Discrete Mathematics and Theoretical Computer Science **5**, 2002, 47–54

$3x+1$ *Minus the* $+$

Kenneth G. Monks

Department of Mathematics, University of Scranton, Scranton, PA, 18510 USA
email: monks@scranton.edu

received Aug 10, 2001, accepted April 9, 2002.

We use Conway's *Fractran* language to derive a function $R:\mathbb{Z}^+\to\mathbb{Z}^+$ of the form

$$R(n)=r_i n \text{ if } n\equiv i \bmod d$$

where d is a positive integer, $0\le i<d$ and $r_0,r_1,\dots r_{d-1}$ are rational numbers, such that the famous $3x+1$ conjecture holds if and only if the R-orbit of 2^n contains 2 for all positive integers n. We then show that the R-orbit of an arbitrary positive integer is a constant multiple of an orbit that contains a power of 2. Finally we apply our main result to show that any cycle $\{x_0,\dots,x_{m-1}\}$ of positive integers for the $3x+1$ function must satisfy

$$\sum_{i\in\mathcal{E}}\left\lfloor\frac{x_i}{2}\right\rfloor=\sum_{i\in\mathcal{O}}\left\lfloor\frac{x_i}{2}\right\rfloor+k.$$

where $\mathcal{O}=\{i:x_i \text{ is odd}\}$, $\mathcal{E}=\{i:x_i \text{ is even}\}$, and $k=|\mathcal{O}|$. The method used illustrates a general mechanism for deriving mathematical results about the iterative dynamics of arbitrary integer functions from *Fractran* algorithms.

Keywords: Collatz conjecture, $3x+1$ problem, *Fractran*, discrete dynamical systems

1 Introduction and Main Results

The famous $3x+1$ conjecture (cf. [3],[4]) states that for every $n\in\mathbb{Z}^+$ there exists $k\in\mathbb{Z}^+$ such that $T^k(n)=1$ where

$$T(n)=\begin{cases}\frac{1}{2}n & \text{if } n \text{ is even}\\ \frac{3}{2}n+\frac{1}{2} & \text{if } n \text{ is odd.}\end{cases}$$

and $T^k=\underbrace{T\circ T\circ\cdots\circ T}_{k}$ denotes the k-fold composition of T with itself. If we let $T_0(x)=\frac{x}{2}$ and $T_1(x)=\frac{3}{2}x+\frac{1}{2}$, then for any n and k, $T^k(n)=T_{v_{k-1}}\circ T_{v_{k-2}}\circ\cdots\circ T_{v_0}(n)$ for some $v_0,\dots v_{k-1}\in\{0,1\}$ and $v_i\equiv T^i(n) \bmod 2$. Several authors (cf. [3]) have given explicit formulas for this composition, e.g.

$$T_{v_{k-1}}\circ T_{v_{k-2}}\circ\cdots\circ T_{v_0}(n)=\frac{3^m}{2^k}n+\sum_{i=0}^{k-1}v_i\frac{3^{v_{i+1}+\cdots+v_{k-1}}}{2^{k-i}} \text{ where } m=\sum_{i=0}^{k-1}v_i.$$

Compare this somewhat unwieldy expression with the much simpler one

$$R_{v_{k-1}} \circ R_{v_{k-2}} \circ \cdots \circ R_{v_0}(n) = \frac{3^m}{2^k} n$$

when $R_0(n) = \frac{1}{2}n$ and $R_1(n) = \frac{3}{2}n$. With this example in mind, it is natural to ask if there is some function of the form

$$R(n) = \begin{cases} r_0 n & \text{if } n \equiv 0 \bmod d \\ r_1 n & \text{if } n \equiv 1 \bmod d \\ \vdots & \vdots \\ r_{d-1} n & \text{if } n \equiv d-1 \bmod d \end{cases} \tag{1.1}$$

where $r_1, \ldots, r_{d-1}$ are rational numbers and $d \geq 2$ such that knowledge of certain R-orbits would settle the $3x+1$ problem, i.e. is there an addition-free variant of the $3x+1$ function whose dynamics encode the conjecture? We answer this question in the affirmative with the following result

Theorem 1 *There are infinitely many functions R of the form (1.1) having the property that the* $3x+1$ *conjecture is true if and only if for all positive integers n the R-orbit of* 2^n *contains* 2*. In particular,*

$$R(n) = \begin{cases} \frac{1}{11}n & \textit{if } 11 \mid n \\ \frac{136}{15}n & \textit{if } 15 \mid n \textit{ and NOTA} \\ \frac{5}{17}n & \textit{if } 17 \mid n \textit{ and NOTA} \\ \frac{4}{5}n & \textit{if } 5 \mid n \textit{ and NOTA} \\ \frac{26}{21}n & \textit{if } 21 \mid n \textit{ and NOTA} \\ \frac{7}{13}n & \textit{if } 13 \mid n \textit{ and NOTA} \\ \frac{1}{7}n & \textit{if } 7 \mid n \textit{ and NOTA} \\ \frac{33}{4}n & \textit{if } 4 \mid n \textit{ and NOTA} \\ \frac{5}{2}n & \textit{if } 2 \mid n \textit{ and NOTA} \\ 7n & \textit{otherwise} \end{cases} \tag{1.2}$$

(where NOTA means "None of the Above" conditions hold) is one such function. Furthermore, for any nonnegative integer n the R-orbit of 2^n *contains the subsequence*

$$2^n, 2^{T(n)}, 2^{T^2(n)}, 2^{T^3(n)} \ldots$$

and these are the only powers of two that occur.

Note that the function R given in the theorem is of the form (1.1) if we take

$$d = \operatorname{lcm}(11, 15, 17, 5, 21, 13, 7, 4, 2) = 1021020$$

since the first condition satisfied by n will also be the first condition satisfied by $n + dj$ for any j.

Proof: The proof is a straightforward application of Conway's *Fractran* language and its mathematical consequences. We refer the reader to [2] for details. A *Fractran program* consists of a finite list of positive rational numbers, $[r_1, \ldots r_t]$. The state of a *Fractran machine* consists of a single positive integer

S. The exponents of the primes in the prime factorization of S are used as registers for storing nonnegative integers. The program is executed by multiplying S by the first rational number in the list for which the product is a nonnegative integer (and halts if no such integer exists). Thus, each *Fractran* program corresponds to a function of the form (1.1) where execution of the program corresponds to iteration of the function.

The *Fractran* program

$$\left[\frac{1}{11},\frac{136}{15},\frac{5}{17},\frac{4}{5},\frac{26}{21},\frac{7}{13},\frac{1}{7},\frac{33}{4},\frac{5}{2},7\right] \tag{1.3}$$

when started with $S=2^n$, will produce $S=2^{T(n)}$ as the next S power of 2 in the orbit. To see this, consider the flowchart for this program indicated in Figure 1. (In what follows we will only be concerned with an initial state that is a power of 2, as required.)

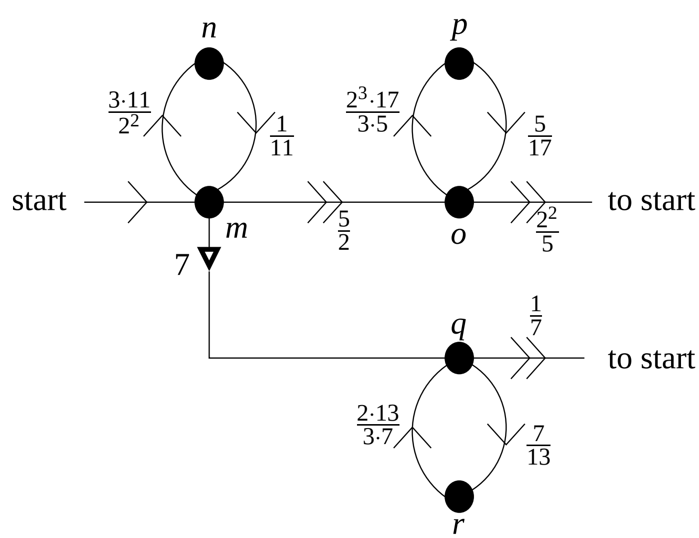

Fig. 1: A Fractran program for T

The edges of the flowchart are labeled in order of decreasing priority using a single arrow, double arrow, and triangle respectively. At a given node, the current state S is multiplied by the fraction labeling the edge of highest priority for which the product is a positive integer. The powers of the primes $5,7,11,13,17$ in S correspond to the nodes o,q,n,r,p respectively, a positive exponent of one of the primes indicating the program is at that node (and it is at node m if it is at no other node). The exponents of 2 and 3 in S are used as registers to compute T. We will refer to these exponents as α and β respectively.

When the program is started with $S=2^n$ at node m, it will execute the loop between nodes m and n exactly $q=\left\lfloor\frac{n}{2}\right\rfloor$ times, each time decreasing α by 2 and incrementing β. This results in $S=2^{n \bmod 2}3^q$.

If n is odd then $n=2q+1$ for some positive integer q and execution proceeds to node o where the state becomes $S=3^q5$. The loop between nodes o and p then produces $S=2^{3q}5$ which is then multiplied by

$\frac{2^2}{5}$ to produce

$$S = 2^{3q+2} = 2^{(6q+4)/2} = 2^{(6q+3+1)/2} = 2^{(3(2q+1)+1)/2} = 2^{(3n+1)/2} = 2^{T(n)}$$

as required.

If n is even, then upon completion of the mn loop S is multiplied by 7 moving execution to node q. The loop between nodes q and r produces $S = 2^q 7$ which is then multiplied by $1/7$ to produce

$$S = 2^q = 2^{n/2} = 2^{T(n)}$$

as required.

Iteration of the function R given in the theorem starting with seed 2^n corresponds exactly to execution of this *Fractran* program (the sequence of states being the R-orbit of 2^n). Since the choice of primes and algorithm used in this program was arbitrary, there are infinitely many such programs, and thus infinitely many such functions. This completes the proof. $\square$

Theorem 1 shows the relationship between the R-orbits of two powers and the $3x+1$ problem. One might ask for its own sake[†] how the iterates of R behave for arbitrary positive integer inputs. We answer this question with the following result.

Theorem 2 *Let R be defined as in (1.2). Then for all $a,b,c,d,e,f,g,h \in \mathbb{N}$*
1. for all $m \in \mathbb{Z}^+$ with $\gcd(m, 2\cdot 3\cdot 5\cdot 7\cdot 11\cdot 13\cdot 17) = 1$,

$$R\left(2^a 3^b 5^c 7^d 11^e 13^f 17^g m\right) = m \cdot R\left(2^a 3^b 5^c 7^d 11^e 13^f 17^g\right)$$

and
2. there exists $k \in \mathbb{N}$ such that $R^k\left(2^a 3^b 5^c 7^d 11^e 13^f 17^g\right) = 2^j$ for some j.

Thus if we iterate R starting with an arbitrary positive integer n, the prime factors of n that are greater than 17 are left unchanged, and the iterates of the remaining factor eventually reach a two power (after which the behavior proceeds as indicated in Theorem 1).

Proof: The proof of part (1) follows immediately from the definition of R, since prime factors greater than 17 are not affected when a positive integer is multiplied by any of the rational numbers listed in (1.3).

To prove part (2), let S be the set of positive integers that are not divisible by a prime greater than 17. Since no prime greater than 17 is a factor of the numerator of any fraction in (1.3), R maps elements of S to elements of S.

Let S' be the subset of S consisting of integers of the form $2^a 3^b$ for some $a,b \in \mathbb{N}$. Let $a,b \in \mathbb{N}$. By the definition of R, $R^2\left(2^{a+2}3^b\right) = 2^a 3^{b+1}$ so that $R^{2b}\left(2^{a+2b}\right) = 2^a 3^b$. Thus any element of S' is in the R-orbit of a power of two. Since the R-orbit of 2^{a+2b} contains infinitely many terms that are powers of two by Theorem 1, so does the R-orbit of $2^a 3^b$ for any $a,b \in \mathbb{N}$. Thus it suffices to show that the R-orbit of any element of S contains an element of S'.

Define $\alpha : S \to \mathbb{N}$ by $\alpha\left(2^{e_1} 3^{e_2} 5^{e_3} 7^{e_4} 11^{e_5} 13^{e_6} 17^{e_7}\right) = \sum_{i=2}^{7} e_i$. We argue by contradiction, and suppose that we have an element n of S so that all iterates $R^k(n) \notin S'$. Then all terms in the R-orbit of n are divisible

[†] Thanks to the anonymous referee of an earlier draft of this paper for suggesting this line of inquiry.

by some prime in $\{5,7,11,13,17\}$. Thus by the definition of R, for all $k\geq 1$, $R^k(n)=r_kR^{k-1}(n)$ for some $r_k\in\left\{\frac{1}{11},\frac{136}{15},\frac{5}{17},\frac{4}{5},\frac{26}{21},\frac{7}{13},\frac{1}{7}\right\}$. For any $k\in\mathbb{N}$, if $r_{k+1}\in\left\{\frac{1}{11},\frac{136}{15},\frac{4}{5},\frac{26}{21},\frac{1}{7}\right\}$ then

$$\alpha\left(R^{k+1}(n)\right)=\alpha\left(r_{k+1}R^k(n)\right)<\alpha\left(R^k(n)\right)$$

and if $r_{k+1}\in\left\{\frac{5}{17},\frac{7}{13}\right\}$ then

$$\alpha\left(R^{k+1}(n)\right)=\alpha\left(r_{k+1}R^k(n)\right)=\alpha\left(R^k(n)\right).$$

So the R-orbit of n has nonincreasing values of α, i.e. the sequence

$$\alpha(n),\alpha(R(n)),\alpha\left(R^2(n)\right),\ldots \tag{1.4}$$

is a nonincreasing. Since none of the terms are a two power (by our assumption), (1.4) is a nonincreasing sequence of positive integers whose terms are all less than or equal to $\alpha(n)$. Thus there must be some $h\geq 0$ such that $\alpha\left(R^k(n)\right)=\alpha\left(R^h(n)\right)$ for all $k\geq h$. So $r_k\in\left\{\frac{5}{17},\frac{7}{13}\right\}$ for all $k\geq h$. But multiplication by these values of r_k decreases the exponent of either 13 or 17 in the prime factorization of an integer, so that repeated multiplication by these fractions eventually produces a non-integer value. This contradicts our assumption and completes the proof. □

Conway [1] used an argument similar to the proof of Theorem 1 to show that there exist functions of the form (1.1) for which the fate of the orbit of an arbitrary positive integer is algorithmically undecidable. In Theorem 1 we turn this method around to obtain a positive result, and now illustrate how this result can be used to obtain mathematical results about the conjecture itself.

2 An Application

Let $x_0,\ldots,x_{n-1}$ be positive integers such that $x_i=T(x_{i-1})$ for $0<i<n$ and $x_0=T(x_{n-1})$. In this situation we say $\{x_0,\ldots,x_{n-1}\}$ is a T-*cycle.* If the $3x+1$ conjecture is true, then the only T-cycle of positive integers is $\{1,2\}$ (the existence of any other positive integer in a T-cycle being a counterexample). Thus it is of interest to study the properties of positive integer T-cyles.

Suppose $\{x_0,\ldots,x_{n-1}\}$ is a T-cycle of positive integers with $x_i=T(x_{i-1})$ for $0<i<n$ and $x_0=T(x_{n-1})$. Then by Theorem 1 the R-orbit of 2^{x_0} is also cyclic and contains $\{2^{x_0},\ldots,2^{x_{n-1}}\}$ as a subset. Thus there exists some positive integer t such that $R^t(x_0)=x_0$. But each application of R is simply multiplication by one of the rational numbers in $\left\{\frac{1}{11},\frac{136}{15},\frac{5}{17},\frac{4}{5},\frac{26}{21},\frac{7}{13},\frac{1}{7},\frac{33}{4},\frac{5}{2},7\right\}$ so that we must have

$$x_0=R^t(x_0)=\left(\frac{1}{11}\right)^a\left(\frac{136}{15}\right)^b\left(\frac{5}{17}\right)^c\left(\frac{4}{5}\right)^d\left(\frac{26}{21}\right)^e\left(\frac{7}{13}\right)^f\left(\frac{1}{7}\right)^g\left(\frac{33}{4}\right)^h\left(\frac{5}{2}\right)^i7^jx_0$$

for some nonnegative integers a,b,c,d,e,f,g,h,i,j with $a+b+c+d+e+f+g+h+i+j=t$. Collecting prime factors on the right hand side and dividing by x_0 gives us

$$2^{3b+2d+e-2h-i}3^{-b-e+h}5^{-b+c-d+i}7^{-e+f-g+j}11^{-a+h}13^{e-f}17^{b-c}=1.$$

This yields the system of linear equations

$$\begin{aligned} 3b+2d+e-2h-i &= 0 \\ -b-e+h &= 0 \\ -b+c-d+i &= 0 \\ -e+f-g+j &= 0 \\ -a+h &= 0 \\ e-f &= 0 \\ b-c &= 0 \end{aligned}$$

which is equivalent to the system

$$\begin{aligned} a &= 2c+i \\ b &= c \\ d &= i \\ e &= c+i \\ f &= c+i \\ g &= j \\ h &= 2c+i. \end{aligned} \tag{2.1}$$

Now define $\mathcal{O} = \{i : x_i \text{ is odd}\}$ and $\mathcal{E} = \{i : x_i \text{ is even}\}$ and let $k = |\mathcal{O}|$ so that $|\mathcal{E}| = n-k$. Then as explained in the proof of Theorem 1 we see that

$$\begin{aligned} i &= k \\ j &= n-k \\ c &= \sum_{i\in\mathcal{O}} \left\lfloor \frac{x_i}{2} \right\rfloor \\ a &= \sum_{i=0}^{n-1} \left\lfloor \frac{x_i}{2} \right\rfloor \end{aligned} \tag{2.2}$$

Substituting (2.2) into $a = 2c+i$ from (2.1) we obtain

$$\sum_{i=0}^{n-1} \left\lfloor \frac{x_i}{2} \right\rfloor = 2\sum_{i\in\mathcal{O}} \left\lfloor \frac{x_i}{2} \right\rfloor + k. \tag{2.3}$$

But $\sum_{i=0}^{n-1} \left\lfloor \frac{x_i}{2} \right\rfloor = \sum_{i\in\mathcal{E}} \left\lfloor \frac{x_i}{2} \right\rfloor + \sum_{i\in\mathcal{O}} \left\lfloor \frac{x_i}{2} \right\rfloor$. Substituting this into (2.3) and simplifying proves

Corollary 1 *If $\{x_0,\ldots,x_{n-1}\}$ is a T-cycle of positive integers and*

$$\mathcal{O} = \{i : x_i \textit{ is odd}\} \textit{ and } \mathcal{E} = \{i : x_i \textit{ is even}\}$$

then

$$\sum_{i\in\mathcal{E}} \left\lfloor \frac{x_i}{2} \right\rfloor = \sum_{i\in\mathcal{O}} \left\lfloor \frac{x_i}{2} \right\rfloor + k.$$

It should be noted that this formula can be proven directly from the known relationship

$$\sum_{i\in\mathcal{E}} x_i = \sum_{i\in\mathcal{O}} x_i + k \tag{2.4}$$

(obtained by noticing that $\{x_0,\ldots,x_{n-1}\} = \{T(x_0),\ldots,T(x_{n-1})\}$ so that $\sum x_i = \sum T(x_i)$ and thus $\sum_{i\in\mathcal{E}} x_i + \sum_{i\in\mathcal{O}} x_i = \sum_{i\in\mathcal{O}} \frac{3x_i+1}{2} + \sum_{i\in\mathcal{E}} \frac{x_i}{2}$ which can be solved to obtain (2.4)). However, the method used here reveals the results of the Corollary without specifically searching for those results. Thus this method provides a general approach for discovering new mathematical results by simply coding different algorithms for computing T (or any other computable integer function) and solving a simple linear system.

References

[1] Conway, J., *Unpredictable Iterations*, Proc. 1972 Number Theory Conference, University of Colorado, Boulder, Colorado (1972) 49-52

[2] Conway, J., *FRACTRAN: A Simple Universal Programming Language for Arithmetic*, Open Problems in Communication and Computation (Ed. T. M. Cover and B. Gopinath), New York: Springer-Verlag, (1987) 4-26

[3] Lagarias, J. C., *The 3x + 1 problem and its generalizations,* Am. Math. Monthly **92** (1985), 3-23

[4] Wirsching, G., *The Dynamical System Generated by the* $3n+1$ *Function*, Lecture Notes in Mathematics **1681**, Springer-Verlag, 1998, ISBN: 3-540-63970-5