Wild and Wooley Numbers

Jeffrey C. Lagarias

1. INTRODUCTION. This paper studies a problem in multiplicative number theory originating from a weakened form of the $3x + 1$ problem. Let $\mathcal{W}_0$ signify the multiplicative semigroup generated by the set of rationals

$$\left\{ \frac{3n+2}{2n+1} : n \geq 0 \right\},$$

which is the set

$$\left\{ \frac{2}{1}, \frac{5}{3}, \frac{8}{5}, \frac{11}{7}, \ldots \right\}.$$

That is, $\mathcal{W}_0$ is the set of all finite products of the generators, allowing repetitions. Let $\mathcal{W}$ signify the larger multiplicative semigroup generated by $\mathcal{W}_0$ together with $\{1/2\}$. We call $\mathcal{W}$ the *wild semigroup* and $\mathcal{W}_0$ the *Wooley semigroup*. The question we consider is: Which integers belong to these semigroups?

The sets of integer elements $\mathcal{W}(\mathbb{Z}) := \mathcal{W} \cap \mathbb{Z}$ and $\mathcal{W}_0(\mathbb{Z}) := \mathcal{W}_0 \cap \mathbb{Z}$ themselves form multiplicative semigroups, which we term the *wild integer semigroup* and *Wooley integer semigroup*, and we refer to their members as "wild integers" and "Wooley integers," respectively. We have the immediate implication that each Wooley integer is a wild integer, but the converse need not hold. The Wooley semigroup $\mathcal{W}_0$ is a semigroup without unit, whereas the wild semigroup $\mathcal{W}$ is a semigroup with unit, and the semigroups $\mathcal{W}_0(\mathbb{Z})$ and $\mathcal{W}(\mathbb{Z})$ inherit these properties. Our particular choice of terminology is explained at the end of the introduction.

An *irreducible element* of a commutative semigroup $\mathcal{T}$ is one that cannot be written as a product of two nonunits (noninvertible elements) in the semigroup (see Gilmer [**8**, sec. 1.6]). We call the irreducible elements of the wild integer semigroup *wild numbers*; similarly we christen the irreducible elements of the Wooley integer semigroup *Wooley numbers*. Thus the wild numbers are a subset of the wild integers, and the Wooley numbers are a subset of the Wooley integers.

It is immediately evident that 2 is both a wild number and a Wooley number. It is also easy to show that 3 is not a wild number, hence not a Wooley number. However the nature of other wild numbers or Wooley numbers is less apparent. The object of this paper is to determine properties of wild numbers and Wooley numbers. It turns out that the Wooley numbers have a complicated and not completely understood structure; by comparison, the wild numbers have a reasonably simple description.

The wild numbers and the Wooley numbers differ in some significant ways. An odd integer w is in the wild integer semigroup if and only if there is a nonnegative integer j such that $2^j w$ is in the Wooley integer semigroup. At the level of irreducible elements, we infer that if w is a wild number, then $2^j w$ is a Wooley number for some j; if $2^j w$ is a Wooley number, however, we cannot (currently) decide whether w must be a wild number. At first glance the Wooley numbers seem to be the simpler objects from a computational perspective. In section 2 we show that there is an effectively computable procedure for deciding whether a given rational number r belongs to $\mathcal{W}_0$.

This leads to an effectively computable procedure to determine whether an integer is a Wooley integer and, if so, whether it is a Wooley number.

In contrast, it is not immediately apparent if there exists an algorithm for recognizing whether a given integer is a wild integer. For a general multiplicative semigroup generated by a recursive set of rational numbers, it seems plausible that the problem of recognizing which integers belong to the semigroup is sometimes an undecidable problem. We refer the reader to Dyson [**5**, Theorem 3] for analogous undecidability results for commutative semigroups. On the other hand, if one were able to characterize directly all members of a semigroup, this could lead to a decision procedure. The "Wild Numbers Conjecture" to be stated shortly provides such a characterization for the wild numbers.

There is a strong connection between these problems and a weakened form of the $3x + 1$ problem, which motivated their study. This notorious problem (see [**11**] or [**17**]) is concerned with the iteration of the function $T : \mathbb{N} \to \mathbb{N}$ defined by

$$T(x) = \begin{cases} \dfrac{3x+1}{2} & \text{if } x \equiv 1 (\text{mod } 2), \\ \dfrac{x}{2} & \text{if } x \equiv 0 (\text{mod } 2). \end{cases}$$

The $3x + 1$ conjecture asserts that for each positive integer n there exists an N such that $T^{(N)}(n) = 1$, where $T^{(N)} = T \circ T \circ \cdots \circ T$ (N terms) is the N-fold iterate of T. It has been verified for all n with $n \leq 10^{17}$, but remains an open problem.

The weakened version of the $3x + 1$ problem to which we alluded earlier was proposed by Herschel Farkas [**7**]. It reads as follows:

Weak $3X + 1$ Conjecture. *Consider the semigroup $\mathcal{S} := \mathcal{W}^{-1} = \{w^{-1} : w \in \mathcal{W}\}$ that is generated by*

$$\left\{ \frac{2n+1}{3n+2} : n \geq 0 \right\}$$

together with the integer 2*. Then $\mathcal{S}$ contains every positive integer.*

Farkas [**7**] observed that the truth of the $3x + 1$ conjecture implies the truth of the weak $3x + 1$ conjecture, because the steps of the $3x + 1$ iteration process can be encoded as certain products of generators using the semigroup multiplication in $\mathcal{S}$. However, there are products of generators in the semigroup $\mathcal{S}$ that do not correspond to the $3x + 1$ iteration, so the Farkas conjecture is potentially easier to resolve than the $3x + 1$ problem.

Investigation of the weak $3x + 1$ conjecture led to questions about the wild integer semigroup $\mathcal{W}(\mathbb{Z})$ as a possible aid in its proof. Conversely, the weak $3x + 1$ conjecture has very strong implications about wild numbers, as explained in section 3, that led to the formulation of the following conjecture:

Wild Numbers Conjecture. *The set of wild numbers consists of all prime numbers, excluding* 3*. Equivalently, the wild integer semigroup $\mathcal{W}(\mathbb{Z})$ consists of all positive integers m not divisible by* 3*.*

This paper studies properties of wild and Wooley integers that bear on these questions. In section 2 we study Wooley integers. We give an effectively computable algorithm for recognizing Wooley integers and Wooley numbers. Using this approach we show that 20 is a Wooley number. We also report computations finding various Wooley integers. We discuss the question of whether the Wooley integer semigroup $\mathcal{W}_0(\mathbb{Z})$ is a free commutative semigroup and present evidence suggesting that it is not. In contrast, in section 3 we show that the weak $3x + 1$ conjecture implies that the wild integer semigroup $\mathcal{W}(\mathbb{Z})$ is a free commutative semigroup with unit.

In section 3 we begin by demonstrating that there are infinitely many wild numbers. Then we show that the weak $3x + 1$ conjecture implies strong restrictions on wild numbers—indeed, we prove that it implies the wild numbers conjecture. This lends strong support to the conviction that the wild numbers conjecture must be true, since the weak $3x + 1$ conjecture itself would follow from the $3x + 1$ conjecture, for which there is extensive evidence. We also deduce a converse assertion to the effect that the wild numbers conjecture implies the weak $3x + 1$ conjecture. As a final result we show that these conjectures completely characterize the structure of the wild semigroup, which turns out to be quite tame.

Based on some of the results derived here, the weak $3x + 1$ conjecture and wild numbers conjecture were subsequently proved by David Applegate and the author in [**1**]. In section 4 we indicate some features of the proof and formulate some open problems about Wooley numbers, which remain mysterious.

The terms "wild semigroup" and "wild number" were suggested by the novel *The Wild Numbers* by Philibert Schogt [**14**]. The novel chronicles the efforts of a mathematics professor to solve the (fictitious) "Beauregard Wild Numbers Problem," while dealing with the ups and downs of life in a university mathematics department. The semigroup problem posed here has some striking resemblances to the information given about the Beaureguard wild numbers problem. Beaureguard wild numbers are described in the novel as certain integers produced at the end of a sequence of elementary operations that involve noninteger rationals at the intermediate steps. Here the semigroup products of $\mathcal{W}$ giving an element of $\mathcal{W}(\mathbb{Z})$ generally consist of rationals whose partial products typically become integers only at the last step. The novel also states [**14**, pp. 34, 37] that 2 is a Beaureguard wild number but 3 is not, and that 67 and 4769 are Beaureguard wild numbers. The wild numbers defined here reproduce nearly all this empirical data. (The one exception is $4769 = 169 \cdot 253$, which belongs to the wild integer semigroup $\mathcal{W}(\mathbb{Z})$ but is not a wild number as we define it. Perhaps the novel has a misprint for 4759 or 4789 or 4967, all primes.) The Beaureguard wild numbers problem is to decide whether there are infinitely many wild numbers [**14**, p. 35]. The terms "Wooley semigroup" and "Wooley numbers" are named after Trevor D. Wooley, in honor of his work in related areas of number theory (for example, [**3**]).

Aside from the definitions of "wild numbers" made in this paper, there have been other concepts of "wild numbers" that possess some of the properties indicated in the foregoing discussion. We refer to sequence A58883 in the *Encyclopedia of Integer Sequences* maintained by Neil Sloane [**15**], and six versions of "pseudo-wild numbers" cited there. The Beaureguard wild numbers problem in Schogt's novel seems to involve iteration, which is not directly present in our semigroup problem. Some iteration problems with a similar flavor to the wild numbers problem come from the "approximate multiplication" maps studied in Lagarias and Sloane [**12**]. A typical example is the map $f(x) = \frac{4}{3}\lceil x \rceil$. The question studied in [**12**] asks whether it is true that, for each positive integer n, some iterate $f^{(N)}(n)$ is again an integer. This iteration problem thus produces a sequence of noninteger rational numbers terminating in an integer. It is currently unsolved and seems likely to be difficult.

2. WOOLEY NUMBERS. We show that the Wooley semigroup $\mathcal{W}_0$ is a recursive semigroup.

Theorem 2.1. *There is an effectively computable procedure that when given any positive rational r determines whether or not it belongs to the Wooley semigroup $\mathcal{W}_0$, and if it does, exhibits it as a product of generators.*

Proof. We cannot represent r unless it is a positive rational number having an odd denominator (in lowest terms). Let

$$g(n) = \frac{3n+2}{2n+1}$$

denote the nth generator of the semigroup $\mathcal{W}_0$, and suppose that

$$r = \prod_{i=1}^{m} g(n_i)$$

with $n_1 \leq n_2 \leq \cdots \leq n_m$. We first bound m above. In fact, since $g(n) > 3/2$ for each n, we must have $r > (3/2)^m$, which delivers an upper bound for m.

Now let m be fixed. We find an upper bound for n_1. We have $r > (3/2)^m$, so

$$r = \left(\frac{3}{2} + \epsilon\right)^m$$

with

$$\epsilon = r^{1/m} - \frac{3}{2} > 0.$$

We claim that $n_1 \leq 1/\epsilon$. If not, then

$$g(n_1) = \frac{3n_1+2}{2n_1+1} = \frac{3}{2} + \frac{1/2}{2n_1+1} < \frac{3}{2} + \epsilon.$$

Since $g(n)$ is a decreasing function of n, we would have

$$r = \prod_{i=1}^{m} g(n_i) \leq g(n_1)^m < \left(\frac{3}{2} + \epsilon\right)^m = r,$$

a contradiction that proves the claim.

Once n_1 is chosen, we can divide out $g(n_1)$ to create a new problem of the same kind with a smaller value $r' = r(g(n_1))^{-1} < 2r/3$, where we ask for a representation using a product of exactly $m-1$ generators. We then show that there is a finite set of choices for n_2, obtaining in the process an explicit upper bound for n_2 as a function of r, m, and n_1. Proceeding by induction on m, we discover that the total allowable set of choices is finite, with an effectively computable upper bound. Searching all of them yields either a relation certifying that r belongs to $\mathcal{W}_0$ or a proof that r does not belong to $\mathcal{W}_0$. ∎

We can carry out this procedure in the simplest cases.

Example. The integers 5 and 10 are not Wooley integers, but 20 is a Wooley integer. As a consequence, 20 is a Wooley number.

Proof. Suppose that 5 were a product of generators of $\mathcal{W}_0$. Since 2 cannot be cancelled from the numerator of any product or 3 from its denominator, any representation of 5 could not use the generators $g(0) = 2/1$, $g(1) = 5/3$, or $g(2) = 8/5$. Any product of three of the remaining generators is no larger than

$$g(3)^3 = \left(\frac{11}{7}\right)^3 = \frac{1331}{243} < 5,$$

so any representation of 5 necessarily includes at least four factors from the generating set of $\mathcal{W}_0$. However, any such product is larger than

$$\left(\frac{3}{2}\right)^4 = \frac{81}{16} > 5,$$

a contradiction.

Suppose that 10 were a product of generators of $\mathcal{W}_0$. Any representation of 10 would use at most five generators, since

$$\left(\frac{3}{2}\right)^6 = \frac{729}{64} > 10.$$

A representation of 10 could not use the generator $2/1$, for if it did this fraction could be removed, yielding a representation of 5, a contradiction. Also, $5/3$ and $8/5$ could not arise as factors because they would add uncancellable terms 3 and 2^3, so the fraction of largest size that could appear in any product is again $11/7$. However

$$\left(\frac{11}{7}\right)^5 < 10,$$

so there can be no such representation.

The number 20 can be expressed as follows:

$$\begin{aligned} 20 &= g(3)^2 \cdot g(5) \cdot g(8) \cdot g(27) \cdot g(32) \cdot g(41) \\ &= \left(\frac{11}{7}\right)^2 \left(\frac{17}{11}\right) \left(\frac{26}{17}\right) \left(\frac{83}{55}\right) \left(\frac{98}{65}\right) \left(\frac{125}{83}\right). \end{aligned}$$

This confirms that 20 belongs to $\mathcal{W}_0$, making it a Wooley integer. To see that 20 is a Wooley number, note that if it is not irreducible, then $20 = n_1 n_2$, where n_1 and n_2 belong to $\mathcal{W}_0(\mathbb{Z})$. At least one of n_1 or n_2 would then be divisible by 5, but the only possibilities are 5 and 10, which have already been ruled out. ■

The algorithm of Theorem 2.1 appears to be very slow, requiring at least exponential time. However, one can find Wooley integers by less exhaustive methods. Table 1 presents additional Wooley integers with identities certifying their membership in $\mathcal{W}_0(\mathbb{Z})$ for certain numbers of the form $2^k p$, where p is prime such that $5 \le p < 50$. These identities were found by Allan Wilks via computer search. Wilks's search used certain heuristics and did not decide whether these products give the minimal power of 2 possible. As a result we can say only that the entries of the table are Wooley integers, not necessarily Wooley numbers.

A commutative semigroup with or without unit 1 is said to be *a free commutative semigroup* if every element of the semigroup except 1 can be factored uniquely (up to ordering of the factors) into a product of irreducible elements. Many such semigroups

Table 1. Members of the Wooley integer semigroup $\mathcal{W}_0(\mathbb{Z})$.

$$2^2 \cdot 5 = \left(\frac{11}{7}\right)^2 \cdot \frac{17}{11} \cdot \frac{26}{17} \cdot \frac{83}{55} \cdot \frac{98}{65} \cdot \frac{125}{83} = g(3)^2 \cdot g(5) \cdot g(8) \cdot g(27) \cdot g(32) \cdot g(41)$$

$$\begin{aligned} 2^2 \cdot 7 &= \frac{11}{7} \cdot \frac{26}{17} \cdot \frac{35}{23} \cdot \frac{215}{143} \cdot \frac{299}{199} \cdot \frac{323}{215} \cdot \frac{371}{247} \cdot \frac{398}{265} \\ &= g(3) \cdot g(8) \cdot g(11) \cdot g(71) \cdot g(99) \cdot g(107) \cdot g(123) \cdot g(132) \end{aligned}$$

$$\begin{aligned} 2^2 \cdot 11 &= \left(\frac{11}{7}\right)^2 \cdot \frac{26}{17} \cdot \frac{35}{23} \cdot \frac{215}{143} \cdot \frac{299}{199} \cdot \frac{323}{215} \cdot \frac{371}{247} \cdot \frac{398}{265} \\ &= g(3)^2 \cdot g(8) \cdot g(11) \cdot g(71) \cdot g(99) \cdot g(107) \cdot g(123) \cdot g(132) \end{aligned}$$

$$\begin{aligned} 2^3 \cdot 13 &= \left(\frac{11}{7}\right)^2 \cdot \left(\frac{17}{11}\right)^3 \cdot \left(\frac{26}{17}\right)^2 \cdot \frac{35}{23} \cdot \frac{215}{143} \cdot \frac{299}{199} \cdot \frac{323}{215} \cdot \frac{371}{247} \cdot \frac{398}{265} \\ &= g(3)^2 \cdot g(5)^3 \cdot g(8)^2 \cdot g(11) \cdot g(71) \cdot g(99) g(107) \cdot g(123) \cdot g(132) \end{aligned}$$

$$\begin{aligned} 2^2 \cdot 17 &= \left(\frac{11}{7}\right)^2 \cdot \frac{17}{11} \cdot \frac{26}{17} \cdot \frac{83}{55} \cdot \frac{98}{65} \cdot \frac{125}{83} \cdot \frac{143}{95} \cdot \frac{215}{143} \cdot \frac{323}{215} \\ &= g(3)^2 \cdot g(5) \cdot g(8) \cdot g(27) \cdot g(32) \cdot g(41) \cdot g(47) \cdot g(71) \cdot g(107) \end{aligned}$$

$$\begin{aligned} 2^5 \cdot 19 &= \left(\frac{11}{7}\right)^4 \cdot \left(\frac{17}{11}\right)^2 \cdot \left(\frac{26}{17}\right)^2 \cdot \frac{38}{25} \cdot \left(\frac{83}{55}\right)^2 \cdot \left(\frac{98}{65}\right)^2 \cdot \left(\frac{125}{83}\right)^2 \\ &= g(3)^4 \cdot g(5)^2 \cdot g(8)^2 \cdot g(12) \cdot g(27)^2 \cdot g(32)^2 \cdot g(41)^2 \end{aligned}$$

$$\begin{aligned} 2^5 \cdot 23 &= \frac{11}{7} \cdot \frac{26}{17} \cdot \frac{35}{23} \cdot \frac{47}{31} \cdot \frac{137}{91} \cdot \frac{155}{103} \cdot \frac{206}{137} \cdot \frac{215}{143} \cdot \left(\frac{299}{199}\right)^2 \cdot \frac{323}{215} \cdot \frac{353}{235} \cdot \frac{371}{247} \cdot \left(\frac{398}{265}\right)^2 \cdot \frac{530}{353} \\ &= g(3) \cdot g(8) \cdot g(11) \cdot g(15) \cdot g(45) \cdot g(51) \cdot g(68) \cdot g(71) \\ &\quad \cdot g(99)^2 \cdot g(107) \cdot g(117) \cdot g(123) \cdot g(132)^2 \cdot g(176) \end{aligned}$$

$$\begin{aligned} 2^5 \cdot 29 &= \left(\frac{11}{7}\right)^4 \cdot \left(\frac{17}{11}\right)^2 \cdot \left(\frac{26}{17}\right)^2 \cdot \frac{29}{19} \cdot \frac{38}{25} \cdot \left(\frac{83}{55}\right)^2 \cdot \left(\frac{98}{65}\right)^2 \cdot \left(\frac{125}{83}\right)^2 \\ &= g(3)^4 \cdot g(5)^2 \cdot g(8)^2 \cdot g(9) \cdot g(12) \cdot g(27)^2 \cdot g(32)^2 \cdot g(41)^2 \end{aligned}$$

$$\begin{aligned} 2^{11} \cdot 31 &= \left(\frac{11}{7}\right)^6 \cdot \left(\frac{17}{11}\right)^3 \cdot \frac{29}{19} \cdot \frac{38}{25} \cdot \frac{62}{41} \cdot \left(\frac{83}{55}\right)^3 \cdot \left(\frac{98}{65}\right)^3 \cdot \left(\frac{125}{83}\right)^3 \cdot \frac{164}{109} \cdot \frac{218}{145} \\ &= g(3)^6 \cdot g(5)^3 \cdot g(8)^3 \cdot g(9) \cdot g(12) \cdot g(20) \cdot g(27)^3 \cdot g(32)^3 \cdot g(41)^3 \cdot g(54) \cdot g(72) \end{aligned}$$

$$\begin{aligned} 2^5 \cdot 37 &= \left(\frac{11}{7}\right)^2 \cdot \left(\frac{26}{17}\right)^2 \cdot \left(\frac{35}{23}\right)^2 \cdot \frac{74}{49} \cdot \left(\frac{215}{143}\right)^2 \cdot \left(\frac{299}{199}\right)^2 \cdot \left(\frac{323}{215}\right)^2 \cdot \left(\frac{371}{247}\right)^2 \cdot \left(\frac{398}{265}\right)^2 \\ &= g(3)^2 \cdot g(8)^2 \cdot g(11)^2 \cdot g(24) \cdot g(71)^2 \cdot g(99)^2 \cdot g(107)^2 \cdot g(123)^2 \cdot g(132)^2 \end{aligned}$$

$$\begin{aligned} 2^{10} \cdot 41 &= \left(\frac{11}{7}\right)^6 \cdot \left(\frac{17}{11}\right)^3 \cdot \left(\frac{26}{17}\right)^3 \cdot \frac{29}{19} \cdot \frac{38}{25} \cdot \left(\frac{83}{55}\right)^3 \cdot \left(\frac{98}{65}\right)^3 \cdot \left(\frac{125}{83}\right)^3 \cdot \frac{164}{109} \cdot \frac{218}{145} \\ &= g(3)^6 \cdot g(5)^3 \cdot g(8)^3 \cdot g(9) \cdot g(12) \cdot g(27)^3 \cdot g(32)^3 \cdot g(41)^3 \cdot g(54) \cdot g(72) \end{aligned}$$

$$\begin{aligned} 2^{11} \cdot 43 &= \left(\frac{11}{7}\right)^5 \cdot \left(\frac{17}{11}\right)^2 \cdot \left(\frac{26}{17}\right)^3 \cdot \frac{29}{19} \cdot \frac{35}{23} \cdot \frac{38}{25} \cdot \left(\frac{83}{55}\right)^2 \cdot \left(\frac{98}{65}\right)^2 \cdot \left(\frac{125}{87}\right)^2 \cdot \frac{215}{143} \\ &\quad \cdot \frac{299}{199} \cdot \frac{305}{203} \cdot \frac{323}{215} \cdot \frac{344}{229} \cdot \frac{371}{247} \cdot \frac{398}{265} \cdot \frac{458}{305} \\ &= g(3)^5 \cdot g(5)^2 \cdot g(8)^3 \cdot g(9) \cdot g(11) \cdot g(12) \cdot g(27)^2 \cdot g(32)^2 \cdot g(41)^2 \cdot g(71) \\ &\quad \cdot g(99) \cdot g(101) \cdot g(107) \cdot g(114) \cdot g(123) \cdot g(132) \cdot g(152) \end{aligned}$$

$$\begin{aligned} 2^{11} \cdot 47 &= \left(\frac{11}{7}\right)^6 \cdot \left(\frac{17}{11}\right)^3 \cdot \frac{29}{19} \cdot \frac{38}{25} \cdot \frac{47}{31} \cdot \frac{62}{41} \cdot \left(\frac{83}{55}\right)^3 \cdot \left(\frac{98}{65}\right)^3 \cdot \left(\frac{125}{83}\right)^3 \cdot \frac{164}{109} \cdot \frac{218}{145} \\ &= g(3)^6 \cdot g(5)^3 \cdot g(8)^3 \cdot g(9) \cdot g(12) \cdot g(15) \cdot g(20) \cdot g(27)^3 \cdot g(32)^3 \cdot g(41)^3 \cdot g(54) \cdot g(72) \end{aligned}$$

arise in number theory (see Knopfmacher [**10**]). Does the Wooley integer semigroup $\mathcal{W}_0(\mathbb{Z})$ have unique factorization? We leave this question unresolved, but it seems likely that it does not. To show this it would suffice to find a Wooley number that contained two distinct odd prime factors. A suggestive example is provided by the number $2^6 \cdot 31 \cdot 41$, which is a Wooley integer expressible in terms of generators by

$$2^6 \cdot 31 \cdot 41 = g(423) \cdot (2^2 \cdot 7)(2^2 \cdot 11)^2,$$

since $g(423) = 1271/847 = (31 \cdot 41)/(7 \cdot 11^2)$, and both $2^2 \cdot 7$ and $2^2 \cdot 11$ belong to $\mathcal{W}_0(\mathbb{Z})$ according to Table 1. If there were a Wooley number of the form $2^c \cdot 31 \cdot 41$, then the semigroup $\mathcal{W}_0(\mathbb{Z})$ would not be a free commutative semigroup, because there would exist four irreducible elements of $\mathcal{W}_0(\mathbb{Z})$—2, $2^a \cdot 31$, $2^b \cdot 41$, $2^c \cdot 31 \cdot 41$—leading to a nonunique factorization of $M = 2^{a+b+c} \cdot 31 \cdot 41$. Such a Wooley number will exist unless there are Wooley integers of form $2^a \cdot 31$ and $2^b \cdot 41$ with $a + b \leq 6$, and this possibility can be tested algorithmically, according to Theorem 2.1.

3. WILD NUMBERS. We begin by showing that there are infinitely many wild numbers.

Theorem 3.1. *The semigroup of wild integers contains infinitely many irreducible elements (i.e., there are infinitely many wild numbers).*

Proof. For

$$n = \frac{5^k - 1}{2}$$

we have

$$g(n) = \frac{\frac{1}{2}(3 \cdot 5^k + 1)}{5^k}.$$

Example 2.2 shows that $2^2 \cdot 5$ is a Wooley number, which implies that 5 is a wild integer, and it is a wild number since it is prime. We conclude that

$$h(k) := \frac{1}{2}(3 \cdot 5^k + 1) = g(n) \cdot 5^k$$

belongs to $\mathcal{W}(\mathbb{Z})$ for each positive integer k.

The sequence $\{h(k) : k \geq 1\}$ satisfies a homogeneous second-order linear recurrence, namely,

$$h(k) = 6h(k-1) - 5h(k-2).$$

This sequence is nondegenerate in the sense of Ward [**16**] (i.e., it does not satisfy a first-order linear recurrence). Accordingly, by the main result of Ward [**16**] the sequence $\{h(k)\}$ contains an infinite number of distinct prime divisors (i.e., the set D of primes p that divide $h(k)$ for at least one k is infinite).

We now argue by contradiction that $\mathcal{W}(\mathbb{Z})$ contains infinitely many irreducible elements. If not, there would exist some prime p in the infinite set D that did not divide any irreducible element. This prime p divides some $h(k)$, which belongs to $\mathcal{W}(\mathbb{Z})$, so there exists a smallest element m of $\mathcal{W}(\mathbb{Z})$ that is divisible by p. This element m is necessarily irreducible, for if not there would be a smaller integer in $\mathcal{W}(\mathbb{Z})$ divisible by p. This gives a contradiction. ■

We obtain much stronger results about the structure of the wild integer semigroup if we assume the truth of the weak $3x + 1$ conjecture.

Theorem 3.2. *Suppose that the weak $3x + 1$ conjecture holds. Then the wild integer semigroup $\mathcal{W}(\mathbb{Z})$ is a free commutative semigroup whose set of generators $\mathcal{P}$ consists entirely of primes. In other words, all wild numbers are prime numbers.*

Proof. The semigroup $\mathcal{W}(\mathbb{Z})$ contains all powers of 2. Also, since 2 is invertible in $\mathcal{W}$, it can be cancelled from all other generators, which therefore must be odd integers. However, the weak $3x + 1$ conjecture says that if n belongs to $\mathcal{W}$, then so does n/k for any positive integer k. Thus, if a composite number n lies in $\mathcal{W}(\mathbb{Z})$, so do all of its prime divisors. It follows that all generators of $\mathcal{W}(\mathbb{Z})$ are primes. The semigroup $\mathcal{W}(\mathbb{Z})$ is now a free commutative semigroup as a consequence of the unique prime factorization of integers. ■

In general, we can certify that a given prime number p is a wild number by finding some j such that $2^j p$ is a Wooley number. For example, 67 is a wild number since $2^{12} \cdot 67$ is a Wooley number. The latter assertion is a consequence of the identity

$$\frac{2^5 \cdot 67}{5 \cdot 37} = g(29) \cdot g(44) \cdot g(69) \cdot g(78) \cdot g(92) \cdot g(104)$$

and the fact, established earlier, that $2^2 \cdot 5$ and $2^5 \cdot 37$ are Wooley numbers.

We next show that the weak $3x + 1$ conjecture implies the wild numbers conjecture.

Theorem 3.3. *If the weak $3x + 1$ conjecture holds, then the wild numbers conjecture is true.*

Proof. We prove the wild numbers conjecture by induction on the nth prime, call it q. The induction hypothesis asserts that all smaller primes except 3 belong to the wild integer semigroup. We call an integer *Y-smooth* if all its prime factors are strictly smaller than Y. Because the wild integers form a semigroup, the induction hypothesis tells us that all q-smooth numbers not divisible by 3 are wild integers. In particular, all integers smaller than q and not divisible by 3 are wild integers. To complete the induction step it suffices to show that q is a wild integer. If so, its primality guarantees that it is irreducible, hence that it is a wild number.

To show that q is a wild integer it suffices to find some multiple mq that is a wild integer, for the weak $3x + 1$ conjecture implies that $1/m$ belongs to the wild semigroup $\mathcal{W}$, making $q = (mq)/m$ a wild integer. We wish to find mq of the form $mq = 3n + 2$ such that $2n + 1$ is a q-smooth number not divisible by 3. If so, then

$$mq = \frac{3n + 2}{2n + 1}(2n + 1)$$

will belong to $\mathcal{W}$, and the desired result will follow. The requirement $mq = 3n + 2$ puts m in a certain residue class modulo 3, and by imposing a condition on m modulo 9 we can guarantee that $2n + 1 \not\equiv 0 \pmod 3$. The resulting integers $2n + 1$ then fall into an arithmetic progression of numbers congruent to r modulo $6q$, with $gcd(r, 6q) =$

1, and we arrive at a special case of the well-studied arithmetic question of finding "smooth numbers" in an arithmetic progression.

We recall general facts on the distribution of "smooth numbers" up to X, namely, those numbers below X having all prime factors smaller than a given bound Y (see, for example, Hildebrand and Tenenbaum [**9**]). Smooth numbers for an appropriate choice of Y play an important role in the design and performance of the fastest known algorithms for factoring large numbers (see Pomerance [**13**]). It is known that the number of integers smaller than X that have all their prime factors below a cutoff value $Y = X^\alpha$ for any fixed α have asymptotically a positive density $\rho(\alpha)Y$, where $\rho(u)$ is a strictly positive function, the *Dickman function*. This function is given by the solution to a certain difference-differential equation and $\rho(u) \approx u^{-u}$. (Here $u = (\log Y)/(\log X)$.) Balog and Pomerance [**2**] carry these bounds over to count the number of Y-smooth numbers in arithmetic progressions modulo N, and their results give an asymptotic formula valid for Y over a large range. In particular, choosing $X \approx q^2$, $N = 6q$, and $Y = q$, which corresponds to $\alpha = 1/2$, one can deduce from their results that the number of q-smooth integers in the first q terms of the arithmetic progression r modulo $6q$ is nonzero whenever $q > C_0$ for some constant C_0 that is, in principle, computable. However, C_0 is not easy to compute, nor is it likely to be small.

To complete our argument we need only demonstrate the existence of a single q-smooth number in the given arithmetic progression r modulo $6q$. This permits us to sidestep the results of Balog and Pomerance and to use instead a direct combinatorial argument. It rests on the observation that if more than half the invertible residue classes modulo N contain Y-smooth numbers, then (by the pigeonhole principle) every invertible residue class r modulo N occurs as a product of two of these residue classes, and consequently contains a Y-smooth number that is the product of Y-smooth numbers from these classes. In more detail, let Σ denote the set of invertible residue classes s modulo N that contain Y-smooth integers. Suppose that r is an arbitrary invertible residue class modulo N. We now define Σ' to consist of those residue classes s' modulo N given by

$$s' \equiv r \cdot s^{-1} (\text{mod } N),$$

where s belongs to Σ. Certainly $|\Sigma'| = |\Sigma|$ and since $|\Sigma|$ and $|\Sigma'|$ exceed half the number of invertible residue classes, there must be some s' in $\Sigma \cap \Sigma'$. Now $r \equiv ss' (\text{mod } N)$, and taking S and S' to be Y-smooth numbers in the classes s and s' modulo N, respectively, we find that $S \cdot S'$ is a Y-smooth number in the class r modulo N.

In our case we have $N = 6q$, which has $\phi(N) = 2(q-1)$ invertible residue classes, and all of these residue classes consist of numbers not divisible by 3. It suffices to demonstrate that more than $q-1$ of these invertible classes contain q-smooth numbers. We show that the set of such residue classes whose least positive residue is smooth exceeds $q-1$ for all sufficiently large q. Now every integer less than $6q$ and relatively prime to $6q$ has all its prime factors smaller than q, except for primes p' with $q < p' < 6q$ and integers of the form $5p'$ with $q \le p' < 6q/5$. Since the number of primes below x is

$$O\left(\frac{x}{\log x}\right),$$

there are at most

$$O\left(\frac{q}{\log q}\right)$$

such integers, hence at least

$$2(q-1) - O\left(\frac{q}{\log q}\right)$$

invertible classes have a least residue that is q-smooth. For large q this gives the result, and by obtaining explicit numerical bounds for the remainder term it is possible to prove that when $q > 10^4$ more than half the invertible residue classes modulo $6q$ are q-smooth. (Such bounds are derived in [**1**].)

Now we can prove the wild numbers conjecture by induction, under the assumption that the weak $3x + 1$ conjecture is valid, with the base case consisting of checking all q such that $q < 10^4$. The base case can be checked by computer. In fact, it suffices to use Table 1 when $q < 50$ and for q in the range $50 < q < 10^4$ to have the computer find directly a smooth number in a suitable arithmetic progression modulo $6q$. ■

Since the $3x + 1$ conjecture appears to be true, Theorem 3.3 provides a powerful argument in favor of the wild numbers conjecture. On the other hand, we have a converse implication:

Theorem 3.4. *If the wild numbers conjecture is true, then the weak* $3x + 1$ *conjecture holds.*

Proof. This implication is proved with an argument similar to the one that established Theorem 3.3. We proceed by induction on the nth prime q, assuming that all primes below q (including 3) belong to the inverse semigroup $\mathcal{W}^{-1}$. We now consider multiples mq, where $m \equiv 1 \pmod 6$. The wild numbers conjecture implies that all such integers are in the wild semigroup $\mathcal{W}$. Writing $mq = 2n + 1$, we look for a case in which $3n + 2$ is a q-smooth number. Expressing m as $m = 6k + 1$, we have

$$3n + 2 = 9kq + \frac{3q+1}{2},$$

which is an arithmetic progression modulo $9q$. As in the earlier result, it suffices to show that for all sufficiently large q more than half of the invertible residue classes modulo $9q$ in the interval $[1, q-1]$ have least positive residues that are q-smooth numbers, which then implies that each arithmetic progression for an invertible residue class modulo $9q$ contains a q-smooth integer smaller than $81q^2$. This holds when $q > 10^5$. Since the $3x + 1$ conjecture has been verified up to 10^5, the base case of the induction is already done. ■

Our final result points out that the truth of the weak $3x + 1$ conjecture completely determines the structure of the wild semigroup $\mathcal{W}$.

Theorem 3.5. *If the weak* $3x + 1$ *conjecture is true, then the wild semigroup* $\mathcal{W}$ *consists of all positive rational numbers* a/b *with* $gcd(a, 3b) = 1$.

Proof. The weak $3x + 1$ conjecture implies that $\mathcal{W}$ contains all fractions $1/p$, where p is prime. By Theorem 3.3 this conjecture implies the wild numbers conjecture, which ensures that $\mathcal{W}$ contains all primes p different from 3. We observed earlier that any rational number $r = a/b$ in lowest terms that belongs to $\mathcal{W}$ has numerator a relatively prime to 3. This gives the result. ■

Theorem 3.5 provides a simple effective decision procedure for membership of a given rational number r in the wild semigroup $\mathcal{W}$, provided that the weak $3x+1$ conjecture is proved.

4. CONCLUDING REMARKS. The results of section 3 demonstrate that the wild numbers conjecture and the weak $3x+1$ conjecture are intertwined: each is implied by the other. David Applegate and the author [**1**] have recently been able to prove both conjectures simultaneously, via a bootstrap induction procedure that uses the truth of one of the conjectures on an interval to extend the truth of the other to a larger interval, and vice versa. The argument of Theorem 3.3 (respectively, Theorem 3.5) provides a way to extend the truth of the conjecture in one direction, provided it holds on a sufficient initial interval in the other. We do not, however, know a way to use the arguments of these theorems simultaneously in both directions. Fortunately, there is another systematic way to find representations of many integers n in the inverse semigroup $\mathcal{W}^{-1}$, which is to iterate the $3x+1$ map starting with n. The argument in [**1**] takes advantage of this fact in constructing the "other" direction of the bootstrap induction. There is an apparent asymmetry in the two directions, in that we are not aware of any dynamical system associated with the wild semigroup that produces relations generating the integers in the wild integer semigroup $\mathcal{W}(\mathbb{Z})$ that is analogous to the use of the $3x+1$ iteration in the inverse semigroup $\mathcal{W}^{-1}$.

The original motivation for studying the wild semigroup came from the weak $3x+1$ conjecture, but the Wooley semigroup $\mathcal{W}_0$ that arose in the process seems interesting in its own right. The Wooley integer semigroup $\mathcal{W}_0(\mathbb{Z})$ appears to be a more complicated object than the wild integer semigroup $\mathcal{W}(\mathbb{Z})$, and there remain many open questions about Wooley integers. One question already raised in section 2 asks whether Wooley integers have unique factorization into irreducibles (i.e., whether the Wooley integer semigroup is a free commutative semigroup). A second question concerns, for each prime p, the behavior of the minimal power $e(p)$ necessary to place $2^{e(p)}p$ in the Wooley integer semigroup. It seems plausible that $e(p)$ is unbounded. The truth of the wild numbers conjecture implies that each number $e(p)$ is finite, so in view of its proof in [**1**], this question is well posed. A third question asks: How does the counting function of the Wooley integers below x grow as $x \to \infty$?

As noted earlier, the wild numbers conjecture is named after the (fictitious) mathematical problem in Philibert Schogt's novel *The Wild Numbers*. In the novel the Beauregard wild numbers problem was presented as a famous unsolved problem, with a long and illustrious history. Its namesake here fails to have either of these attributes. The wild numbers conjecture has a short history, and the infinitude of the wild numbers was established by Theorem 3.1. Nevertheless, our terminology seems fitting, for the novel asserts that there is "a fundamental relationship between wild numbers and prime numbers" [**14**, p. 36], and the wild numbers of this paper coincide with the prime numbers, excluding 3. Understanding the behavior of prime numbers is one of the great quests of mathematics, with a history as long and impressive as one could hope for; see Derbyshire [**4**] or du Sautoy [**6**] for recent accounts.

ADDED IN PROOF. In work done at the University of Minnesota-Duluth REU program in Summer 2005, Ana Carlani answered a question raised in section 4. She showed that the Wooley integer semigroup does not have unique factorization into irreducibles.

ACKNOWLEDGMENTS Most of this work was done while I was at A.T. & T. Labs-Research, whom I thank for support. I am indebted to Allan Wilks for computing the data in Table 1 and to Jim Reeds for recommending the novel *The Wild Numbers*. I thank the reviewers for helpful comments.

1. D. Applegate and J. C. Lagarias, The $3x + 1$ semigroup, *J. Number Theory* (in press).
2. A. Balog and C. Pomerance, The distribution of smooth numbers in arithmetic progressions, *Proc. Amer. Math. Soc.* **115** (1992) 33–43.
3. A. Balog and T. D. Wooley, On strings of consecutive integers with no large prime factor, *J. Australian Math. Soc.* **64** (1998) 266–276.
4. J. Derbyshire, *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*, Joseph Henry Press, Washington, D.C., 2003.
5. V. H. Dyson, On the decision problem for theories of finite models, *Israel J. Math.* **2** (1964) 55–70.
6. M. du Sautoy, *The Music of the Primes: Searching to solve the greatest mystery in mathematics*, Harper-Collins, New York, 2003.
7. H. Farkas, Variants of the $3N + 1$ problem and multiplicative semigroups, in *Geometry, Spectral Theory, Groups, and Dynamics: Proceedings in Memory of Robert Brooks*, M. Entov, Y. Pinchover, and M. Sageev, eds., American Mathematical Society, Providence, 2005.
8. R. Gilmer, *Commutative Semigroup Rings*, University of Chicago Press, Chicago, 1984.
9. A. Hildebrand and G. Tenenbaum. Integers without large prime factors, J. Théorie des Nombres, Bordeaux **5** (1993) 411–484.
10. J. Knopfmacher, *Abstract Analytic Number Theory.* 2nd Ed., Dover Publications, New York 1990.
11. J. C. Lagarias, The $3X + 1$ problem and its generalizations, this MONTHLY **85** (1985) 1–21.
12. J. C. Lagarias and N. J. A. Sloane, Approximate squaring, *Experimental Math.* **13** (2004) 113–128.
13. C. Pomerance, The role of smooth numbers in number-theoretic algorithms, *Proc. Intl. Congress of Mathematicians* (Zurich, 1994), 2 vols., Birkhäuser, Basel 1995, pp. 411–422.
14. P. Schogt, *The Wild Numbers*, Four Walls Eight Windows, New York, 1998.
15. N. J. A. Sloane, *The On-line Encyclopedia of Integer Sequences*, published electronically at `http://www.research.att.com/~njas/sequences/`.
16. M. Ward, Prime divisors of second order recurring sequences, *Duke Math. J.* **21** (1954) 607–614.
17. G. J. Wirsching, *The Dynamical System Generated by the* $3n + 1$ *Function*, Lecture Notes in Math. no. 1681, Springer-Verlag, Berlin, 1998.

JEFFREY C. LAGARIAS received his Ph.D. in mathematics from M. I. T. in 1974. He worked from 1974–2004 at Bell Laboratories, A.T. & T. Bell Laboratories, and A. T. & T. Laboratories. He recently joined the faculty at the University of Michigan, where Trevor D. Wooley is a colleague. He is a frequent contributor to this MONTHLY.

Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, MI 48109-1043
email: lagarias@umich.edu