ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Number Theory 117 (2006) 146–159
JOURNAL OF Number Theory
www.elsevier.com/locate/jnt

The $3x+1$ semigroup

David Applegate[a], Jeffrey C. Lagarias[b,*]

[a] *AT&T Laboratories, Florham Park, NJ 07932-0971, USA*
[b] *Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1109, USA*

Received 6 June 2004; revised 20 April 2005

Available online 13 September 2005

Communicated by D. Goss

Abstract

The $3x+1$ semigroup is the multiplicative semigroup $\mathscr{S}$ of positive rational numbers generated by $\{\frac{2k+1}{3k+2} : k \geqslant 0\}$ together with $\{2\}$. This semigroup encodes backwards iteration under the $3x+1$ map, and the $3x+1$ conjecture implies that it contains every positive integer. This semigroup is proved to be the set of positive rationals $\frac{a}{b}$ in lowest terms with $b \not\equiv 0 \pmod 3$, and so contains all positive integers.

MSC: primary 11B83; secondary 11Y16; 58F13

1. Introduction

The $3x+1$ problem concerns the behavior under iteration of the $3x+1$ function $T : \mathbb{Z} \to \mathbb{Z}$ given by

$$T(n) = \begin{cases} \dfrac{n}{2} & \text{if } n \equiv 0 \pmod 2 \\[2ex] \dfrac{3n+1}{2} & \text{if } n \equiv 1 \pmod 2 \end{cases} \tag{1}$$

* Corresponding author.
E-mail addresses: david@research.att.com (D. Applegate), lagarias@umich.edu (J.C. Lagarias).

0022-314X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jnt.2005.06.010

The $3x+1$ Conjecture asserts that for each $n \geqslant 1$, some iterate $T^{(k)}(n) = 1$. This is a notoriously hard problem, work on which is surveyed in Lagarias [2] and Wirsching [7]. It has been verified for all $n < 2.8 \times 10^{17}$ (see Oliveira e Silva [4]) but remains unsolved.

Recently Farkas [1] proposed an interesting weakening of the $3x+1$ problem, as follows. Let $\mathcal{S}$ denote the multiplicative semigroup of positive rational numbers generated by $\{\frac{n}{T(n)} : n \geqslant 0\}$, i.e. by 2 and by $\{\frac{2k+1}{3k+2} : k \geqslant 0\}$. We call $\mathcal{S}$ the $3x+1$ semigroup, and write

$$\mathcal{S} := \left\langle 2, \frac{1}{2}, \frac{3}{5}, \frac{5}{8}, \frac{7}{11} \cdots \right\rangle.$$

Farkas formulated the following conjecture.

Weak 3*x*+1 Conjecture. *The* $3x+1$ *semigroup* $\mathcal{S}$ *contains every positive integer.*

The semigroup $\mathcal{S}$ encodes inverse iteration by the $3x+1$ function. That is, the semigroup $\mathcal{S}$ contains $1 = 2 \cdot \frac{1}{2}$, and has the property that if $T(n) \in \mathcal{S}$, then also $n \in \mathcal{S}$, because each $\frac{n}{T(n)}$ is a generator of $\mathcal{S}$. It follows that if the $3x+1$ iteration eventually takes n to 1, then n belongs to $\mathcal{S}$. Thus the $3x+1$ conjecture implies the weak $3x+1$ conjecture.

The weak $3x+1$ conjecture appears a potentially easier question to resolve than the $3x+1$ conjecture, since the semigroup $\mathcal{S}$ permits some representations of integers as products of generators not corresponding to $3x+1$ iteration. Indeed, the object of this paper is to prove the following result characterizing all elements of the $3x+1$ semigroup, which implies the weak $3x+1$ conjecture.

Theorem 1.1. *The* $3x+1$ *semigroup* $\mathcal{S}$ *equals the set of all positive rationals* $\frac{a}{b}$ *in lowest terms having the property that* $b \not\equiv 0 \pmod 3$. *In particular, it contains every positive integer.*

In order to prove this result, we shall need to study the inverse semigroup $\mathcal{W} := \mathcal{S}^{-1}$ generated by $\{\frac{T(n)}{n} : n \geqslant 1\}$, i.e. by $\frac{1}{2}$ and by $\{\frac{3k+2}{2k+1} : n \geqslant 0\}$. That is,

$$\mathcal{W} := \mathcal{S}^{-1} = \left\langle \frac{1}{2}, \frac{2}{1}, \frac{5}{3}, \frac{8}{5}, \ldots \right\rangle.$$

We call this semigroup the *wild semigroup*, following the terminology used in a paper [3] of the second author, which was inspired by the novel "The Wild Numbers" [6]. The paper [3] formulated the following conjecture.

Wild Numbers Conjecture. *The integers in the wild semigroup* $\mathcal{W}$ *consist of all integers* $m \geqslant 1$ *with* $m \not\equiv 0 \pmod 3$. *Equivalently, the* $3x+1$ *semigroup* $\mathcal{S}$ *contains all unit fractions* $\frac{1}{m}$ *such that* $m \not\equiv 0 \pmod 3$.

Theorem 1.1 is equivalent to the truth of both the weak $3x + 1$ conjecture and the wild numbers conjecture. In [3] the two conjectures were shown to be equivalent, so to deduce Theorem 1.1 it would suffice to prove either one of them separately. However in the approach taken here we consider them together, and prove them simultaneously using an inductive method in which the truth of the conjectures to given bounds implies their truth to a larger bound. We use a see-saw method that increases the bound first of one, then the other.

In Section 2 we show the relevance of the $3x + 1$ iteration to the weak $3x + 1$ conjecture. This is the new ingredient introduced here relative to [3]. In Section 3 we then prove properties of integers in $\mathcal{W}$ and in Section 4 we complete the argument for Theorem 1.1.

2. Modified $3x + 1$ iterations

To prove the weak $3x + 1$ conjecture by induction on the size of the integer m, it would suffice to prove that under forward iteration of the $3x + 1$ map starting at a given $m \geqslant 2$, we eventually arrive at a smaller integer m', which would belong to the semigroup $\mathcal{S}$ by the induction hypothesis. The sequence of reverse $3x+1$ iterates going from m' back to m are multiplications by elements of $\mathcal{S}$, and this would establish that $m \in \mathcal{S}$. However, if this argument could be carried out, it would prove more, namely the $3x + 1$ conjecture itself. Since this problem seems out of reach, we considered a modification of this approach.

We take advantage of the fact that the $3x + 1$ iteration decreases "almost all" integers, in the sense of [2, Theorem A]. We recall that forward iteration of the $3x + 1$ function $T(\cdot)$ for j steps is known to decrease the value of an integer n in most congruence classes $n \pmod{2^j}$. Recall that the first j steps of the $3x + 1$ iteration are uniquely determined by the class $n \pmod{2^j}$ and that every symbol pattern of even and odd integers of length j occurs in some trajectory of length j, cf. Lagarias [2, Theorem B]. A residue class $s \pmod{2^j}$ is said to have a *strong stopping time* $k \leqslant j$ if the smallest integer $s \geqslant 2$ in the residue class decreases after k steps of iteration. This property is then inherited by all members $\geqslant 2$ of the residue class. As j increases the fraction of integers not having a strong stopping time goes to zero, but there still remain exponentially many residue classes $(\bmod\, 2^j)$ not having this decreasing property [2, Theorems C, D].

The semigroup $\mathcal{S}$ permits the possibility of going "uphill" by taking an initial value n to a value mn via some integer multiplier m, provided $m \in \mathcal{S}^{-1} = \mathcal{W}$. That is, if $\frac{1}{m} \in \mathcal{S}$ and if we know $mn \in \mathcal{S}$ then we may deduce $n = \frac{1}{m} \cdot mn \in \mathcal{S}$. We pay a price in going "uphill" of increasing the initial size of the integer, but in doing so, we may move from a "bad" residue class $s \pmod{2^j}$ to a "good" residue class $ms \pmod{2^j}$ which under iteration results in such a large decrease in the size of the number that it overcomes the added multiplicative factor m and arrives at an integer smaller than n in $\leqslant j$ steps. One can use this procedure only for j steps ahead because the members of the residue class only possess the same symbolic dynamics for j steps, and we wish the property of decrease to hold for all members of the residue class. If so, we can carry out the induction step for all integers in this particular "bad" residue class $s \pmod{2^j}$. Another variation of this idea is to multiply by various m's in the middle of the first j steps of the iteration; there is no reason why the multiplication must be done only at the first step, one may still gain by switching the residue class in the middle of the iteration.

One can now ask: is there a finite j and a finite list $\{m_1, m_2, \dots m_r\}$ of integer elements in $\mathcal{W}$ such that suitable multiplications by elements of this list will decrease elements in every residue class $\pmod{2^j}$ in this fashion? If so, this would yield a proof of the weak $3x + 1$ conjecture by induction on n.

This approach comes very close to succeeding, but there is an obstruction that in principle prevents complete success. We found by computer search, for small values of j, multiplier lists that established decrease for every residue class $\pmod{2^j}$ except for the class $-1 \pmod{2^j}$. These searches revealed that the class $-1 \pmod{2^j}$ resisted elimination for $12 \leqslant j \leqslant 30$. We then looked for and found the following proof that the class $-1 \pmod{2^j}$ can never be eliminated by this method. The iterates of a positive integer n in the congruence class $-1 \pmod{2^j}$ will behave exactly the same way as -1 does for the first j steps, allowing multipliers. We may write the jth iterate of -1 obtained using multipliers as $\frac{m_1 m_2 \cdots m_j a(-1)+b}{2^j}$, in which m_k is the multiplier used at the kth step (we allow $m_k = 1$), a is a power of 3, and b is a positive integer. For this multiplier sequence any $n \equiv -1 \pmod{2^j}$ will map to $\frac{m_1 m_2 \cdots m_j a n+b}{2^j}$ after j steps. However we must have

$$\frac{(m_1 m_2 \cdots m_j)a(-1) + b}{2^j} \leqslant -1,$$

because all iterates of -1, times multipliers, remain negative. Rearranging this inequality gives

$$(m_1 m_2 \cdots m_j)a \geqslant 2^j + b.$$

Now, for positive n, multiplying both sides by $\frac{n}{2^j}$ yields

$$\frac{(m_1 m_2 \cdots m_j)an + b}{2^j} \geqslant n + \frac{b(n+1)}{2^j} > n.$$

It follows that decrease cannot have occurred after j steps, and an argument for no decrease at any intermediate step is similar.

We conclude that to get an inductive proof of the weak $3x + 1$ conjecture along these lines, a new method will be needed to handle integers in the "bad" congruence class $-1 \pmod{2^j}$, and it will be necessary to consider an infinite set of multipliers in $\mathcal{W}$.

We now prove the decrease mentioned above for all residue classes $\pmod{4096}$ except the class $-1 \pmod{4096}$, using a fixed finite set H of multipliers given below; these are residue classes $\pmod{2^j}$ for $j = 12$. In what follows it will be important that the decrease is by a constant factor strictly smaller than one. In Section 3 we will verify the hypothesis $H \subset \mathcal{W}$ made in this lemma.

Lemma 2.1. *If* $H = \{5, 7, 11, 13, 23, 29, 43\} \subset \mathcal{W}$, *then for every integer* $x > 1$ *with* $x \not\equiv -1 \pmod{4096}$ *there exists* $s \in \mathcal{W}$ *such that* $sx \in \mathbb{Z}$ *and* $sx \leqslant \frac{76}{79}x$.

Proof. This is established case by case in Tables 1 and 2. Every path shown consists of iterations of $T(\cdot)$ and multiplications by integers in H, and thus consists of iterations of multiplications by elements of $\mathcal{W}$. The iteration takes k steps, where k is the given number of bits, and for integers n in the class $s \pmod{2^k}$, one has $T^{(k)}(n) = c(s)n + d(s)$, with

$$c(s) = \frac{3^l m_1 m_2 \cdots m_k}{2^k},$$

in which the m_i are the multipliers at each step and l is the number of odd elements in the resulting trajectory, and $d(s) \geqslant 0$. The quantity $c(s)$ is the "asymptotic ratio" reported in the second column of the tables.

The "class bits" presented in these tables are binary strings comprising the binary expansion of the residue class written in reverse order. The set of these binary strings together form a prefix code which by inspection certifies that every residue class $\pmod{4096}$ is covered except $-1 \pmod{4096}$. The data on the far right in the table gives the action on the smallest positive element in the congruence class (resp. second smallest element for the class containing $n = 1$). In each case the factor of decrease on all elements of the progression (excluding the element $n = 1$), reported as the "worst-case ratio" in the table, is that given by the decrease on this particular element. $\square$

To deal with the residue class $-1 \pmod{2^j}$, we next show that there always exists a simple (but infinite) sequence of multipliers having the property that, starting from $n \equiv -1 \pmod{2^j}$, with $n > 0$ one arrives at a final integer n' that is only slightly larger than the initial starting point n. We will later make use of this to eliminate the congruence class $-1 + 2^j \pmod{2^{j+1}}$, in an induction on j.

Lemma 2.2. *Let* x, k, *and* j *be positive integers such that* $x \equiv -1 \pmod{2^k}$, *with* $1 \leqslant j \leqslant k$ *and* $j \equiv 1, 5 \pmod 6$. *Then the multiplier* $m = \frac{2^j+1}{3}$ *is an integer satisfying* $m \equiv 1, 5 \pmod 6$, *with the property that the jth iterate of* mx *satisfies the bound*

$$T^j(mx) = x + \frac{x+1}{2^j} \leqslant \frac{2^j+2}{2^j} x \ ,$$

and $T^j(mx) \equiv -1 \pmod{2^{k-j}}$. *If in addition* $x \not\equiv -1 \pmod{2^{k+1}}$, *then* $T^j(mx) \not\equiv -1 \pmod{2^{k+1-j}}$.

Table 1
Decreasing weak $3x+1$ paths, for $x \not\equiv 15 \pmod{16}$

Residue class Class bits	Asymptotic ratio	Worst-case ratio	Path
$0 \pmod 2$ 0	$\frac{1}{2}$ 0.5000	$\frac{1}{2}{}^*$ 0.5000	$2 \to 1$
$1 \pmod 4$ 10	$\frac{3}{4}$ 0.7500	$\frac{4}{5}{}^*$ 0.8000	$5 \to 8 \to 4$
$3 \pmod{16}$ 1100	$\frac{9}{16}$ 0.5625	$\frac{2}{3}$ 0.6667	$3 \to 5 \to 8 \to 4 \to 2$
$11 \pmod{32}$ 11010	$\frac{27}{32}$ 0.8438	$\frac{10}{11}$ 0.9091	$11 \to 17 \to 26 \to 13 \to 20 \to 10$
$27 \pmod{128}$ 1101100	$\frac{117}{128}$ 0.9141	$\frac{25}{27}$ 0.9259	$27 \to 41 * 13 = 533 \to 800 \to 400 \to$ $\to 200 \to 100 \to 50 \to 25$
$91 \pmod{256}$ 11011010	$\frac{225}{256}$ 0.8789	$\frac{80}{91}$ 0.8791	$91 * 25 = 2275 \to 3413 \to 5120 \to$ $\to 2560 \to 1280 \to 640 \to 320 \to$ $\to 160 \to 80$
$219 \pmod{256}$ 11011011	$\frac{243}{256}$ 0.9492	$\frac{209}{219}$ 0.9543	$219 \to 329 \to 494 \to 247 \to 371 \to$ $\to 557 \to 836 \to 418 \to 209$
$59 \pmod{128}$ 1101110	$\frac{81}{128}$ 0.6328	$\frac{38}{59}$ 0.6441	$59 \to 89 \to 134 \to 67 \to 101 \to$ $\to 152 \to 76 \to 38$
$123 \pmod{256}$ 11011110	$\frac{189}{256}$ 0.7383	$\frac{91}{123}$ 0.7398	$123 * 7 = 861 \to 1292 \to 646 \to$ $\to 323 \to 485 \to 728 \to 364 \to$ $\to 182 \to 91$
$251 \pmod{256}$ 11011111	$\frac{207}{256}$ 0.8086	$\frac{203}{251}$ 0.8088	$251 * 23 = 5773 \to 8660 \to 4330 \to$ $\to 2165 \to 3248 \to 1624 \to 812 \to$ $\to 406 \to 203$
$7 \pmod{64}$ 111000	$\frac{45}{64}$ 0.7031	$\frac{5}{7}$ 0.7143	$7 * 5 = 35 \to 53 \to 80 \to 40 \to 20 \to$ $\to 10 \to 5$
$39 \pmod{128}$ 1110010	$\frac{105}{128}$ 0.8203	$\frac{32}{39}$ 0.8205	$39 * 35 = 1365 \to 2048 \to 1024 \to$ $\to 512 \to 256 \to 128 \to 64 \to 32$
$103 \pmod{512}$ 111001100	$\frac{351}{512}$ 0.6855	$\frac{71}{103}$ 0.6893	$103 \to 155 \to 233 * 13 = 3029 \to$ $\to 4544 \to 2272 \to 1136 \to 568 \to$ $\to 284 \to 142 \to 71$
$359 \pmod{512}$ 111001101	$\frac{315}{512}$ 0.6152	$\frac{221}{359}$ 0.6156	$359 * 35 = 12565 \to 18848 \to 9424 \to$ $\to 4712 \to 2356 \to 1178 \to 589 \to$ $\to 884 \to 442 \to 221$
$231 \pmod{256}$ 11100111	$\frac{135}{256}$ 0.5273	$\frac{122}{231}$ 0.5281	$231 * 5 = 1155 \to 1733 \to 2600 \to$ $\to 1300 \to 650 \to 325 \to 488 \to$ $\to 244 \to 122$
$23 \pmod{32}$ 11101	$\frac{27}{32}$ 0.8438	$\frac{20}{23}$ 0.8696	$23 \to 35 \to 53 \to 80 \to 40 \to 20$

Table 2
Decreasing weak $3x+1$ paths, for $x \equiv 15 \pmod{16}$

Residue class Class bits	Asymptotic ratio	Worst-case ratio	Path
15 (mod 128) 1111000	$\frac{81}{128}$ 0.6328	$\frac{10}{15}$ 0.6667	$15 \to 23 \to 35 \to 53 \to 80 \to 40 \to$ $\to 20 \to 10$
79 (mod 256) 11110010	$\frac{243}{256}$ 0.9492	$\frac{76}{79}$ 0.9620	$79 \to 119 \to 179 \to 269 \to 404 \to$ $\to 202 \to 101 \to 152 \to 76$
207 (mod 256) 11110011	$\frac{225}{256}$ 0.8789	$\frac{182}{207}$ 0.8792	$207 * 5 = 1035 \to 1553 * 5 = 7765 \to$ $\to 11648 \to 5824 \to 2912 \to$ $\to 1456 \to 728 \to 364 \to 182$
47 (mod 128) 1111010	$\frac{117}{128}$ 0.9141	$\frac{43}{47}$ 0.9149	$47 * 13 = 611 \to 917 \to 1376 \to$ $\to 688 \to 344 \to 172 \to 86 \to 43$
111 (mod 128) 1111011	$\frac{99}{128}$ 0.7734	$\frac{86}{111}$ 0.7748	$111 * 11 = 1221 \to 1832 \to 916 \to$ $\to 458 \to 229 \to 344 \to 172 \to 86$
31 (mod 64) 111110	$\frac{33}{64}$ 0.5156	$\frac{16}{31}$ 0.5161	$31 * 11 = 341 \to 512 \to 256 \to$ $\to 128 \to 64 \to 32 \to 16$
63 (mod 128) 1111110	$\frac{99}{128}$ 0.7734	$\frac{49}{63}$ 0.7778	$63 * 11 = 693 \to 1040 \to 520 \to$ $\to 260 \to 130 \to 65 \to 98 \to 49$
127 (mod 256) 11111110	$\frac{129}{256}$ 0.5039	$\frac{64}{127}$ 0.5039	$127 * 43 = 5461 \to 8192 \to 4096 \to$ $\to 2048 \to 1024 \to 512 \to 256 \to$ $\to 128 \to 64$
255 (mod 512) 111111110	$\frac{387}{512}$ 0.7559	$\frac{193}{255}$ 0.7569	$255 * 43 = 10965 \to 16648 \to 8224 \to$ $\to 4112 \to 2056 \to 1028 \to$ $\to 514 \to 257 \to 386 \to 193$
511 (mod 1024) 1111111110	$\frac{783}{1024}$ 0.7646	$\frac{391}{511}$ 0.7652	$511 \to 767 * 29 = 22243 \to 33365 \to$ $\to 50048 \to 25024 \to 12512 \to$ $\to 6256 \to 3128 \to 1564 \to$ $\to 782 \to 391$
1023 (mod 2048) 11111111110	$\frac{1089}{2048}$ 0.5317	$\frac{544}{1023}$ 0.5318	$1023 * 11 = 11253 \to 16880 \to$ $\to 8440 \to 4220 \to 2110 \to$ $\to 1055 * 11 = 11605 \to 17408 \to$ $\to 8704 \to 4352 \to 2176 \to$ $\to 1088 \to 544$
2047 (mod 4096) 111111111110	$\frac{3267}{4096}$ 0.7976	$\frac{1633}{2047}$ 0.7978	$2047 * 11 = 22517 \to 33776 \to$ $\to 16888 \to 8444 \to 4222 \to$ $\to 2111 * 11 = 23221 \to 34832 \to$ $\to 17416 \to 8708 \to 4354 \to$ $\to 2177 \to 3266 \to 1633$

Proof. Since $j \equiv 1, 5 \pmod 6$, $2^j \equiv 2, 5 \pmod 9$, so $m = \frac{2^j+1}{3}$ is an odd integer and $m \not\equiv 0 \pmod 3$. Since $mx \equiv -m \pmod{2^k}$, mx is odd, so

$$T(mx) = \frac{3mx+1}{2} = \frac{2^j x + x + 1}{2} .$$

Since $x \equiv -1 \pmod{2^k}$ and $k \geqslant j$, $2^j x + x + 1 \equiv 0 \pmod{2^j}$. Thus

$$T^j(mx) = \frac{2^j x + x + 1}{2^j} = x + \frac{x+1}{2^j} .$$

Now $\frac{x+1}{2^j} \equiv 0 \pmod{2^{k-j}}$, so $T^j(mx) \equiv -1 \pmod{2^{k-j}}$. If in addition $x \not\equiv -1 \pmod{2^{k+1}}$, then $\frac{x+1}{2^j} \not\equiv 0 \pmod{2^{k+1-j}}$, so $T^j(mx) \not\equiv -1 \pmod{2^{k+1-j}}$. $\square$

To make use of Lemma 2.2 in an inductive proof, we need to establish that after using it on n to obtain $n' = T^{(j)}(n)$ a single $3x+1$ iteration applied to n' produces an integer n'' smaller than n. This is the aim of the following lemma, which gives an inductive method of eliminating the class $-1 + 2^j \pmod{2^{j+1}}$ using a suitable integer multiplier m, assuming that m is a wild integer.

Lemma 2.3. *Suppose $H \subset \mathcal{W}$. Let $x \equiv -1 \pmod{2^k}$ and $x \not\equiv -1 \pmod{2^{k+1}}$, for a fixed $k \geqslant 12$. Now choose j so that $j \equiv 1 \pmod 6$ and $k - 10 \leqslant j \leqslant k - 5$. Then $m = (2^j+1)/3$ is an integer, and if $m \in \mathcal{W}$, then there exists $s \in \mathcal{W}$ such that $sx \in \mathbb{Z}$ and $sx \leqslant \frac{1235}{1264}x$.*

Proof. First, note that for all $k \geqslant 12$, since $j \equiv 1 \pmod 6$ and $j \geqslant k - 10 \geqslant 2$, we have $j \geqslant 7$. From Lemma 2.2, $m \in \mathbb{Z}$ and $T^j(mx) = x + \frac{x+1}{2^j}$, so there exists $s_1 \in \mathcal{W}$ such that $s_1 mx \in \mathbb{Z}$, $s_1 mx = x + \frac{x+1}{2^j}$, $s_1 mx \equiv -1 \pmod{2^{k-j}}$, and $s_1 mx \not\equiv -1 \pmod{2^{k+1-j}}$. But $k - j \leqslant 10$, so $s_1 mx \not\equiv -1 \pmod{2^{11}}$. Thus from Lemma 2.1, there exists $s_2 \in \mathcal{W}$ such that $s_2(s_1 mx) \in \mathbb{Z}$ and $s_2(s_1 mx) \leqslant \frac{76}{79}(s_1 mx)$. Now $j \geqslant 7$ and the bound of Lemma 2.2 gives

$$x + \frac{x+1}{2^j} \leqslant \frac{2^j+2}{2^j} x \leqslant \frac{130}{128} x,$$

so that $s_2 s_1 mx \leqslant \frac{76}{79}\frac{130}{128}x = \frac{1235}{1264}x$. $\square$

3. Wild integers

The wild integers are the integers in the wild semigroup $\mathcal{W}$. The "multiplier" approach begun in Section 2 required the use of multipliers that are wild integers, and indicated that in taking this approach one would need to consider an infinite set of

Table 3
Membership certificates in $\mathcal{W}$ for members of H

5	=	$(\frac{1}{2})^2 \cdot (\frac{11}{7})^2 \cdot \frac{17}{11} \cdot \frac{26}{17} \cdot \frac{83}{55} \cdot \frac{98}{65} \cdot \frac{125}{83}$
	=	$(\frac{1}{2})^2 \cdot g(3)^2 \cdot g(5) \cdot g(8) \cdot g(27) \cdot g(32) \cdot g(41)$
7	=	$(\frac{1}{2})^2 \cdot \frac{11}{7} \cdot \frac{26}{17} \cdot \frac{35}{23} \cdot \frac{215}{143} \cdot \frac{299}{199} \cdot \frac{323}{215} \cdot \frac{371}{247} \cdot \frac{398}{265}$
	=	$(\frac{1}{2})^2 \cdot g(3) \cdot g(8) \cdot g(11) \cdot g(71) \cdot g(99) \cdot g(107) \cdot g(123) \cdot g(132)$
11	=	$(\frac{1}{2})^2 \cdot (\frac{11}{7})^2 \cdot \frac{26}{17} \cdot \frac{35}{23} \cdot \frac{215}{143} \cdot \frac{299}{199} \cdot \frac{323}{215} \cdot \frac{371}{247} \cdot \frac{398}{265}$
	=	$(\frac{1}{2})^2 \cdot g(3)^2 \cdot g(8) \cdot g(11) \cdot g(71) \cdot g(99) \cdot g(107) \cdot g(123) \cdot g(132)$
13	=	$(\frac{1}{2})^3 \cdot (\frac{11}{7})^2 \cdot (\frac{17}{11})^3 \cdot (\frac{26}{17})^2 \cdot \frac{35}{23} \cdot \frac{215}{143} \cdot \frac{299}{199} \cdot \frac{323}{215} \cdot \frac{371}{247} \cdot \frac{398}{265}$
	=	$(\frac{1}{2})^3 \cdot g(3)^2 \cdot g(5)^3 \cdot g(8)^2 \cdot g(11) \cdot g(71) \cdot g(99) \cdot g(107) \cdot g(123) \cdot g(132)$
23	=	$(\frac{1}{2})^5 \cdot \frac{11}{7} \cdot \frac{26}{17} \cdot \frac{35}{23} \cdot \frac{47}{31} \cdot \frac{137}{91} \cdot \frac{155}{103} \cdot \frac{206}{137} \cdot \frac{215}{143} \cdot (\frac{299}{199})^2 \cdot \frac{323}{215} \cdot \frac{353}{235} \cdot \frac{371}{247}$ · · $(\frac{398}{265})^2 \cdot \frac{530}{353}$
	=	$(\frac{1}{2})^5 \cdot g(3) \cdot g(8) \cdot g(11) \cdot g(15) \cdot g(45) \cdot g(51) \cdot g(68) \cdot g(71) \cdot g(99)^2$ · · $g(107) \cdot g(117) \cdot g(123) \cdot g(132)^2 \cdot g(176)$
29	=	$(\frac{1}{2})^5 \cdot (\frac{11}{7})^4 \cdot (\frac{17}{11})^2 \cdot (\frac{26}{17})^2 \cdot \frac{29}{19} \cdot \frac{38}{25} \cdot (\frac{83}{55})^2 \cdot (\frac{98}{65})^2 \cdot (\frac{125}{83})^2$
	=	$(\frac{1}{2})^5 \cdot g(3)^4 \cdot g(5)^2 \cdot g(8)^2 \cdot g(9) \cdot g(12) \cdot g(27)^2 \cdot g(32)^2 \cdot g(41)^2$
43	=	$(\frac{1}{2})^{11} \cdot (\frac{11}{7})^5 \cdot (\frac{17}{11})^2 \cdot (\frac{26}{17})^3 \cdot \frac{29}{19} \cdot \frac{35}{23} \cdot \frac{38}{25} \cdot (\frac{83}{55})^2 \cdot (\frac{98}{65})^2 \cdot (\frac{125}{87})^2 \cdot \frac{215}{143}$ · · $\frac{299}{199} \cdot \frac{305}{203} \cdot \frac{323}{215} \cdot \frac{344}{229} \cdot \frac{371}{247} \cdot \frac{398}{265} \cdot \frac{458}{305}$
	=	$(\frac{1}{2})^{11} \cdot g(3)^5 \cdot g(5)^2 \cdot g(8)^3 \cdot g(9) \cdot g(11) \cdot g(12) \cdot g(27)^2 \cdot g(32)^2 \cdot g(41)^2$ · · $g(71) \cdot g(99) \cdot g(101) \cdot g(107) \cdot g(114) \cdot g(123) \cdot g(132) \cdot g(152)$

multipliers. This in turn seems to require understanding the complete structure of the integer elements in $\mathcal{W}$, which leads to investigation of the wild numbers conjecture.

In this section we establish properties of wild integers, giving criteria for establishing their existence. We first show that the elements in H in Section 2 are wild integers. Here we write $g(n) = \frac{3n+2}{2n+1}$.

Lemma 3.1. *The set* $H = \{5, 7, 11, 13, 23, 29, 43\}$ *is contained in the wild semigroup* $\mathcal{W} = \mathcal{S}^{-1}$.

Proof. Table 3 gives certificates showing that the elements in H belong to $\mathcal{W}$, representing them in terms of the generators of $\mathcal{W}$. The table uses the notation $g(n) = \frac{3n+2}{2n+1}$, for $n \geqslant 1$. Aside from $p = 5$, these identities were found by computer search by Allan Wilks, see Section 2 of [3]. $\square$

The following lemma uses the truth of the weak $3x + 1$ conjecture on an initial interval to extend the range on which the wild numbers conjecture holds.

Lemma 3.2. *Suppose that the weak $3x + 1$ conjecture holds for $1 \leqslant n \leqslant 2^j - 2$ and that the wild numbers conjecture holds for $1 \leqslant m \leqslant \frac{2^j-1}{189}$, with $j \geqslant 16$. Then the wild numbers conjecture holds for $1 \leqslant m \leqslant \frac{2^{j+1}-1}{189}$.*

Proof. It suffices to prove that every prime q with $\frac{2^j-1}{189} < q \leqslant \frac{2^{j+1}-1}{189}$ lies in $\mathcal{W}$. Proceeding by induction on increasing q, we may assume every prime p with $3 < p < q$ lies in $\mathcal{W}$. It now suffices to prove: *there exists a positive integer $n \leqslant 2^j - 2$ with $nq \in \mathcal{W}$*. For if so, then the induction hypothesis implies that $n \in \mathcal{S}$ so that $q = \frac{1}{n} \cdot nq \in \mathcal{W}$. In establishing this we will consider only those n such that $nq \equiv -1 \pmod 9$. Then $nq = 3l + 2$ for some positive integer l, and $nq = t \cdot (2l + 1)$, where $t = \frac{3l+2}{2l+1} \in \mathcal{W}$. Thus it will suffice to show $2l + 1 \in \mathcal{W}$.

To carry this out, define a as the least positive residue with $aq \equiv -1 \pmod 9$, so that $0 < a < 9$. For n in the arithmetic progression $n = 9k + a$, setting $nq = 3l + 2$, we have

$$2l + 1 = 2\left(\frac{nq-2}{3}\right) + 1 = \frac{2}{3}((9k + a)q - 2) + 1 = 6qk + r \quad \text{with} \quad r := \frac{2aq - 1}{3}.$$

The condition $aq \equiv -1 \pmod 9$ gives $r \equiv -1 \pmod 6$, and $r \pmod{6q}$ is invertible $\pmod{6q}$. For the given prime q the values a and r are determined, and we need to find a suitable value of k. If $0 \leqslant k < 6q$ then:

$$n = 9k + a \leqslant 9(6q - 1) + a < 54q \leqslant 54\left(\frac{2^{j+1}-1}{189}\right) = \frac{2}{7}(2^{j+1} - 1) \leqslant 2^j - 2,$$

so $n \in \mathcal{S}$ by hypothesis. Therefore it suffices to prove: *for each prime q with $\frac{2^j-1}{189} < q \leqslant \frac{2^{j+1}-1}{189}$ there exists an integer $0 \leqslant k < 6q$ such that $6qk + r \in \mathcal{W}$.*

Define a positive integer to be q-smooth if all its prime factors are smaller than q. Let Σ_q denote the set of q-smooth integers s with $0 < s < 6q$ and $gcd(s, 6q) = 1$. Then every $s \in \Sigma_q$ is a product of primes p with $5 \leqslant p < q$, and the induction hypothesis implies that $s \in \mathcal{W}$.

Claim. *If $q \geqslant 256$ then $|\Sigma_q| > q - 1$.*

Assuming the claim is true, we can apply it in our situation because $q > \frac{2^j-1}{189} \geqslant \frac{2^{16}-1}{189} > 346$. The claim implies that Σ_q contains more than half of the invertible residue classes $\pmod{6q}$, since $\phi(q) = 2(q - 1)$. Therefore, in the group of invertible residue classes $\pmod{6q}$, the sets Σ_q and $r \cdot \Sigma_q^{-1}$ must meet, since each contains more than half of the classes. Therefore $s_1 \equiv r \cdot s_2^{-1} \pmod{6q}$, for some $s_1, s_2 \in \Sigma_q$. Now $s_1 s_2 \equiv r \pmod{6q}$, and we may define $k \geqslant 0$ by setting $s_1 s_2 = 6qk + r$. Since each $s_i \in \Sigma_q \subseteq \mathcal{W}$ we have $6qk + r \in \mathcal{W}$. Since $s_1, s_2 < 6q$ we find that $k < 6q$, as required. Thus the proof of Lemma 3.2 will be complete once the claim is established.

To prove the claim, since $\phi(6q) = 2q - 2$ we may reformulate it as the assertion: *there are at most* $q - 2$ *invertible residue classes below* $6q$ *which are not* q*-smooth*. The non-q-smooth numbers below $6q$ relatively prime to q consist of the primes q' with $q < q' < 6q$ together with the integers $5q'$ where q' is prime with $q < q' < \frac{6}{5}q$. Thus we must show that for $q > 256$,

$$(\pi(6q) - \pi(q)) + \pi\left(\tfrac{6}{5}q\right) - \pi(q) \leqslant q - 2. \tag{2}$$

The left side of (2) is $O(\frac{q}{\log q})$ by the prime number theorem, so (2) holds for all sufficiently large q; it remains to establish the specific bound. We use explicit inequalities for prime counting functions due to Rosser and Schoenfeld [5, Theorems 1,2], which state that for all $x \geqslant 17$,

$$\frac{x}{\log x} < \pi(x) < \frac{x}{\log x - \frac{3}{2}}, \tag{3}$$

and also that, for all $x \geqslant 114$,

$$\pi(x) < \frac{5}{4}\frac{x}{\log x}.$$

The first of these inequalities gives

$$\pi(6x) \leqslant \frac{6x}{\log(6x) - \frac{3}{2}} \leqslant \frac{6x}{\log x}$$

since $\log 6 \geqslant \frac{3}{2}$. The second gives, for $x \geqslant 256$,

$$\begin{aligned}\pi\left(\tfrac{6}{5}x\right) &< \frac{5}{4}\left(\frac{\frac{6}{5}x}{\log(\frac{6}{5}x)}\right)\\ &< \frac{3}{2}\frac{x}{\log x}\left(\frac{\log x}{\log x + \log\frac{6}{5}}\right)\\ &< \frac{3}{2}\frac{x}{\log x}\left(1 - \frac{\frac{1}{6}}{\log x + \frac{1}{6}}\right)\\ &< \frac{3}{2}\frac{x}{\log x} - 2,\end{aligned}$$

where we used $\log\frac{6}{5} > \frac{1}{6}$, and $x \geqslant 256$ was used at the last step. Combining these bounds gives, for $x \geqslant 256 > e^{11/2}$,

$$\pi(6x) + \pi\left(\frac{6}{5}x\right) - 2\pi(x) < \frac{11}{2}\frac{x}{\log x} - 2 \leqslant x - 2,$$

which proves the claim. $\square$

4. Completion of proofs

Proof of Theorem 1.1. The theorem is equivalent to the truth of the weak $3x+1$ conjecture and the wild numbers conjecture. Together these two conjectures imply that the semigroup $\mathcal{S}$ contains all rationals $\frac{a}{b} = a \cdot \frac{1}{b}$ with $b \not\equiv 0 \pmod 3$. However $\mathcal{S}$ contains no rational $\frac{a}{b}$ in lowest terms with $b \equiv 0 \pmod 3$, because no generator of $\mathcal{S}$ contains a multiple of 3 in its denominator. Conversely, if $\mathcal{S}$ contains all such rationals, then both conjectures hold.

We prove the weak $3x+1$ conjecture and wild numbers conjecture simultaneously by induction on $k \geqslant 12$, using the following three inductive hypotheses.

(1) For each integer $x > 1$ with $x \not\equiv -1 \pmod{2^k}$ there is an element $s \in \mathcal{W}$ such that sx is an integer and $sx \leqslant \frac{1235}{1264}x$.
(2) The weak $3x+1$ conjecture is true for $1 \leqslant n \leqslant 2^k - 2$.
(3) The wild integers conjecture is true for $1 \leqslant m \leqslant \frac{2^k-1}{189}$.

We treat the induction step first, and the base case afterwards. We suppose the inductive hypotheses hold for some $k \geqslant 12$, and must show they then hold for $k+1$.

Hypotheses (2) and (3) for k permit Lemma 3.2 to apply, whence for $k \geqslant 16$ we conclude that inductive hypothesis (3) holds for $k+1$. For the remaining cases $12 \leqslant k \leqslant 15$ we verify inductive hypothesis (3) for $k+1$ directly by computation, which is included in the base case below.

Inductive hypothesis (1) for k gives that all elements smaller than $2^{k+1} - 1$ except possibly $2^k - 1$ can be decreased by multiplication by an element of $\mathcal{W}$ to a smaller integer. We wish to apply Lemma 2.3 to show that all elements in the congruence class $-1 + 2^k \pmod{2^{k+1}}$ can also be decreased by multiplication by an element of $\mathcal{W}$ to an integer smaller by the multiplicative factor $\frac{1235}{1264}$. First, Lemma 3.1 shows that the elements of H belong to $\mathcal{W}$, establishing one hypothesis of Lemma 2.3. Second, the other multiplier $m = \frac{2^j+1}{3}$ in the hypothesis of Lemma 2.3 has $j = k - 5 - (k \pmod 6)$, and satisfies $m \leqslant \frac{2^{(k+1)-6}+1}{3} \leqslant \frac{2^{k+1}-1}{189}$, so $m \in \mathcal{W}$ by inductive hypothesis (3), which is already established to hold for $k+1$. Thus all the hypotheses of Lemma 2.3 are satisfied, and its conclusion verifies the inductive hypothesis (1) for $k+1$.

Next, inductive hypothesis (1) for $k+1$ establishes the decreasing property for all integers $1 < n \leqslant 2^{k+1} - 2$, hence the weak $3x+1$ conjecture follows for all integers in this range. This verifies inductive hypothesis (2) for $k+1$, and so completes the induction step.

It remains to treat the base case, which is $k = 12$ for hypotheses (1) and (2), and $k = 16$ for hypothesis (3). For $k = 12$ inductive hypothesis (1) is verified by Lemma 2.1. The inductive hypothesis (2) for $k = 12$ is verified by the fact that the $3x+1$ conjecture has been checked over the range $1 \leqslant n \leqslant 2^{12} = 4096$.

Finally we must verify inductive hypothesis (3) for $k = 16$. This requires verifying the wild numbers conjecture for $1 \leqslant m \leqslant \frac{2^{16}-1}{189} = \frac{65535}{189} < 400$. It suffices to do this for all primes below 400, except $p = 3$. Representations in the generators of $\mathcal{W}$ for all such primes below 50, are given in [3]. (Table 3 gives representations for some of these primes.) For primes $50 \leqslant p \leqslant 400$, one can check the criterion by computer using the method of Lemma 3.2, finding by computer search a q-smooth number in the appropriate arithmetic progression, for each prime q in the interval, and using the truth of the $3x+1$ conjecture for $1 < x < 10^5$. In fact, this q-smooth calculation can be carried out by computer for every q with $11 < q < 400$, and only the certificates for $p = 5$, $p = 7$, $p = 11$ in Table 3 are needed to begin the induction. As an example, for $q = 13$ we have $a = 2$ and $r = \frac{2aq-1}{3} = 17$, and the arithmetic progression $78k + 17$ contains $875 = 5^3 \cdot 7$. $\square$

5. Concluding remarks

The proofs in this paper are computer-intensive. Computer experimentation played an important role in the discovery of the patterns underlying the induction. This included the efficacy of using multipliers to eliminate congruence classes $(\bmod 2^k)$ in Section 2, and in uncovering the existence of the "intractable" residue class $-1 \pmod{2^k}$. If one had studied the problem without using the computer, the "intractable" case $-1 \pmod{2^k}$ could have been uncovered first, and this might have discouraged further investigation of this proof approach. It was also important to have the evidence detailed in [3], which provided a strong element of confidence in the truth of the weak $3x+1$ conjecture and wild numbers conjecture.

Extensive computations were needed to find the data in the tables. Once found, this data in the tables provides "succinct certificates" for checking correctness of the congruence class properties, which can be verified by hand. Similarly the induction step is in principle checkable by hand.

The proof methods developed in this paper should apply more generally in determining the integers in various multiplicative semigroups of rationals having a similar nature.

Acknowledgements

The authors thank Allan Wilks for helpful computations concerning the wild semigroup in Lemma 3.1. They thank the reviewer for many useful comments, including a suggested revision of the proof of Lemma 3.2 incorporated in this version. The second author did most of his work on this paper while employed at AT&T Labs.

References

[1] H. Farkas, Variants of the $3N+1$ conjecture and multiplicative semigroups, in: M. Entov, Y. Pinchover, M. Sageev (Eds.), Geometry, Spectral Theory, Groups and Dynamics: Proceedings in Memory of Robert Brooks, Contemporary Mathematics, vol. 387, American Mathematical Society, Providence, RI, 2005.

[2] J.C. Lagarias, The $3x + 1$ problem and its generalizations, Amer. Math. Monthly 92 (1985) 3–21.

[3] J.C. Lagarias, Wild and Wooley numbers, Amer. Math. Monthly 113 (2006), to appear.

[4] T. Oliveira e Silva, Maximum excursion and stopping time record-holders for the $3x + 1$ problem: computational results, Math. Comp. 68 (1) (1999) 371–384 (see website: http://www.ieeta.pt/ tos/3x+1.html for additional computations).
[5] J.B. Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, Illinois J. Math. 6 (1962) 64–94.
[6] P. Schogt, The Wild Numbers, Four Walls Eight Windows, New York, 1998.
[7] G.J. Wirsching, The Dynamical System on the Natural Numbers Generated by the $3n + 1$ Function, Lecture Notes in Mathematics, vol. 1681, Springer, Berlin, 1988.