denote the path B taken in reverse order. Then we have our result for the non-buffer-exponent case. The buffer-exponent case follows from the fact that the buffer exponent is one among an infinity of exponents of one parity. Thus y is mapped to by an infinite sequence of odd, positive integers. We then simply apply the above argument.. □