

MULTIPLICATIVE SEMIGROUPS RELATED TO THE $3x + 1$ PROBLEM

ANA CARAIANI

ABSTRACT. Recently Lagarias introduced the Wild semigroup, which is intimately connected to the $3x + 1$ Conjecture. Applegate and Lagarias proved a weakened form of the $3x + 1$ Conjecture while simultaneously characterizing the Wild semigroup through the Wild Number Theorem. In this paper, we consider a generalization of the Wild semigroup which leads to the statement of a weak $qx + 1$ conjecture for q any prime. We prove our conjecture for $q = 5$ together with a result analogous to the Wild Number Theorem. Next, we look at two other classes of variations of the Wild semigroup and prove a general statement of the same type as the Wild Number Theorem.

1. INTRODUCTION

The $3x + 1$ iteration is given by the function on the integers

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

The $3x + 1$ conjecture asserts that iteration of this function, starting from any positive integer n , eventually reaches the integer 1. This is a famous unsolved problem.

Farkas [2] formulated a semigroup problem which represents a weakening of the $3x + 1$ conjecture. He associated to this iteration the multiplicative semigroup W generated by all the rationals $\frac{n}{T(n)}$ for $n \geq 1$. We'll call this the $3x + 1$ *semigroup*, following the nomenclature in [1]. This generating set is easily seen to be

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

because the iteration can be written $T(2n+1) = 3n+2$ and $T(2n) = n$. Farkas observes that $1 = \frac{1}{2} \cdot 2 \in W$ and that if $T(n) \in W$ then $n = \frac{n}{T(n)} \cdot T(n) \in W$. Therefore the truth of the $3x + 1$ conjecture implies that all positive integers belong to W_3 . He raised the question that W contained all positive integers, a problem later termed by Lagarias [7] the Weak $3x + 1$ Conjecture.

In the course of studying Farkas's conjecture, Lagarias [7] was led to study a similar semigroup question concerning which integers occur in the inverse semigroup $W^{-1} = \{g^{-1} : g \in W\}$. We shall refer to this as the *inverse $3x + 1$ semigroup* (it is also known as the *wild semigroup* in [1]). The inverse $3x + 1$ semigroup has generators

$$G^{-1} = \left\{ \frac{3n+2}{2n+1} : n \geq 0 \right\} \cup \left\{ \frac{1}{2} \right\}.$$

He conjectured that the integers contained in W^{-1} are all the positive integers that are not divisible by 3 ("the Wild Numbers Conjecture") and proved that this new

conjecture is equivalent to the Weak $3x + 1$ Conjecture. Applegate and Lagarias [1] subsequently proved both of these conjectures. Their result gave a complete characterization of the elements of the semigroup W , showing that it consisted of all positive rationals whose denominator is not divisible by 3.

The results of Applegate and Lagarias establish that W is a very large semigroup inside the (infinitely generated) abelian group of all rational numbers \mathbb{Q}^* . Indeed, if we let W^+ be the semigroup generated by $W \cup \{\frac{1}{3}, -1\}$ then $W^+ = \mathbb{Q}^*$. This latter fact turns out to have a simplified direct proof, as we will show in a later section.

In this paper we study the structure of certain semigroups $S = S(A, B, C, D)$ associated to similar iteration problems. These semigroups all have generating sets of the form

$$S(A, B, C, D) = \left\langle \left\{ \frac{An + B}{Cn + D} : n \geq 0 \right\} \right\rangle,$$

together with a finite set of additional generators, specific to each iteration problem. We address the question of when some of these semigroups become equal to the group of rational numbers \mathbb{Q}^* . The specific cases we treat are the following ones.

First, we consider semigroups obtained from generalized “ $qx + 1$ conjectures.” If q is an odd prime, we consider the iteration of the $qx + 1$ function

$$T_q(x) = \begin{cases} \frac{x}{2} & \text{for } x \text{ even} \\ \frac{qx+1}{2} & \text{for } x \text{ odd.} \end{cases}$$

Related to the iteration above, we construct the multiplicative semigroup W_q generated by all rationals $\frac{T_q(n)}{n}$ for $n \geq 1$. Then it is easy to see that

$$W_q = \left\langle \left\{ \frac{qn + \frac{q+1}{2}}{2n + 1} : n \geq 0 \right\} \cup \left\{ \frac{1}{2} \right\} \right\rangle.$$

We conjecture that the semigroup W_q is “large” for every odd prime q , in the sense that after adding a finite number of generators to W_q , we can obtain the entire group \mathbb{Q}^* .

We can conceive of a “ $qx + 1$ conjecture” as stating that sufficient iteration of this function, starting from any positive integer n , eventually reaches 1. As in the case $q = 3$, the fact that W_q is large would follow from the “ $qx + 1$ conjecture.” The $qx + 1$ conjecture is false in general. For example, it fails for $q = 5$, since the iteration starting at 13 goes through the cycle 13, 33, 83, 208, 104, 52, 26, 13 and never reaches 1. The heuristic analysis of the $5x + 1$ problem suggests that the set of integers which iterate to 1 is very sparse [8]. In this paper, we nevertheless prove that $W_5[5, -1]$ is equal to \mathbb{Q}^* . Thus the semigroup problem associated to the $5x + 1$ problem has a “positive” answer. Thus, our findings indicate that the results of these semigroup problems shed no information on the truth or falsity of the $3x + 1$ problem, or the $5x + 1$ problem.

The second class of semigroups we consider are the semigroups

$$S_q = \left\langle \left\{ \frac{2qn - 1}{qn - 1} : n \geq 1 \right\} \right\rangle$$

where q is a prime, which were suggested by work of Farkas [2]. In the concluding section of the paper we discuss the semigroups

$$V_q = \left\langle \left\{ \frac{2qn + 1}{qn + 1} : n \geq 0 \right\} \right\rangle$$

which can be treated similarly to S_q . These are motivated by Farkas's treatment of V_2 [2, 6]. To obtain results on such a semigroup S_q we assume we have added to it as extra generators a suitably chosen finite set of rational numbers. Under the hypothesis that the prime q has 2 as a primitive root and that q is not the largest prime with this property, we prove that S_q will be equal to \mathbb{Q}^* after adding only a finite number of extra generators. The proof will essentially show that $n \in S_q$ by induction on the positive integer n and the reason for adding the extra generators is to ensure the base case of the induction.

2. FIRST PROPERTIES OF W_q AND S_q

In this section we derive some basic properties of both classes of semigroups we consider, proving parallel results for W_q and S_q . The main idea we highlight is that such a multiplicative semigroup contains many positive integers if and only if it contains many inverses of positive integers.

Recall that for an odd prime q we considered the semigroup W_q given by

$$W_q = \left\langle \left\{ \frac{qn + \frac{q+1}{2}}{2n + 1} : n \geq 0 \right\} \cup \left\{ \frac{1}{2} \right\} \right\rangle.$$

The case $q = 3$ is the ‘‘Wild semigroup’’ considered by Lagarias [7] and Applegate and Lagarias [1]. Every rational number in the generating set of W_q is a ratio between the value of the $qx + 1$ function at some positive integer n and the integer n . If the $qx + 1$ iteration were to eventually reach 1 then we could express $\frac{1}{n}$ as a product of generators of W_q . This motivates the following conjecture.

Conjecture 2.1 (Weak $qx + 1$ Conjecture). *For each positive integer n the fraction $\frac{1}{n}$ is in W_q .*

We also make the following parallel conjecture.

Conjecture 2.2 (Inverse $qx + 1$ Conjecture). *A positive integer n is in W_q if and only if $\gcd(n, q) = 1$.*

The special case $q = 3$ is the ‘‘Wild Numbers Conjecture’’ which was proved by Applegate and Lagarias [1]. In this section of the paper, we will prove that the two conjectures are close to being equivalent; namely they become equivalent after we add finitely many extra generators. In the next section of the paper we prove both of these conjectures in the case $q = 5$.

Our study of the semigroup W_q will involve selectively adding to it a finite number of extra generators. As a notation for describing semigroups with extra generators, define for $B \geq 2$ the semigroup

$$W_q[B] = \left\langle W_q \cup \left\{ \frac{1}{q} \right\} \cup \left\{ p, \frac{1}{p} : p \leq B, p \neq q \right\} \right\rangle$$

where p runs over prime numbers. In particular, $W_q[B]$ contains all positive integers $n \leq B$ which are prime to q .

Recall that for any prime q we have also defined the multiplicative semigroup

$$S_q = \left\langle \left\{ \frac{2qn - 1}{qn - 1} : n \geq 1 \right\} \right\rangle.$$

We conjecture that this semigroup is “large”, so that it will equal \mathbb{Q}^* after adjoining to it only finitely many generators. In order to describe the extra generators, we define for $B \geq 2$ the semigroup

$$S_q[B] = \left\langle S_q \cup \left\{ p, \frac{1}{p} : p \leq B, p \neq q \right\} \right\rangle.$$

Sometimes the extra generators being added this way are already in the semigroup (as we shall see in the following section in the case of W_5). We would hope to not have to add any generators at all, thus proving the strongest possible result on the semigroups W_q or S_q . This can be avoided for specific values of q via a computer calculation. However, to obtain results for general q it seems necessary to add a finite number of extra generators.

We now introduce several very useful ideas which will highlight our basic strategy in studying both W_q and S_q .

Definition 2.3. Let p be a prime number. We call a positive integer p -smooth if it is a product of primes $r < p$.

This property is useful because if a multiplicative semigroup S contains all primes $r < p$ then it contains all p -smooth numbers as well.

Definition 2.4. Let S be any of the multiplicative semigroups we are considering. We say that a positive integer n is *reduced* to m in S if $\frac{n}{m} \in S$. That is, if one can prove $m \in S$, it follows that $n \in S$.

We say that a positive integer n is *I-reduced* to m if $\frac{m}{n} \in S$. Thus, if $\frac{1}{m} \in S$ then it follows that $\frac{1}{n} \in S$.

Our strategy will be to reduce or I-reduce n to some $m < n$ and use induction to prove that S contains all positive integers relatively prime to some modulus or the reciprocals of these positive integers. Notice that because S is multiplicative the relation “reduced to” between n and m is transitive. (Note however that this relation is not symmetric.)

Note 2.5. Definition 2.4 allows an integer n to be reduced to an integer $m > n$. The notion of “reduced” does not inherently refer to reducing the size of the integer n .

Definition 2.6. Let k be a positive integer. We call k a *multiplier* for a multiplicative semigroup S if $\frac{1}{k} \in S$. We call k an *inverse multiplier*, or an *I-multiplier* for S if $k \in S$.

If $\frac{1}{k} \in S$, then once we prove that $kn \in S$ for some n it will also follow that $n \in S$. If $k \in S$, then proving $\frac{1}{kn} \in S$ implies that $\frac{1}{n} \in S$.

Our strategy for proving $n \in S$ will be to find an appropriate multiplier k for n and then reduce kn to some $m < n$. Then $m \in S$ will imply $kn \in S$ which in turn will imply $n \in S$. This basic strategy gives us the following theorem.

Theorem 2.7. *Let $q \geq 5$ be a prime and set $b(q) = 50q^3$. Then for all $B \geq b(q)$ the following conditional result holds: if $\frac{1}{m} \in W_q[B]$ for all integers $1 \leq m \leq M$ then*

$$n \in W_q[B] \text{ for all positive integers } n \leq \frac{M}{2q^3} \text{ with } \gcd(n, q) = 1.$$

In particular, if $\frac{1}{m} \in W_q[B]$ for all $m \geq 1$ then $n \in W_q[B]$ for all positive integers n with $\gcd(n, q) = 1$.

Proof. To prove that $n \in W_q[B]$ for all integers $n \leq \frac{M}{q^2}$ which are prime to q , it suffices to prove it for primes $p \neq q$ which are less than $\frac{M}{q^2}$. The proof will be by induction on the prime p . The fact that all such primes $p \leq B$ are already in $W_q[B]$ gives us the base case of the induction.

Let $p > B$ be the next prime for which we want to prove $p \in W_q[B]$. All primes $r < p$, $r \neq q$ satisfy $r \in W_q[B]$ by the induction hypothesis. Therefore, all p -smooth numbers which are prime to q will be contained in $W_q[B]$ as well. We are done if we can find a multiplier k so that kp reduces to a p -smooth number.

For k to be a multiplier, we want $\frac{1}{k} \in W_q[B]$ and by hypothesis this holds for any $k < M$. We wish to find a generator

$$\frac{kp}{2n+1} = \frac{qn + \frac{q+1}{2}}{2n+1} \in W_q$$

so that $2n+1$ is p -smooth. The condition on k is $2kp - 1 = q(2n+1)$, so $2kp \equiv 1 \pmod{q}$. Since $\gcd(2p, q) = 1$, this congruence can be satisfied for certain values of k and the condition puts k in an arithmetic progression of common difference q .

If k is any term in this progression, then kp reduces to $2n+1 = \frac{2kp-1}{q}$. If this number is p -smooth and not divisible by q then we are done. There are q possible residue classes for k modulo q^2 and $q-1$ of them will produce an integer $2n+1$ which is not divisible by q . We pick any class for k modulo q^2 out of the $q-1$. Then k can be any term in an arithmetic progression of common difference q^2 . As it runs through the terms of that arithmetic progression, $2n+1$ will run through the terms of an arithmetic progression of common difference $2pq$.

We claim that for $p > 50q^3$, we can find a p -smooth number $2n+1$ in this arithmetic progression which is less than $(2pq)^2$. We defer the proof of this claim to the next lemma. Now we complete the proof assuming the claim is true. Then $k = \frac{q(2n+1)+1}{2p} < 2pq^3$. Since p is bounded by $\frac{M}{2q^3}$ we get $k < M$ so k satisfies a sufficient condition for being a multiplier. Thus we are done. \square

Lemma 2.8. *Let p and q be prime numbers with $q \geq 5$. Assume that $p > 50q^3$. Then any arithmetic progression of common difference $2pq$ whose terms are relatively prime to $2pq$ contains a p -smooth number less than $(2pq)^2$.*

Proof. It suffices to show that more than half of the invertible residue classes modulo $2pq$ contain a p -smooth number between 1 and $2pq$. Indeed, once more than half the residue classes are p -smooth, we can obtain any invertible residue class as a product of two p -smooth residue classes. In all cases, the product will be a product of p -smooth numbers and so a p -smooth number itself. It is easy to see that the product will always be less than $(2pq)^2$.

We shall now use a counting argument. There are a total of $(p-1)(q-1)$ invertible residue classes modulo $2pq$. Out of these, the only ones whose members between 1 and $2pq$ are not p -smooth are those of the form $a \cdot r$, where $r \geq p$ is

a prime and a is an odd integer between 1 and $2q - 1$ inclusive. If we let $\pi(x)$ denote the prime counting function, then for any fixed a the number of such r is $\pi\left(\frac{2pq}{a}\right) - \pi(p - \epsilon)$ for positive $\epsilon \ll 1$, so if we sum over all a then it suffices to show that

$$\sum_{a=1, a \text{ odd}}^{2q-1} \left(\pi\left(\frac{2pq}{a}\right) - \pi(p - \epsilon) \right) < \frac{(p-1)(q-1)}{2},$$

or, rearranging slightly, that

$$\sum_{a=1, a \text{ odd}}^{2q-1} \pi\left(\frac{2pq}{a}\right) < \frac{(p-1)(q-1)}{2} + q\pi(p - \epsilon).$$

We now use the bounds cited in [1] that $\frac{x}{\log(x)} < \pi(x) < \frac{x}{\log(x)-1.5}$ for $x \geq 17$, and assume that $p > 17$. The left hand side of this equality is then bounded above by $\sum \frac{2pq/a}{\log(2pq/a)-1.5} < \frac{2pq}{\log(p)-1.5} \sum \frac{1}{a}$, and using the easy bound $\frac{1}{1} + \frac{1}{3} + \dots + \frac{1}{2q-1} < \frac{\log(2q-1)}{2} + \frac{1}{2} < \frac{\log(2q)+1}{2}$ we have

$$\sum_{a=1, a \text{ odd}}^{2q-1} \pi\left(\frac{2pq}{a}\right) < pq \cdot \frac{\log(2q) + 1}{\log(p) - 1.5}.$$

Similarly, the right hand side is bounded below by

$$\begin{aligned} \frac{(p-1)(q-1)}{2} + q\pi(p - \epsilon) &> \frac{(p-1)(q-1)}{2} + \frac{(p-\epsilon)q}{\log(p-\epsilon)} \\ &= pq \left(\frac{(1-\frac{1}{p})(1-\frac{1}{q})}{2} + \frac{1-\frac{\epsilon}{p}}{\log(p-\epsilon)} \right), \end{aligned}$$

and if we assume $q \geq 5$ and $p > 17$ then it is further bounded below by $pq \left(\frac{32}{85} + \frac{1}{\log(p)} \right)$ as $\epsilon \rightarrow 0$. Dividing by pq , we see that it suffices to pick $p > 17$ large enough so that

$$\begin{aligned} \frac{\log(2q) + 1}{\log(p) - 1.5} &< \frac{32}{85} + \frac{1}{\log(p)} \\ \log(2q) &< \frac{32}{85} \log(p) - \frac{48}{85} - \frac{1.5}{\log(p)}. \end{aligned}$$

Letting $P = \log(p)$ and $Q = \log(2q)$ for simplicity, we need to pick P sufficiently large so that $\frac{32}{85}P^2 - (Q + \frac{48}{85})P - 1.5 > 0$. This happens whenever P is bigger than the larger root of this quadratic:

$$P > \frac{Q + \frac{48}{85} + \sqrt{(Q + \frac{48}{85})^2 + \frac{192}{85}}}{64/85}$$

and since $Q \geq \log(10)$, we have $Q + \frac{80}{85} \geq \sqrt{(Q + \frac{48}{85})^2 + \frac{192}{85}}$, so it suffices to take $P > \frac{85}{32}Q + 2$. Exponentiating this, we need $p > e^2(2q)^{85/32}$. Therefore the lemma is true for all $p > 50q^3$, since

$$50q^3 > e^2(2q)^{85/32} = 46.58 \cdot q^{85/32}.$$

□

TABLE 1. Proof that W_5 contains reciprocals of all primes $p \leq 13$ and contains all primes $p \leq 23$.

$\frac{1}{3} = \left(\frac{1}{2}\right)^3 \cdot \frac{5 \cdot 1 + 3}{2 \cdot 1 + 1}$	$3 = \frac{5 \cdot 0 + 3}{2 \cdot 0 + 1}$
$\frac{1}{7} = \frac{1}{2} \cdot \left(\frac{1}{3}\right)^2 \cdot \frac{5 \cdot 3 + 3}{2 \cdot 3 + 1}$	$19 = 3 \cdot \frac{5 \cdot 26 + 3}{2 \cdot 26 + 1} \cdot \frac{5 \cdot 10 + 3}{2 \cdot 10 + 1}$
$\frac{1}{11} = \left(\frac{1}{2}\right)^2 \cdot \frac{1}{7} \cdot \frac{5 \cdot 5 + 3}{2 \cdot 5 + 1}$	$11 = 3 \cdot 19 \cdot \frac{1}{13} \cdot \frac{5 \cdot 28 + 3}{2 \cdot 28 + 1}$
$\frac{1}{13} = \frac{1}{3} \cdot \frac{1}{11} \cdot \frac{5 \cdot 6 + 3}{2 \cdot 6 + 1}$	$7 = 11 \cdot \left(\frac{1}{2}\right)^2 \cdot \frac{5 \cdot 5 + 3}{2 \cdot 5 + 1}$
$\frac{1}{5} = \frac{1}{13} \cdot \frac{5 \cdot 2 + 3}{2 \cdot 2 + 1}$	$2 = 7 \cdot \left(\frac{1}{3}\right)^2 \cdot \frac{5 \cdot 3 + 3}{2 \cdot 3 + 1}$
	$13 = 3 \cdot 19 \cdot \frac{1}{11} \cdot \frac{5 \cdot 28 + 3}{2 \cdot 28 + 1}$
	$17 = 3^3 \cdot \left(\frac{1}{2}\right)^2 \cdot \frac{5 \cdot 13 + 3}{2 \cdot 13 + 1}$
	$23 = 3^2 \cdot \frac{5 \cdot 4 + 3}{2 \cdot 4 + 1}$

Note 2.9. It can be shown in a similar fashion that for $B \geq 50q^3$ and assuming $n \in W_q[B]$ for all $n \geq 1$ relatively prime to q , then we may conclude $\frac{1}{n} \in W_q[B]$ for all positive integers n . This implies that the Weak $qx + 1$ Conjecture and the Inverse $qx + 1$ Conjecture are equivalent under the hypothesis that $p \in W_q$ and $\frac{1}{p} \in W_q$ for all primes $p \leq 50q^3$.

We have the analogous result for the semigroups S_q .

Theorem 2.10. *Let q be an odd prime and assume $B \geq 50q^3$. Then the following conditional result holds. If $m \in S_q[B]$ for $1 \leq m \leq M$ with $\gcd(m, q) = 1$ then*

$$\frac{1}{n} \in S_q[B] \text{ for all positive integers } n \leq \frac{M}{2q^2} \text{ with } \gcd(n, q) = 1.$$

Proof. We argue just as in the proof of Theorem 2.7. Thus we only need to show that $\frac{1}{p} \in S_q[B]$ for all prime numbers $p < \frac{M}{2q^2}$, $p \neq q$. We do this by induction on the prime p . From the definition of $S_q[B]$ we know this is true as long as $p \leq B$. Let p be the smallest prime for which we want to prove that $\frac{1}{p} \in S_q[B]$. Then $p > b(q)$ and by the induction hypothesis, all p -smooth numbers m which are relatively prime to q already satisfy $\frac{1}{m} \in S_q[B]$. Our strategy will be to find an I-multiplier $k \leq M$ such that kp I-reduces to a p -smooth number which is prime to q .

We want $kp = qn - 1$ for some n such that $2qn - 1$ is p -smooth, since $\frac{2qn-1}{qn-1} \in S_q[B]$. Then $kp \equiv -1 \pmod{q}$ so the possible values of k are in an arithmetic progression of difference q . The p -smooth number will have to be $2kp + 1$, taking values in an arithmetic progression of difference $2qp$. By Lemma 2.8, one can find a p -smooth number less than $(2pq)^2$ in this progression. Then we choose k such that $2kp + 1 < (2pq)^2$ is p -smooth, which means $k < \frac{(2pq)^2}{2p}$. As long as $p \leq \frac{M}{2q^2}$ we have that $k \leq M$ so the chosen k will be an I-multiplier. \square

3. THE $5x + 1$ SEMIGROUP

Now we will focus our attention on W_5 , the multiplicative semigroup generated by $\frac{5n+3}{2n+1}$ for $n \geq 0$ and $\frac{1}{2}$.

Theorem 3.1. *The multiplicative semigroup W_5 contains all $n \geq 1$ with $\gcd(n, 5) = 1$ and all fractions $\frac{1}{n}$ with $n \geq 1$.*

This theorem implies that both the Weak $5x + 1$ Conjecture and the Inverse $5x + 1$ Conjecture are true. These are special cases of Conjectures 2.1 and 2.2, respectively.

The proof of this theorem will be based on three lemmas, which will make up an inductive argument. Concretely, we will prove by induction on the positive integer n that $\frac{1}{n} \in W_5$. Since we have $1 \in W_5$ (it is equal to $2 \cdot \frac{1}{2}$) it suffices to I-reduce n to some positive integer $m < n$. We will provide a systematic way of doing this through the following lemmas.

Lemma 3.2. *For any positive integer $n < 12^{11} - 1$ we have $\frac{1}{n} \in W_5$.*

Proof. We will show that any $n < 12^{11} - 1$ I-reduces to some $m < n$. We may easily check that $\frac{1}{p} \in W_5$ for all primes $p \leq 13$ and that W_5 contains all primes $p \leq 23$ with $p \neq 5$; these computations are done in Table 1. We can use all odd $k \leq 23$ as I-multipliers. Thus for any odd $k \leq 23$ it suffices to I-reduce kn to some m , since this will imply that n also I-reduces to m . After picking a specific value of k we I-reduce kn to some $F(kn)$ such that $\frac{F(kn)}{kn} \in W_5$. However, we only allow one of three choices for $F(kn)$.

$$F(kn) = \begin{cases} \frac{kn}{2} & \text{if } kn \text{ is even} \\ \frac{kn}{3} & \text{if } kn \text{ is odd, but } 3 \mid kn \\ \frac{5kn+1}{2} & \text{otherwise.} \end{cases}$$

Notice that these choices make sense, since $\frac{F(kn)}{kn}$ will always equal $\frac{1}{2}$, $\frac{1}{3}$ or a fraction of the form $\frac{5N+3}{2N+1}$, all of which belong to W_5 . If n is even or a multiple of 3 we can pick $k = 1$ and immediately I-reduce n to $F(n)$ with $F(n) < n$. In this case we are done.

If $\gcd(n, 12) = 1$, the first step of the I-reduction does not give us $F(kn) < n$. Still, if $n \equiv 7 \pmod{12}$ we can again pick $k = 1$ and I-reduce n to $F(F(F(n))) = \frac{5n+1}{12} < n$. This is an example of our basic strategy. For any $n < 12^{11} - 1$ we can pick a successive sequence of I-multipliers k_1, \dots, k_j such that $m = F(k_j F(\dots F(k_1 n)))$ satisfies $m < n$. We can find the values of the I-multipliers k_1, \dots, k_j for each $n < 12^{11} - 1$ using a computer search. Thus, each $n < 12^{11} - 1$ I-reduces to some $m < n$ and the computer search records the largest value of the ratio $\frac{m}{n}$ as $\frac{694}{695} \approx 0.99856$. \square

Remark 3.3. The computer search finds a sequence of I-multipliers for each n with the property that going from n to m doesn't require dividing by more than 12^{11} . This means that any positive integer $n' \equiv n \pmod{12^{11}}$ gives us some $m' < n'$ from the exact same sequence of I-multipliers. Thus, instead of working with positive integers, we are working with residue classes modulo 12^{11} .

Lemma 3.4. *Let $M \geq 12^{11} - 2$. Assume that $\frac{1}{n} \in W_5$ for all $n \leq M$. Then $m \in W_5$ for all $m \leq \frac{M}{250}$ with $\gcd(m, 5) = 1$.*

Proof. We follow the argument from the proof of Theorem 2.7. The main point is to show that primes $p \leq \frac{M}{250}$ with $p \neq 5$ belong to W_5 . To do this, it suffices to prove that a p -smooth number less than $(10p)^2$ appears in every arithmetic progression with difference $10p$. By Lemma 2.8, this follows from the inequality

$$\pi(10p) + \pi\left(\frac{10p}{3}\right) + \pi\left(\frac{10p}{5}\right) + \pi\left(\frac{10p}{7}\right) + \pi\left(\frac{10p}{9}\right) < 2(p-1) + 5\pi(p-\epsilon).$$

This inequality holds for $167 \leq p \leq 700$ by computer search.

Using the approximation $\frac{x}{\log(x)} < \pi(x) < \frac{x}{\log(x)-1.5}$ for $x \geq 17$ given in [1], we see that the inequality above follows for $p > 17$ from the inequality

$$\sum \frac{10p/a}{\log(10p/a) - 1.5} < 2(p-1) + \frac{5(p-\epsilon)}{\log(p-\epsilon)},$$

where the sum is taken over odd $a < 10$. Letting $\epsilon \rightarrow 0$, this inequality is true for $p = 701$, and the difference between the right side and the left side is easily shown to be an increasing function for $p \geq 700$ by computing derivatives. We conclude that we can find the desired p -smooth number whenever $167 \leq p \leq \frac{M}{250}$. For the primes $p < 167$ other than 5 we show directly by means of a computer search that they belong to W_5 . This completes the proof of the lemma. \square

Lemma 3.5. *Let $j \geq 11$. Assume that every $n \not\equiv -1 \pmod{12^j}$ can be I-reduced to some $m < n$. Then every $n \not\equiv -1 \pmod{12^{j+1}}$ can be I-reduced to some $m < n$.*

Proof. From the hypothesis of the lemma, we can I-reduce all the integers in all the residue classes modulo 12^{j+1} except for $l \cdot 12^j - 1$ with $1 \leq l < 12$. It is enough to show how to handle these extra residue classes. Let $n = l \cdot 12^j - 1$ with $1 \leq l < 12$. We use $k = \frac{12^i+1}{5}$ as an I-multiplier, where i is the largest integer satisfying $j-10 \leq i \leq j-3$ and $i \equiv 2, 6, 14$ or $18 \pmod{20}$. First, notice that the smallest value of such an i is 6, obtained for $j = 11, \dots, 16$. Now, for i in those congruence classes, $k = \frac{12^i+1}{5}$ is always an integer and it is always prime to 5. Also, since $i \leq j-3$ we have that $k < \frac{12^j-2}{12^3} < \frac{12^j-2}{250}$, so by Lemma 3.4, $k \in W_5$. Therefore we can use it as an I-multiplier.

For the chosen k , $2F(kn) = (12^i + 1)(l \cdot 12^j - 1) + 1$ is divisible by 12^i . This means that kn can be I-reduced to $m_0 = l \cdot 12^j + l \cdot 12^{j-i} - 1$. Since $j-i \leq 10$ and $1 \leq l < 12$, we have $m_0 \not\equiv -1 \pmod{12^{11}}$. This means that we can further I-reduce m_0 to some $m < m_0$ as in Lemma 3.2. Moreover,

$$m_0 < \frac{12^i + 2}{12^i} n \leq \frac{12^6 + 2}{12^6} n.$$

Even though m_0 might not be in fact smaller than n , it is close enough to it that once we I-reduce m_0 to m we get

$$m < \frac{694}{695} m_0 < \frac{694}{695} \cdot \frac{12^6 + 2}{12^6} \cdot n < n.$$

Therefore, we can I-reduce all n of the form $l \cdot 12^j - 1$ with $1 \leq l < 12$ to some $m < n$. This proves that W_5 contains $\frac{1}{n}$ for all $n < 12^{j+1} - 1$. Any positive integer $n' \equiv n \pmod{12^{j+1}}$ can give us some $m' < n'$ from the exact same sequence of I-multipliers. This finishes the proof of our lemma. \square

Proof of Theorem 3.1. We now put together the three lemmas. We prove by induction on $j \geq 11$ that $\frac{1}{n} \in W_5$ for all $n < 12^j - 1$ and $m \in W_5$ for all $m < \frac{12^j-1}{250}$ satisfying $\gcd(m, 5) = 1$. Lemmas 3.2 and 3.4 give us the base case of the induction for $j = 11$. Assume now that the induction hypothesis is satisfied for some $j \geq 11$. Then lemma 3.5 implies that $\frac{1}{n} \in W_5$ for all $n < 12^{j+1} - 1$. This makes the hypothesis of lemma 3.4 for $j+1$ true and as a result $m \in W_5$ for all $m < \frac{12^{j+1}-1}{250}$ which satisfy $\gcd(m, 5) = 1$. As j becomes larger and larger we see that W_5 contains all fractions $\frac{1}{n}$ and all positive integers m which are prime to 5. \square

4. THE SEMIGROUPS S_q

We return to the study of the semigroups S_q which we have defined for every prime number q

$$S_q = \left\langle \left\{ \frac{2qn - 1}{qn - 1} : n \geq 1 \right\} \right\rangle.$$

Remember also that for $B \geq 2$, $S_q[B]$ is the semigroup obtained from S_q by adding a finite number of extra generators (specifically prime numbers $p \neq q$ which are less than B and their inverses). We make the following conjecture.

Conjecture 4.1. *Let q be a prime number. There exists some $B \geq 2$ such that $S_q[B]$ contains all positive integers n and their inverses $\frac{1}{n}$ which satisfy $\gcd(n, q) = 1$.*

We will prove this conjecture in a few special cases.

Theorem 4.2. *The semigroup S_2 generated by rationals of the form $\frac{4N+3}{2N+1}$ for $N \geq 0$ contains all odd integers $n > 1$.*

Proof. The proof is by induction on the odd integer n . First, we notice that for $N = 0$ we get $3 \in T_2$. Let $G(n) = \frac{n-1}{2}$ for n odd. Now let n be an odd integer greater than 1 so that $n' \in S_2$ for all $1 < n' < n$. Let 2^j be the largest power of 2 that divides $n - 1$, so that $n \equiv 2^j + 1 \pmod{2^{j+1}}$, and let $k = 2^j - 1$. Then $kn \equiv -1 \pmod{2^{j+1}}$, so kn will reduce to $G(kn) \equiv -1 \pmod{2^j}$ which can be further reduced. We eventually get kn to $m_0 = G(G \dots G(kn))$, where G is composed with itself j times. We can compute m_0 explicitly as $\frac{k(n-1)}{2^j}$. Let $m = \frac{n-1}{2^j} < n$. From the choice of j , we know that m is odd. We have $\frac{kn}{m_0} \in S_2$, but this fraction is the same as $\frac{n}{m}$. If $m = 1$ we find directly that $n \in S_2$. Otherwise, $m \in S_2$ by the induction hypothesis and $\frac{n}{m} \in S_2$, so $n \in S_2$ as well. \square

Corollary 4.3. *Conjecture 4.1 is true for $q = 2$.*

Proof. This follows from Theorems 2.10 and 4.2. \square

We will now study the semigroup S_3 generated by all rationals of the form $\frac{6n+5}{3n+2}$ with $n \geq 0$ a positive integer. Let us add the extra generators 2 and $\frac{1}{5}$. We notice that $2, \frac{5}{2}, \frac{11}{5}, \frac{23}{11}$ can be obtained from the generators, so all the primes 2, 5, 11, 23 can also be obtained. Similarly $\frac{1}{2}, \frac{1}{3}$ and $\frac{1}{5}$ can be obtained, so the semigroup with the extra generators is actually $S_3[5]$. We prove the following theorem.

Theorem 4.4. *The semigroup $S_3[5]$ contains all integers not divisible by 3 and their inverses.*

Proof. We shall prove that $n \in S_3[5]$ by induction on n . We claim $S_3[5]$ contains both n and $\frac{1}{n}$ for all $n \leq 2041$ with $\gcd(n, 3) = 1$. This can be verified through a computer search. For the induction step, it suffices to show we can reduce each integer n to some $m < n$ with $\frac{n}{m} \in S_3[5]$. As we've seen in the proof of Theorem 3.1, we can reduce an entire residue class simultaneously once we've managed to reduce its smallest representative.

Our argument will be based on two crucial lemmas, the proofs of which we give below. The first lemma says that, assuming $\frac{1}{m} \in S_3[5]$ for all $m \leq 2^{j+1}$ prime to 3, we can reduce all residue classes modulo $3 \cdot 2^{j+1}$ except for the residue class of 1.

The second lemma says that, assuming $\frac{1}{m} \in S_3[5]$ for all $m \leq 7^{l+2}$ prime to 3, we can reduce all $n \equiv 1 \pmod{3 \cdot 2^{j+1}}$ as long as $n \not\equiv 1 \pmod{3 \cdot 7^{l+1}}$. Putting the two lemmas together, we find that as long as $\frac{1}{m} \in S_3[5]$ for all $m \leq \max(2^{j+1}, 7^{l+1})$ prime to 3 we can reduce all positive integers $n \not\equiv 1 \pmod{3 \cdot 2^{j+1} \cdot 7^{l+1}}$ which are prime to 3.

Assuming the truth of the two lemmas, we prove the result up to the bound $M = 7^{l+2}$ by induction on l . If $S_3[5]$ contains all inverses of prime to 3 integers up to the bound M , then it must also contain all prime to 3 integers $n \leq \frac{M^2}{14} \leq 2^{j+1} \cdot 7^{l+1}$ where j satisfies $2^j \leq M < 2^{j+1}$. Since $M \geq 2041 > 50 \cdot 3^3$ we can apply Theorem 2.10. This means that $S_3[5]$ will contain the inverses of all the integers up to $\frac{M^2}{14 \cdot 2q^2} = \frac{M^2}{252} > 7M$ which are prime to 3. Now we can go through the same argument again, this time using $7M = 7^{l+3}$ as our bound instead of M . The theorem follows by induction on the integer l . \square

Lemma 4.5. *Let $j \geq 3$. Assume that $\frac{1}{m} \in S_3[5]$ for all $m \leq 2^{j+1}$ with $\gcd(m, 3) = 1$. Then all the integers n in all the residue classes modulo $3 \cdot 2^{j+1}$ except for the class of 1 can be reduced to smaller integers using multipliers no greater than 2^{j+1} .*

Proof. We first note that any $n \equiv 2 \pmod{3}$ can be very easily reduced to $G(n) = \frac{n-1}{2} < n$ if it is odd or to $\frac{n}{2} < n$ if it is even. From now on we shall consider only the case $n \equiv 1 \pmod{6}$. Observe that for such an n any odd multiplier k which satisfies $k \equiv 2 \pmod{3}$ allows us to reduce kn to $G(kn) = \frac{kn-1}{2}$. (This is true because we have $kn = 6n' + 5$ which can be reduced to $3n' + 2$.)

It is not hard to see that we can reduce most residue classes modulo 8. Indeed, we can perform a reduction by taking the multiplier $k = 5$. If $n \equiv 5 \pmod{8}$ then we can reduce n to $\frac{5n-1}{8} < n$ and we are done. If $n \equiv 7 \pmod{8}$ then we reduce it to $m = \frac{G(G(5n))}{8} = \frac{5n-3}{8} < n$. Finally, $n \equiv 3 \pmod{8}$ reduces to $m = G(G(G(5n))) = \frac{5n-7}{8} < n$. Thus, the only residue class modulo 8 that we haven't reduced is 1.

We show now how to reduce all residue classes modulo 2^{j+1} except for the class of 1. Let $i \leq j$ be the greatest integer such that $n \equiv 1 \pmod{2^i}$, or equivalently $n \equiv 2^i + 1 \pmod{2^{i+1}}$. Two cases arise depending on the parity of i . First, if $2 \mid i$ then $2^i + 1 \equiv 2 \pmod{3}$, so we can choose $k = 2^i + 1$. It is easy to see that $k < 2^{j+1}$ which implies $\frac{1}{k} \in S_3[5]$, so k actually is a multiplier. Furthermore, $kn - 1 \equiv 2^{2i} + 2^{i+1} \equiv 0 \pmod{2^{i+1}}$. Therefore, we can reduce n to $m = \frac{kn-1}{2^{i+1}} < \frac{kn}{2^{i+1}} < n$ and we are done with the first case. Second, if $2 \nmid i$ then $2^i + 3 \equiv 2 \pmod{3}$ and also $2^i + 3 < 2^{j+1}$, so we can choose $k = 2^i + 3$ as our multiplier. Then $kn - 3 \equiv (2^i + 1)(2^i + 3) - 3 \equiv 0 \pmod{2^{i+1}}$. In this case we can reduce n to $m = \frac{G(G(kn))}{2^{i+1}} = \frac{kn-3}{2^{i+1}} < n$. Thus we have shown that we can reduce all residue classes modulo 2^{j+1} except for the residue class of 1. \square

Lemma 4.6. *Let $j \geq 3$. Assume that $\frac{1}{m} \in S_3[5]$ for all $m \leq 7^{l+2}$ with $\gcd(m, 3) = 1$. Then any $n \equiv 1 \pmod{3 \cdot 2^{j+1}}$ can be reduced to a smaller integer as long as $n \not\equiv 1 \pmod{3 \cdot 7^{l+1}}$.*

Proof. We already know that $\frac{1}{7} \in S_3[5]$, so if at some point in reducing n we reach a multiple of 7, we can simply divide it by 7 and get a smaller integer. Furthermore, for reducing integers we shall keep using the iterating function $G(n) = \frac{n-1}{2}$, noticing

in addition that if $n \equiv 5 \pmod{6}$ and $7^i \mid G(n)$ then $\frac{G(n)}{7^i} \equiv G(n) \pmod{6}$ is either even or congruent to 5 (mod 6).

Now assume that $n \equiv 1 \pmod{2^{j+1}}$. We shall try to reduce most residue classes that n could belong to modulo 7^{l+1} . First, we show that if $n \not\equiv 1 \pmod{7}$ then it is quite easy to reduce n . Indeed, for $n \equiv 2 \pmod{7}$ we use 11 as a multiplier, going from n to $11n$ and further to $\frac{11n-1}{2}$ which is a multiple of 7, so we reduce n to $m = \frac{11n-1}{14} < n$. For $n \equiv 3 \pmod{7}$ we use 5 as a multiplier and get to $\frac{5n-1}{2}$, which is still divisible by 2 (as $n \equiv 1 \pmod{8}$) so we get the odd integer $\frac{5n-1}{4}$. This integer is divisible by 7 so we can further reduce it to $m = \frac{5n-1}{28} < n$. For $n \equiv 4 \pmod{7}$ we use 23 as a multiplier and get from n to $\frac{23n-1}{14} \equiv 5 \pmod{6}$, so we can apply G one more time and get to $\frac{23n-15}{28} < n$. For $n \equiv 5 \pmod{7}$ we use 17 as a multiplier and notice that $17n - 1 \equiv 0 \pmod{28}$ so we reduce to $m = \frac{17n-1}{28} < n$. For $n \equiv 6 \pmod{7}$ we first reduce to $m_0 = \frac{5n-1}{4} = \frac{1}{2}G(5n) \equiv 2 \pmod{7}$ in order to switch to a better residue class modulo 7. Finally we reduce n to $m = \frac{11m_0-1}{14} < \frac{11}{14} \frac{5n-1}{4} < \frac{55}{56}n < n$. In this way, we have seen that if $n \not\equiv 1 \pmod{7}$ then n can be reduced. This will be the base case for an argument by induction.

Assume that any $n \equiv 1 \pmod{2^{j+1}}$, $n \not\equiv 1 \pmod{7^i}$ for some $i \leq l$ can be reduced to some $m < n$. Let $n \equiv 1 \pmod{7^i}$ such that $7^{i+1} \nmid n - 1$. Then n is congruent modulo 7^{i+1} to some $t \cdot 7^i + 1$ with $1 \leq t \leq 6$, so we can prove that n is reducible by providing the reduction steps for each t . In each case, we want to use an integer k congruent to both $n^{-1} \equiv (7-t) \cdot 7^i + 1 \pmod{7^{i+1}}$ and 5 (mod 6) as a multiplier. In fact it will suffice to let $k = k' \cdot 7^{i+1} + (7-t) \cdot 7^i + 1$ for some $k' < 7$. For $n \equiv 7^i + 1 \pmod{7^{i+1}}$, we let the multiplier be $k = 4 \cdot 7^{i+1} + 6 \cdot 7^i + 1$, and we have $kn - 1 \equiv 0 \pmod{7^{i+1}}$ and $kn - 1 \equiv 2 \pmod{4}$. We divide $G(kn)$ by 7^{i+1} and apply G again. If $G\left(\frac{G(kn)}{7^{i+1}}\right)$ is even, then we can divide by 2 and get a number $m < \frac{kn}{8 \cdot 7^{i+1}} < n$. If $G\left(\frac{G(kn)}{7^{i+1}}\right)$ is odd, then it is also congruent to 2 (mod 3) so we can apply G once more. Applying G sends an integer to less than half that integer, so again we get $m < \frac{kn}{8 \cdot 7^{i+1}} < n$. So in the case $n \equiv 7^i + 1 \pmod{7^{i+1}}$ we are done.

If $n \equiv 2 \cdot 7^i + 1 \pmod{7^{i+1}}$ then the inverse of n modulo 7^{i+1} is $5 \cdot 7^i + 1 \equiv 0 \pmod{6}$. We use $k = 5 \cdot 7^{i+1} + 5 \cdot 7^i + 1$ as a multiplier in order get to $G(kn)$ divisible by 7^{i+1} . We notice that $kn - 1 \equiv k - 1 \equiv 8 \cdot 5 \cdot 7^i \pmod{2^{j+1}}$ so $8 \mid kn - 1$. This means that $4 \mid G(kn)$ in addition to $7^{i+1} \mid F(kn)$ by construction. Then we can take $m = \frac{G(kn)}{4 \cdot 7^{i+1}} < \frac{k}{8 \cdot 7^{i+1}}n < n$.

In the remaining cases, we have $n \equiv t \cdot 7^i + 1 \pmod{7^{i+1}}$ with $3 \leq t \leq 6$. Taking $k = (6t - 14) \cdot 7^i + 1$ as a multiplier, it is easy to verify that $kn \equiv 1 \pmod{7^{i+1}}$ and $k \equiv 5 \pmod{6}$. Then $m_0 = \frac{G(kn)}{7^{i+1}} < \frac{k}{14 \cdot 7^i}n$ is either even or congruent to 5 (mod 6). If it is even we divide by 2 and if it is odd we apply G again, so in either case we have reduced m_0 to an integer $m \leq \frac{m_0}{2} < \frac{k}{28 \cdot 7^i}n < \frac{(6t-13) \cdot 7^i}{28 \cdot 7^i}n < n$. Thus we have shown that we can reduce all residues but 1 (mod 7^{i+1}), so the induction step is complete. \square

5. SEMIGROUPS S_q , FOR q HAVING 2 AS A PRIMITIVE ROOT

In this section we prove conjecture 4.1 for a substantial class of primes, namely those when q satisfies two hypotheses:

H1: 2 is a primitive root modulo q ;

H2: There is a prime $p > q$ such that p has 2 as a primitive root.

The second hypothesis is not as important, since it is believed there are infinitely many primes for which 2 is a primitive root. In fact, Hooley [4] has shown that assuming suitable Riemann hypotheses are true, then Artin's conjecture on primitive roots holds, and in particular 2 is a primitive root for infinitely many primes q . The results of Hooley imply that 2 should be a primitive root for a positive proportion of primes, with proportion given by Artin's constant $C = 0.3739\dots$. Moreover, Heath-Brown [3] showed unconditionally that there are at most two exceptional primes for which Artin's conjecture fails. If 2 is not one of these exceptions, then we will have proved our conjecture for infinitely many primes. Putting together the hypotheses, we get the following result.

Theorem 5.1. *Let q be a prime which satisfies (H1) and (H2). Then there exists some finite $B \geq 2$ depending on q such that $S_q[B]$ contains all positive integers which are prime to q and their inverses.*

The question of finding B arises. Our general argument allows us to find an upper bound for it, which is of the form $pq^2 \cdot 2^{p+3q-2}$ where p is the smallest prime greater than q with 2 as a primitive root. In practice, for small q it is easy to find much better bounds for B . For $q = 3$ we've seen that it is enough to take $B = 5$. One can use a computer search to find B for $q = 5, 11, 13, 19$.

We shall now give the general argument for $q > 3$ as long as it satisfies the hypotheses of the theorem. We make use of a few technical lemmas.

Lemma 5.2. *Let $n \equiv -1 \pmod{q}$ be an odd integer. Then n can be reduced to $G(n) = \frac{n-1}{2}$.*

Proof. It's easy to see that $n = 2qN - 1$ for some $N \geq 1$, so it can be reduced to $G(n) = qN - 1$. \square

Notation 5.3. Let n and j be positive integers, with n odd. Let n_j denote the smallest positive integer for which $n \cdot n_j \equiv 1 \pmod{2^j}$.

In particular, $n_j < 2^j$. If we think of it as a residue class, n_j is the inverse in $\mathbb{Z}/2^j\mathbb{Z}$ of the residue class of n .

Lemma 5.4. *Assume that $j \geq q - 1$ and that $\frac{1}{k} \in S_q[B]$ for all $k \leq 2^j$ which are prime to q . Then any integer n in a residue class modulo $q \cdot 2^j$ for which both n_j and $n_j n - 1$ are prime to q can be reduced to a smaller integer using multipliers no greater than 2^j .*

Proof. If $q \nmid n_j$ then n_j belongs to an invertible residue class modulo q . Since 2 is a primitive root modulo q , the integers of the form $2^i - 1$ cover all invertible residue classes modulo q except for -1 . As a result, one can pick i between 1 and $q-2$ such that $(2^i - 1) \cdot n_j \cdot n \equiv -1 \pmod{q}$ as long as $n_j \cdot n \not\equiv 1 \pmod{q}$. We will multiply n by $(2^i - 1)n_j$. The resulting integer $m_0 = (2^i - 1)n_j \cdot n$ is congruent to $2^i - 1$ modulo 2^j and to $-1 \pmod{q}$. Therefore, we can reduce m_0 to $G(m_0) = \frac{m_0 - 1}{2}$, which is still congruent to $-1 \pmod{q}$ and congruent to $2^{i-1} - 1$ modulo 2^{j-1} . We can apply G a total of i times and we get

$$G(\dots G(m_0)) = \frac{m_0 - (2^i - 1)}{2^i}.$$

This integer is still divisible by 2^{j-i} , so after a total of j steps we get $\frac{(2^i-1)n_j n - (2^i-1)}{2^j}$. In fact, what we've shown is that $\frac{(2^i-1)n_j n}{(2^i-1)m} \in S_q[B]$, where $m = \frac{n_j n - 1}{2^j}$. Thus $\frac{n_j n}{m} \in S_q[B]$ and we can reduce n directly to $m = \frac{n_j n - 1}{2^j} < n$ using $n_j < 2^j$ as a multiplier. \square

When we find a series of steps to reduce simultaneously all the integers in a given residue class to smaller integers, we refer to this series of steps as a way of reducing the entire residue class. Lemma 5.4 allows us to reduce many residue classes modulo $q \cdot 2^j$. Indeed, assume that some residue class n modulo $q \cdot 2^j$ has not been reduced modulo $q \cdot 2^i$ for any $q-1 \leq i \leq j$. Let a denote the inverse of the residue class of n modulo q . Then n_i has to be either 0 or a modulo q for all $q-1 \leq i \leq j$. On the other hand, each $n_i \equiv n_{q-1} \pmod{2^{q-1}}$ so n_j can be written as $n_{q-1} + \sum_{i=q}^j \epsilon_i 2^{i-1}$ where ϵ_i is either 0 or 1. Truncating the sum after the term $\epsilon_i 2^{i-1}$ gives us the value of n_i . We deduce that the only powers of 2 with coefficient 1 are those congruent to a or $-a$ modulo q . Moreover, a and $-a$ have to appear alternately in the sum, to ensure that the sum at any point is either 0 or a modulo q . Also, since 2 is a primitive root modulo q , the difference between the exponents of 2 which give a and $-a$ as residues is exactly $\frac{q-1}{2}$. Note that for $q > 3$ we have $\frac{q-1}{2} \geq 2$. We will show that the residue classes that do not reduce must have $n_j = n_{q-1}$.

Lemma 5.5. *Assume that $n_i = 2^{i-1} + n_{i-1}$ and that $q \mid n_{i-1}$. Then n can be reduced to a smaller integer using a multiplier less than 2^{i+1} and so can its entire residue class modulo $q \cdot 2^{i+1}$.*

Proof. Let $0 < a < q-1$ be the inverse of the residue class of n modulo q . Since $n_{i-1} \equiv 0 \pmod{q}$ we must have that $2^{i-1} \equiv a \pmod{q}$, otherwise we are done. Moreover, $n_{i+1} = n_i$, since the difference between consecutive exponents of nonzero terms in n_j has to be at least 2 for $q > 3$. Then $(2^l - 1)(2^{i-1} + n_{i-1}) \equiv n_{i-1}(2^l - 1) - 2^{i-1} \pmod{2^{i+1}}$ for $l \geq 2$. If we pick l such that $2^{i+1} > (2^l - 1)n_{i-1} > 2^{i-1}$, then $n_{i-1}(2^l - 1) - 2^{i-1}$ is the residue class of $n_{i+1}(2^l - 1)$ modulo 2^{i+1} . Moreover, if we let $k = n_{i-1}(2^l - 1) - 2^{i-1}$ then $k \equiv -a \pmod{q}$, so $kn \equiv -1 \pmod{q}$. Following exactly the same steps of reduction as in Lemma 5.4 we see that we can reduce n to $\frac{kn - (2^l - 1)}{2^{i+1}} < n$. Moreover, this reduction process only depends on the residue class of n modulo q and modulo 2^{i+1} .

We're only left to check that we can pick the appropriate l . This is easy to check by placing n_{i-1} between 2^h and 2^{h+1} for some $h \leq i-1$. Let $l = i - h$ and notice that $(2^l - 1)n_{i-1} > 2^{l-1+h} = 2^{i-1}$ (hence $l \geq 2$, or we would have $n_{i-1} > 2^{i-1}$) and $(2^l - 1)n_{i-1} < 2^{l+h+1} = 2^{i+1}$, as desired. \square

Lemma 5.5 reduces all the residue classes with more than one 2^i term added after n_{q-1} , using multipliers no greater than 2^{j+1} . Indeed, if at least two terms are added, then one of these terms will be added to a multiple of q . Assume $n_j \neq n_{q-1}$. The only possibility is that n_{q-1} is prime to q and there is just one term added after it, so $n_j = 2^i + n_{q-1}$. Then n_j will be a multiple of q and $2^i \equiv -a \pmod{q}$. Let $l = i + q - 1$. We shall reduce this case, using multipliers no greater than 2^{j+q-1} .

Proposition 5.6. *Let $C_1 = 2^{q-1}$. For the semigroup $S_q[B]$, with q satisfying hypotheses (H1) and (H2), the only residues modulo 2^j which do not reduce using*

potential multipliers less than $C_1 \cdot 2^j$ are those for which $n_j = n_{q-1}$. The number of such residues is bounded as $j \rightarrow \infty$, since n_{q-1} takes values less than 2^{q-1} .

Proof. Notice that $n_l = n_{i+1} = n_j$, since otherwise $n_l = n_{i+1} + 2^{i+\frac{q-1}{2}}$ and we use Lemma 5.5. In this case, we notice that $(2^l - 1)n_j$ is congruent to $k = 2^l - n_j$ modulo 2^l . Also, $k \equiv 2^l \equiv -a \pmod{q}$ so $kn \equiv -1 \pmod{q}$. Thus, we reduce n to $\frac{kn - (2^l - 1)}{2^l} < n$. The multiplier used is at most $2^l < 2^{j+q-1} = C_1 \cdot 2^j$. \square

Let $p > q$ be a prime number for which 2 is a primitive root and assume $p \in S_q[B]$ (for example if $B \geq p$). Let j be sufficiently large and let n be a positive integer for which $n_j = n_{q-1}$. Then we will show that we can reduce n modulo p^l using multipliers not larger than $C_2 p^l$ as long as n is not congruent to $(n_{q-1})^{-1}$ modulo p^l . It is clear that if $p \mid n$ then n can be reduced to $\frac{n}{p}$, so we shall assume that n is in an invertible residue class modulo p^l .

We give an explicit way to construct the appropriate multiplier for n if $n_{q-1} \cdot n - 1 \not\equiv 0 \pmod{p^l}$. First, notice that by the same argument as in Lemma 5.4, instead of requiring the multiplier to be exactly $-a$ modulo q it is enough to find a multiplier $k \equiv n_{q-1} \pmod{2^j}$ that is not congruent to 0 or a modulo q . If k is such a multiplier, then we can multiply it by some $2^h - 1$ to get $(2^h - 1)k \equiv -a \pmod{q}$ and obtain $\frac{(2^h - 1)kn - (2^h - 1)}{2^j}$ after j steps. The $2^h - 1$ cancels as before and we are left with $\frac{kn - 1}{2^j}$.

Unfortunately, we may have $k > 2^j$, but we can still try to select k such that $p^i \mid kn - 1$ for some large enough i . In that case, knowing that $p \in S_q[B]$ allows us to divide by it, so we would be able to reduce n to $\frac{kn - 1}{2^j \cdot p^i} < n$. Indeed, assume that $i - 1 < l$ is the largest exponent such that $p^{i-1} \mid n_{q-1} \cdot n - 1$.

We want to find a multiplier of the form $k = 2^j \cdot t \cdot p^{i-1} + n_{q-1}$ such that $p^i \mid kn - 1$ and $t < p$. This condition is equivalent to $2^j \cdot t \cdot n \equiv -\frac{n_{q-1} \cdot n - 1}{p^{i-1}} \pmod{p}$, or that $2^j \cdot t$ covers all invertible residues modulo p . At the same time, we have the restriction that $2^j \cdot t$ can not be congruent to $-n_{q-1}$ or $a - n_{q-1}$ modulo q . Now, n_{q-1} is either 0 or a modulo q , so in either case, one of the restricted residues is 0. If we require t to be invertible modulo q we are left with only one restricted residue. We use the following result.

Lemma 5.7. *Assume that 2 is a primitive root modulo $q \geq 5$ and that there exists a prime $p > q$ such that 2 is a primitive root modulo p . Then such a prime p can be chosen so that for any invertible residues r, s modulo p, q respectively we can choose $q - 1 \leq j < p + q - 2$ and $t < p$ with $q \nmid t$ such that $2^j \cdot t \equiv r \pmod{p}$ but $2^j \cdot t \not\equiv s \pmod{q}$.*

Proof. Since 2 is a primitive root modulo p we know that $2^j \pmod{p}$ covers all invertible residues modulo p when j varies between $q - 1$ and $p + q - 3$. There are $p - 1$ such residues and they are in bijective correspondence with the values of $t \equiv r \cdot 2^{-j} \pmod{p}$. Thus each value of j determines a value of $t = t(j)$. Assume that for all such pairs t and j we get $2^j \cdot t \equiv s \pmod{q}$. Then if $j \leq p - 2$ and $j' = j + q - 1$ we must have either $q \mid t(j)$ or $q \mid t(j')$ or $q \mid t(j) - t(j')$. There are less than $\left\lfloor \frac{p}{q} \right\rfloor$ positive integers less than p which are divisible by q , so there are at most $2 \left\lfloor \frac{p}{q} \right\rfloor$ values of j that fall under the first two cases. On the other hand, assume that $t(j') = t(j) + bq$, where b can be positive or negative and thus take

any of $2 \left\lfloor \frac{p}{q} \right\rfloor$ values. Then by our assumption $2^{j+q-1}(t(j) + bq) \equiv 2^j t(j) \pmod{p}$ or $2^{j+q-1} \cdot bq \equiv r(1 - 2^{q-1}) \pmod{p}$. Thus, each value of b determines at most one value of j . This accounts for another $2 \left\lfloor \frac{p}{q} \right\rfloor$ values of n . The total number of values j can take between $q - 1$ and $p - 2$ is $p - q$. If we can find p such that 2 is a primitive root modulo p and $p > q + 4 \left\lfloor \frac{p}{q} \right\rfloor$ then we are done.

For $q = 5$ we can simply choose $p = 19$. For $q \geq 11$ and $p > 2q$ the inequality $p > q + 4 \left\lfloor \frac{p}{q} \right\rfloor$ is satisfied. If there is no prime $p > 2q$ with 2 a primitive root modulo p , then we can choose such a p between q and $2q$. The inequality we want becomes $p > q + 4$. If $p = q + 4$, however, then 16 divides either $p^2 - 1$ or $q^2 - 1$, so quadratic reciprocity then implies that 2 is a quadratic residue either modulo p or modulo q . This contradicts the assumption that 2 is a primitive root modulo both p and q , so the only way that $p > q + 4$ might not be satisfied is when $p = q + 2$.

Suppose that $p = q + 2$. Then the only pair of invertible residues modulo p with difference q is $(1, q + 1)$. The values of j and j' corresponding to this pair must have difference a multiple of $q - 1$ and must be between $q - 1$ and $2q - 1$. The only such pairs are $(q - 1, 2q - 2)$ and $(q, 2q - 1)$. For both these pairs, t and t' are 1 and -1 . Also, the exponents $q - 1, q, 2q - 2, 2q - 1$ become $-2, -1, -4, -3$ if reduced modulo $p - 1$. This implies we have either $2^{-2} \equiv -2^{-4} \pmod{p}$ or $2^{-1} \equiv -2^{-3} \pmod{p}$, which both give $p = 5$. However, $p > q \geq 5$ so we have reached a contradiction, and thus it follows that $p > q + 4$ as desired. \square

Lemma 5.7 implies that as j varies between $q - 1$ and $p + q - 1$, $2^j \cdot t$ covers all invertible residues modulo p while always avoiding one specified residue modulo q . As a result, we can always pick t and j to get our desired multiplier k , and k will be less than $2^{p+q-2} p^l$. Let $C_2 = 2^{p+q-2}$. Then we can reduce almost all residues modulo p^l using multipliers no higher than $C_2 p^l$.

The only integers that can not be reduced are those congruent to $(n_{q-1})^{-1}$ modulo 2^j and modulo p^l . Therefore they are at least as large as $\frac{2^j p^l + 1}{2^{q-1}}$, and so we conclude that all of the integers up to $\frac{2^j p^l + 1}{2^{q-1}}$ that are not divisible by q belong to $S_q[B]$. We can now use induction to get the proof of Theorem 5.1 for $q > 3$.

Proof of Theorem 5.1. Suppose $\frac{1}{m} \in S_q[B]$ for all integers $m \leq M$ which are prime to q (For the base case of the induction, we pick $M = B$). We set $j = \lfloor \log_2 \frac{M}{C_1} \rfloor$ and $l = \lfloor \log_p \frac{M}{C_2} \rfloor$. Then we have shown above that all integers n which are prime to q and less than $\frac{M^2}{2^p C_1 C_2 2^{q-1}}$ can be reduced to smaller integers using multipliers no larger than M . Therefore, $S_q[B]$ contains all these integers $n \leq \frac{M^2}{2^p C_1 C_2 2^{q-1}}$. Set

$$C = 2^{p+3q-2} p q^2.$$

From Theorem 2.10, we find $\frac{1}{m} \in S_q[B]$ for all integers $m \leq \frac{M^2}{C}$ prime to q . It suffices to start out with $B > C$, to ensure that each bound $M \geq B$ can be replaced by $\frac{M^2}{C} > M + 1$. The theorem follows by induction. \square

6. CONCLUDING REMARKS

The basic argument used in proving that $S_q[B]$ contains almost all rationals for large enough B can be adapted to prove a similar result for the semigroups

$$V_q = \left\langle \left\{ \frac{2qn+1}{qn+1} : n \geq 0 \right\} \cup \{2\} \right\rangle,$$

where q is a prime number. In order to account for extra generators, we set

$$V_q[B] = \left\langle V_q \cup \left\{ \frac{1}{p}, p : p \leq B, p \neq q \right\} \right\rangle.$$

For large enough B and as long as q satisfies (H1) and (H2) we claim that $V_q[B]$ contains $n, \frac{1}{n}$ for all positive integers n such that $\gcd(n, q) = 1$. The proof of this fact is analogous to the proof of the corresponding statement for $S_q[B]$. The only essential difference is that we use the reduction of $m \equiv 1 \pmod{2q}$ to $n = \frac{m+1}{2}$. For a general $m \not\equiv 0 \pmod{q}$ we try to find multipliers k such that $km \equiv 1 \pmod{2q}$. The basic strategy remains that of reducing residues modulo higher and higher powers of 2 and of r . One question that arises from verifying these results is what happens when 2 is not a primitive root modulo q and whether we can adapt the above strategy to this case.

Another promising direction is adapting the basic strategy to various semigroups obtained by adjoining extra generators to W_q, S_q, V_q . In doing so, we have a found an easy proof of a weakened version of the Wild Numbers Theorem. Let

$$W_3^+ = \langle W_3 \cup \{3, -1\} \rangle.$$

By the Wild Numbers Theorem, $W_3^+ = \mathbb{Q}^*$. We explain an easier, direct proof of this fact, based on the following two lemmas.

Lemma 6.1. *Let $M \gg 0$ and suppose that $\frac{1}{m} \in W_3^+$ for all positive $m \leq M$. Then $n \in W_3^+$ for all positive $n \leq \frac{M^2}{147}$.*

Proof. The main step is reducing n . Let $G(n) = \frac{2n-1}{3}$ for all $n \equiv 2 \pmod{3}$. We clearly have $\frac{n}{G(n)} \in W_3^+$, so $n \equiv 2 \pmod{3}$ can be reduced easily. Suppose now that $n \equiv 1 \pmod{3^j}$ but $n \not\equiv 1 \pmod{3^{j+1}}$. We use $k = \frac{1}{2}(1 + 3^j)$ as a multiplier. If $n \equiv 1 + 2 \cdot 3^j \pmod{3^{j+1}}$ then $G(kn)$ will be divisible by 3^j so we reduce n to

$$m = \frac{G(kn)}{3^j} < \frac{3^j + 1}{3^{j+1}} \cdot n < n.$$

If instead $n \equiv 1 + 3^j \pmod{3^{j+1}}$ then we reduce n to $m' = \frac{G(kn)}{3^{j+1}}$. However, $m' \equiv 2 \pmod{3}$ so it can be reduced in turn to $m = G(m') < \frac{2}{3}m'$. We've managed to reduce n to

$$m < \frac{2}{3} \cdot \frac{3^j + 1}{3^j} \cdot n < n.$$

Thus, we can reduce all residue classes modulo 3^{j+1} except for the residue class of 1 using multipliers no greater than 3^{j+1} . Similarly, we can reduce all residue classes modulo 7^{i+1} other than the residue class of 1 using multipliers no greater than 7^{i+1} . Putting these two facts together, we see that $n \in W_3^+$ for all $n < \frac{M}{3} \cdot \frac{M}{49}$. \square

The next lemma is proved similarly, by I-reducing residue classes modulo higher and higher powers of 2 and of 5.

Lemma 6.2. *Let $M \gg 0$ and suppose that $m \in W_3^+$ for all positive $m \leq M$. Then $\frac{1}{n} \in W_3^+$ for all positive $n \leq \frac{M^2}{50}$.*

Putting the two lemmas together, along with a computer search to verify that W_3^+ contains $n, \frac{1}{n}$ for all $n \leq 10^4$ we get the following result.

Proposition 6.3. *The semigroup W_3^+ contains all positive rational numbers.*

This proposition is weaker than the Wild Numbers Theorem; however, it gives a partial answer to a question posed in [7]. The argument of Lagarias and Applegate uses a bootstrap induction which deduces the validity of the Wild Numbers Theorem over some interval from the validity of the Weak $3x + 1$ Conjecture over a larger interval. Lagarias asks whether there is an argument that could be applied symmetrically to both directions. It turns out that if we add 3 as a generator to W_3 then the function $G(n) = \frac{2n-1}{3}$ for $n \equiv 2 \pmod{3}$ generates a dynamical system similar to the one obtained from iterating the $3x + 1$ function. This suggests that adding q as a generator to W_q, S_q (or V_q) makes it considerably easier to prove the desired results.

Acknowledgement. This research was done at the University of Minnesota Duluth with financial support from University of Minnesota Duluth and Princeton University. The author would like to thank Joe Gallian, David Arthur, Reid Barton, Geir Helleloid, Ricky Liu, Phil Matchett Wood, and Steven Sivek for their help in completing the work. The author would also like to thank Jeff Lagarias and the anonymous referees for providing helpful references and detailed advice about the organization of this paper.

REFERENCES

- [1] D. Applegate and J. C. Lagarias, The $3x + 1$ semigroup, *J. Number Theory* **117** (2006), no. 1, 146–159.
- [2] H. Farkas, Variants of the $3x + 1$ problem and multiplicative semigroups, pp.121–127 in: *Geometry, Spectral Theory, Groups and Dynamics*, Contemporary Math Vol. 387, Amer. Math. Soc., Providence, RI, 2005.
- [3] D.R. Heath-Brown, Artin’s conjecture for primitive roots, *Quart. J. Math. Oxford Ser. (2)* **37** (1986), no. 145, 27–38.
- [4] C. Hooley, On Artin’s conjecture, *J. Reine Angew. Math.* **225** (1967), 209–220.
- [5] “IMO 1998/A3 solution.” <http://www.kalva.demon.co.uk/imo/isoln/isoln983.html>.
- [6] A. Kontorovich and J. C. Lagarias, Stochastic models for the $3x + 1$ and $5x + 1$ problems, preprint: arxiv:0910.1944.
- [7] J. C. Lagarias, Wild and Wooley numbers, *Amer. Math. Monthly* **113** (2006), no. 2, 97–108.
- [8] S. Volkov, A probabilistic model for the $5k + 1$ problem and related maps, *Stochastic Process. Appl.* **116** (2006), no. 4, 662–674.

Current address: Harvard Mathematics Dept.

E-mail address: caraiani@math.harvard.edu