Strongly sufficient sets and the distribution of arithmetic sequences in the $3x+1$ graph

Keenan Monks *
Harvard University
Cambridge, MA 02138

Kenneth G. Monks †
University of Scranton
Scranton, PA 18510

Kenneth M. Monks ‡
Colorado State University
Fort Collins, CO 80523

Maria Monks §
University of California
Berkeley, CA 94720

April 20, 2012

Abstract

The $3x+1$ Conjecture asserts that the T-orbit of every positive integer contains 1, where T maps $x \mapsto x/2$ for x even and $x \mapsto (3x+1)/2$ for x odd. A set S of positive integers is *sufficient* if the orbit of each positive integer intersects the orbit of some member of S. In [9] it was shown that every arithmetic sequence is sufficient.

In this paper we further investigate the concept of sufficiency. We construct sufficient sets of arbitrarily low asymptotic density in the natural numbers. We determine the structure of the groups generated by the maps $x \mapsto x/2$ and $x \mapsto (3x+1)/2$ modulo b for b relatively prime to 6, and study the action of these groups on the directed graph associated to the $3x+1$ dynamical system. From this we obtain information about the distribution of arithmetic sequences and obtain surprising new results about certain arithmetic sequences. For example, we show that the forward T-orbit of every positive integer contains an element congruent to 2 mod 9, and every non-trivial cycle and divergent orbit contains an element congruent to 20 mod 27. We generalize these results to find many other sets that are strongly sufficient in this way.

Finally, we show that the $3x+1$ digraph exhibits a surprising and beautiful self-duality modulo 2^n for any n, and prove that it does not have this property for any other modulus. We then use deeper previous results to construct additional families of nontrivial strongly sufficient sets by showing that for any $k < n$, one can "fold" the digraph modulo 2^n onto the digraph modulo 2^k in a natural way.

Keywords: Collatz conjecture, arithmetic sequences, group actions, sufficient sets, 3x+1 digraph

*email: monks@college.harvard.edu
†email: monks@scranton.edu ; Corresponding author
‡email: monks@math.colostate.edu
§email: monks@math.berkeley.edu

1 Introduction

The $3x+1$ Conjecture, also known as the Collatz Conjecture, is a famous open problem in discrete dynamics. Attributed to L. Collatz in the 1930's, the conjecture states that if $T : \mathbb{Z} \to \mathbb{Z}$ is defined to be

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

then for any positive integer x, there is a nonnegative integer k for which $T^k(x) = 1$. In other words, the T-orbit of x (i.e. the sequence $x, T(x), T(T(x)), \ldots$) contains the number 1 among its elements.

Historically, the Collatz problem has been broken down into two smaller conjectures:

> **Nontrivial Cycles Conjecture:** There are no T-cycles of positive integers other than the cycle containing 1.
>
> **Divergent Orbits Conjecture:** There are no divergent T-orbits of positive integers.

Together, these two statements suffice to prove the $3x+1$ conjecture. Both remain unresolved.

There has recently been progress towards reducing the $3x+1$ problem to a seemingly simpler problem. We say that positive integers x and y *merge* if there exist nonnegative integers k and j for which $T^k(x) = T^j(y)$. A set of positive integers S is said to be *sufficient* if for every positive integer x, there is an element $y \in S$ that merges with x. Notice that to prove the $3x+1$ conjecture, it suffices to show that the T-orbit of every element of some sufficient set S contains 1 since every integer that merges with an element of S must also enter the cycle $\overline{1,2}$ as well. In [9], the third author shows that every arithmetic sequence is sufficient.

We can visualize these notions by drawing a directed graph associated to the dynamical system $T : \mathbb{Z} \to \mathbb{Z}$. Let $T_0(x) = x/2$ and $T_1(x) = (3x+1)/2$, and define the $3x+1$ *graph* $\mathcal{G}$ to be the two-colored directed graph on the positive integers having a black edge from x to z if $T_0(x) = z$, and a red edge (which are also dashed in the images in this paper for the benefit of those reading a black and white printout) from x to z if $T_1(x) = z$. (See Figure 1.1.) Then two integers merge if and only if they are in the same connected component of $\mathcal{G}$, and a sufficient set is a set of nodes which intersects every connected component of $\mathcal{G}$. The $3x+1$ conjecture is true if and only if $\mathcal{G}$ is connected.

In this paper we undertake a deeper investigation into the distribution of arithmetic sequences in the $3x+1$ graph and properties of sufficient sets in general. Define a *forward tracing path* to be a path in $\mathcal{G}$ along the directed arrows (which is simply an initial segment of a T-orbit), and a *back tracing path* to be a path in $\mathcal{G}$ against the direction of these arrows.

For brevity, we write "$a \underset{b}{\equiv} c$" in place of "$a \equiv c \pmod b$" throughout. Using this notation, we recall the well-known fact that if $x \underset{3}{\equiv} 0$, then the only way to form a back tracing path starting from x is by applying T_0^{-1} repeatedly, forming a single reverse chain of black arrows in $\mathcal{G}$. Moreover, every element of this chain is also divisible by 3. Conversely,

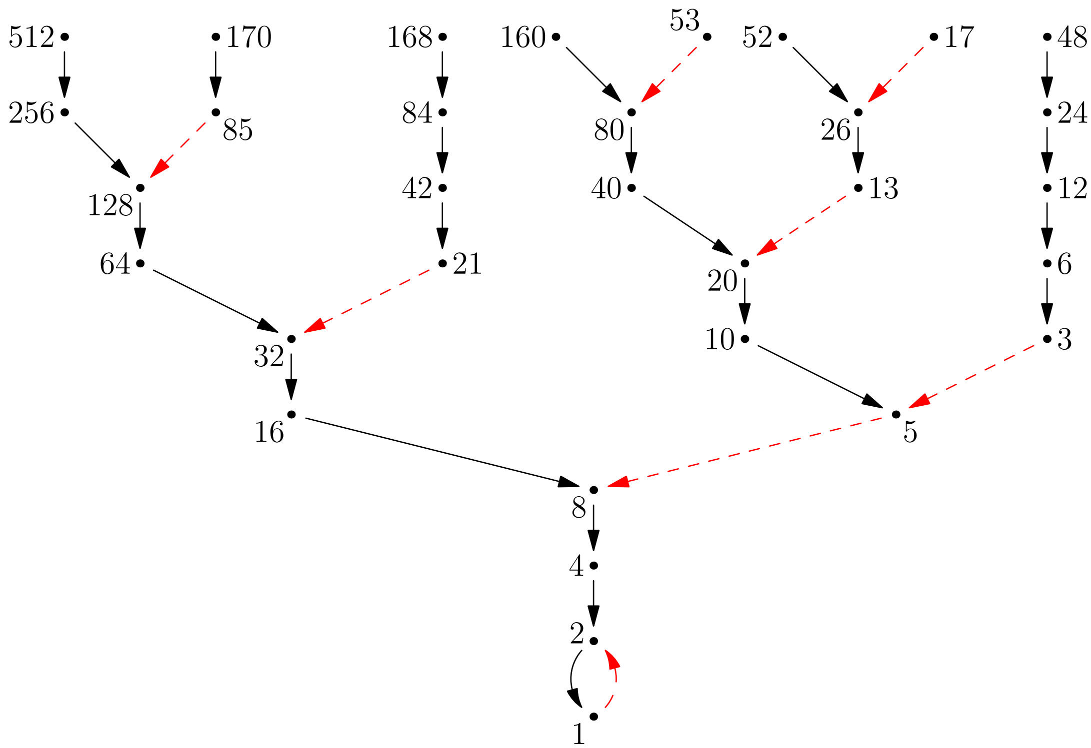

Figure 1.1: A portion of the $3x+1$ graph $\mathcal{G}$ near 1.

every positive integer x forward traces to a number that is not divisible by 3, at which point all future points in its orbit are not divisible by 3. Thus, the back tracing paths from multiples of 3 are well-understood, and for convenience we also define the *pruned* $3x+1$ *Graph*, denoted $\widetilde{\mathcal{G}}$, to be the subgraph of $\mathcal{G}$ consisting of the positive integers relatively prime to 3. A portion of the $\widetilde{\mathcal{G}}$ is illustrated in Figure 5.1.

In [9], in order to prove that every arithmetic sequence is sufficient, the third author shows that for any positive integers a and d and for every node x in the pruned $3x+1$ graph $\widetilde{\mathcal{G}}$, there is a back tracing path from x to some y in $\mathcal{G}$ with $y \in \{a + dn \mid n \in \mathbb{Z}\}$. (Since any x in the general $3x+1$ graph has a forward tracing path to some x' in the pruned graph, this shows that every x in $\mathcal{G}$ merges with some y in the given arithmetic sequence.) In section 4, we strengthen these results by finding bounds on the minimum number of red arrows needed to back trace from any positive integer x to an integer y in a desired arithmetic sequence. These results rely on an understanding of the $3x+1$ groups G_b generated by T_0 and T_1 modulo an integer b relatively prime to 6. We completely classify these groups in section 3.

For specific moduli d, we can obtain even stronger and more surprising results. In section 5, we describe infinite back tracing paths as elements of an inverse limit and show that every such back tracing path must contain an integer congruent to 2 mod 9. We use similar methods to show that in fact every T-orbit of positive integers must also contain an element congruent to 2 mod 9. For this reason, we say that the set of integers congruent to 2 mod 9 is *strongly sufficient*.

In Section 6, we define Γ_d to be the finite directed graph obtained by taking the $3x+1$ graph $\mathcal{G}$ mod d. We use these graphs to determine graph-theoretic criteria for a set to be strongly sufficient, and we provide several families of such sets S that must intersect every nontrivial T-orbit or infinite back tracing path. We also demonstrate that finding strongly sufficient sets is a plausible way to approach the Nontrivial Cycles Conjecture.

Finally, in Section 7, we show that the graphs Γ_{2^n} exhibit a surprising and beautiful self-duality (given by a map first defined in [8]) and that these are the only Γ_d having this property. We also use results from [10] and [6] to show that for any $k < n$, one can fold Γ_{2^n} onto Γ_{2^k} in a natural way. We combine these deeper tools with our results on strong sufficiency to obtain an infinite family of strongly sufficient sets consisting of unions of residue classes modulo a power of 2.

2 Sparse sufficient sets

Since every arithmetic sequence is sufficient, the arithmetic sequences form a family of sufficient sets with members having arbitrarily small positive density in the integers. It is natural to ask if there is a sufficient set of density zero in the positive integers. We answer this question in the affirmative as follows.

Theorem 2.1. *For any function $f : \mathbb{N} \to \mathbb{N}$ and any positive integers a and d, the set of integers*

$$\{2^{f(n)}(a + dn) \mid n \in \mathbb{N}\}$$

is a sufficient set.

Proof. We know that the set $\{a + dn \mid n \in \mathbb{N}\}$ is sufficient. So, given a positive integer x, there is a number of the form $a + dN$ that merges with x. Thus, the positive integer $2^{f(N)}(a + dN)$, which maps to $a + dN$ after $f(N)$ iterations of T, also merges with x. This completes the proof. □

By taking $f(n)$ to be sufficiently large relative to n, we can use this to produce infinitely many sufficient sets of density zero in the positive integers. We state one family of these below.

Corollary 2.2. *For any fixed a and d, the sequence $(a + dn) \cdot 2^n$ is a sufficient set with asymptotic density zero in the positive integers.*

Thus, to prove the $3x+1$ conjecture, it suffices to show that the T-orbit of every number in the density-zero set $\{(a + dn) \cdot 2^n \mid n > 0\}$ contains 1. This method can also be used to find arbitrarily sparse sufficient sets containing only odd numbers (for example, the set $\left\{\left(2^{2f(n)+1}(a + 3dn) - 1\right)/3 \mid n \in \mathbb{N}\right\}$ for $a \underset{3}{\equiv} 2$ and $f : \mathbb{N} \to \mathbb{Z}_+$). Notice, however, that the elements of any such sufficient sets eventually map to $a + dn$, so one still effectively needs to show that the elements of the arithmetic sequence $\{a + dn\}$ map to 1 under T. Thus we turn our attention to this problem next by investigating the actual distribution of such sequences in the $3x + 1$ graph.

3 Classification of the groups G_b

For any positive integer b relatively prime to 6, the functions T_0 and T_1 act as permutations on $\mathbb{Z}/b\mathbb{Z}$. We begin by completely classifying the permutation groups G_b generated by these two permutations.

Let C_r denote the cyclic group of order r. Also let

$$\mathrm{AGL}(1,b) = \{x \mapsto cx + d \mid d \in \mathbb{Z}/b\mathbb{Z}, c \in (\mathbb{Z}/b\mathbb{Z})^*\}$$

be the group of affine maps mod b under composition, and note that G_b can be viewed as the subgroup of $\mathrm{AGL}(1,b)$ generated by $T_0(x) = x/2$ and $T_1(x) = (3x+1)/2$. Moreover, the subgroup $\{x \mapsto x + a \mid a \in \mathbb{Z}/b\mathbb{Z}\}$ is isomorphic to the cyclic group C_b, and is a normal subgroup of $\mathrm{AGL}(1,b)$. It will also be useful to consider the element $P(x) = x + 1$ of G_b.

Theorem 3.1. *Let b be a positive integer relatively prime to 2 and 3. Write the prime factorization $b = p_1^{e_1} p_2^{e_2} \cdots p_n^{e_n}$. For all $i \in \{1, \ldots, n\}$, let s_i be the multiplicative order of* $2 \bmod p_i^{e_i}$, *let t_i be the multiplicative order of* $3 \bmod p_i^{e_i}$, *and let $a_i = \mathrm{lcm}(s_i, t_i)$. Then*

$$G_b \cong C_b \rtimes M$$

where

$$M = C_{a_1} \times C_{a_2} \times \cdots \times C_{a_n}.$$

In the statement above, the action on C_b defining the semidirect product is the action of $\mathrm{AGL}(1,b)$ by conjugation on the subgroup

$$\{x \mapsto x + a \mid a \in \mathbb{Z}/b\mathbb{Z}\} \cong C_b.$$

Proof. In [9], it was shown that the function $P(x) = x + 1$ is an element of the group generated by T_0 and T_1. This function clearly has order b and generates the cyclic subgroup $C_b = \{x \mapsto x + d : d \in \mathbb{Z}/b\mathbb{Z}\}$. This is a normal subgroup, since it is a fixed set under conjugation.

Since $x \mapsto x+1$ is in G_b, and T_0 and T_1 can be expressed in terms of the maps $x \mapsto x+1$, $x \mapsto 2x$, and $x \mapsto 3x$, we have that $G_b \subseteq \langle x \mapsto x+1, x \mapsto 2x, x \mapsto 3x \rangle$ as a subgroup of $\mathrm{AGL}(1,b)$. Moreover, $2x$ and $3x$ can be generated using T_0, T_1, and $x+1$, so in fact

$$G_b = \langle x \mapsto x+1, x \mapsto 2x, x \mapsto 3x \rangle .$$

The first generator corresponds to the cyclic subgroup C_b. We now only need to see what we obtain from multiplication by 2 and 3.

By the Chinese Remainder Theorem, we have $\mathbb{Z}/b\mathbb{Z} \cong C_{{p_1}^{e_1}} \times C_{{p_2}^{e_2}} \times \cdots \times C_{{p_n}^{e_n}}$. Thus we can look at the action of multiplication by 2 and 3 on each component, and the action on the whole group G_b will be the direct product of each of these.

Let $p \in \{p_1, p_2, \ldots, p_n\}$ and e be the corresponding exponent. Since b is relatively prime to 2, we know that p is an odd prime. Thus $(Z/p^eZ)^*$, the group of units of Z/p^eZ, is cyclic. Let s be the order of 2 and t the order of 3 in $(Z/p^eZ)^*$. Since the subgroup lattice of $(Z/p^eZ)^*$ is isomorphic to the divisor lattice of $\phi(p^e)$, we have that $\| \langle 2, 3 \rangle \| = \mathrm{lcm}(s,t)$, which concludes the proof. □

While Theorem 3.1 describes the overall structure of the groups, it would be useful to understand it as a finitely generated group in terms of the generators T_0 and T_1 (mod b). We begin by calculating the orders of T_0 and T_1 in G_b. To do so, we introduce the auxiliary function E.

Definition. Let $E_0(x) = 3x/2$ and $E_1(x) = (x+1)/2$, and define $E : \mathbb{Z}_+ \to \mathbb{Z}_+$ by

$$E(x) = \begin{cases} E_0(x) & x \text{ is even} \\ E_1(x) & x \text{ is odd} \end{cases}.$$

Let $P(x) = x + 1$. Then straightforward calculation shows that $E = PTP^{-1}$, and in particular that $E_0 = PT_1P^{-1}$ and $E_1 = PT_0P^{-1}$. Thus, the E-orbit of a positive integer $x > 1$ can be obtained by taking the T-orbit of $x - 1$ and adding 1 to each element of the orbit. Therefore the $3x + 1$ conjecture is equivalent to showing that the E-orbit of any positive integer $x > 1$ contains the integer 2.

Remark 3.2. The conjugacy between T and E via P makes computation of orbits somewhat easier: to compute the E-orbit of a positive integer $x > 1$, one first replaces any 2's in the prime factorization of x with 3's, one at a time, until the result is odd. At that point, one divides by 2 and rounds up to the nearest integer, and repeats the process. For instance, the E-orbit of 8 is

$$8, 12, 18, 27, 14, 21, 11, 6, 9, 5, 3, 2, \ldots,$$

which corresponds to the T-orbit of 7:

$$7, 11, 17, 26, 13, 20, 10, 5, 8, 4, 2, 1, \ldots.$$

Lemma 3.3. *Let b be a positive integer relatively prime to 2 and 3. The order of T_0 in G_b is equal to the order of 2 modulo b, and the order of T_1 in G_b is equal to the order of $\frac{3}{2}$ modulo b.*

Proof. The order of T_0 in G_b is equal to the order of $1/2$ modulo b, which is equal to the order of 2 modulo b.

For T_1, we have that T is conjugate to E on the positive integers via the map $x \mapsto x+1$. Therefore, T and E are also conjugate when considered as maps on $\mathbb{Z}/b\mathbb{Z}$. In particular, the conjugacy sends T_0 to E_1 and T_1 to E_0.

Now, the order of T_1 in G_b is equal to the order of E_0 in G_b by the conjugacy, and the order of E_0 is equal to the order of $\frac{3}{2}$ modulo b (since $E_0(x) = \frac{3}{2}x$). This completes the proof. □

In [9], it was shown that G_b acts transitively on $\mathbb{Z}/b\mathbb{Z}$, by showing that the map $P(x) = x + 1$ is in G_b. It is easy to check that the shortest representation for P map in terms of the generators T_0 and T_1 is

$$T_0^{-2} T_1 T_0 T_1^{-1} T_0 = P \tag{3.1}$$

and the corresponding result for the map E is

$$E_1^{-2} E_0 E_1 E_0^{-1} E_1 = P.$$

4 Uniformity in the distribution of arithmetic sequences

For this section and the next, we require some notation and basic results involving back tracing. We generally follow the notation used in Wirsching's book [12].

Define the set of *feasible vectors* to be

$$\mathcal{F} = \bigcup_{k=0}^{\infty} \mathbb{N}^{k+1}.$$

Let $s \in \mathcal{F}$. Then $s = (s_0, s_1, \ldots, s_k)$ for some nonnegative integers k and $s_0, s_1, \ldots, s_k$. The *length* of s, written $l(s)$, is k. The norm of s, written $||s||$, is $l(s) + \sum_{i=0}^{l(s)} s_i$.

For $s \in \mathcal{F}$ with $s = (s_0, s_1, \ldots, s_k)$, Wirsching calls the function $v_s : \mathbb{Z}^+ \to \mathbb{Q}$ given by $v_s = T_0^{-s_0} \circ T_1^{-1} \circ T_0^{-s_1} \circ T_1^{-1} \circ \cdots \circ T_1^{-1} \circ T_0^{-s_k}$ a *back tracing function.* If $v_s(x) \in \mathbb{Z}^+$ then we say s is an *admissible vector* for x, and that the corresponding back tracing function is a *admissible* for x. Define

$$\mathcal{E}(x) = \{s \in \mathcal{F} : s \text{ is admissible for } x\}.$$

Wirsching also shows that if $l(s) = m > 0$, then there is a unique congruence class a mod 3^m with a relatively prime to 3 such that, if x is any positive integer, s is an admissible vector for x if and only if $x \underset{3^m}{\equiv} a$.

Naturally it would be useful to strengthen the existence theorems in [9] to determine how a given arithmetic sequence is distributed in the $3x+1$ graph. More precisely, we wish to determine bounds for how far away the nearest element in a given arithmetic sequence $a + d\mathbb{N}$ is to a given positive integer x. We do so by first making precise the general bounds that follow from the proof of [9], Lemma 3.8, and strengthen those bounds for the special case where d is relatively prime to 2 and 3. In every case we show that the bounds obtained are independent of the choice of x, proving that arithmetic sequences are in this sense uniformly distributed in the $3x+1$ graph.

4.1 Back tracing modulo an arbitrary modulus d

We begin with the following bound for the length of a back tracing sequence to any modulus d.

Theorem 4.1. *Let $d > 1$ be a positive integer, and write $d = 2^n 3^m b$ where b is relatively prime to 2 and 3. Let $a \in \mathbb{N}$ with $a < d$, and let f be the order of $3/2$ modulo b. Then any $x \in \mathbb{N} - 3\mathbb{N}$ back traces to an element of $a + d\mathbb{N}$ via an admissible sequence of length at most $2(b-1)f + n + 1$.*

Remark 4.2. This bound depends only on the modulus d and not on the starting position x. This shows that the arithmetic sequence $a + d\mathbb{N}$ is, in some sense, "evenly distributed" throughout the $3x+1$ graph.

In order to prove this we first prove the case where $n = m = 0$, obtaining a stronger bound in this situation. The construction follows that of [9], Lemma 2.8. We sketch the proofs here and refer the reader to [9] for details.

Lemma 4.3. *Let b be a positive integer relatively prime to 2 and 3, and let $a \in \{0, 1, \ldots, b-1\}$ be any residue modulo b. Let f be the order of $3/2$ modulo b. Then for any positive integer x relatively prime to 3, there exists a admissible vector $s \in \mathcal{E}(x)$ for which $v_s(x) \underset{b}{\equiv} a$ and $v_s(x) \underset{3}{\not\equiv} 0$, such that*

$$l(s) \leq (b-1)f. \tag{4.1}$$

Proof. From equation (3.1) we have that $T_0^{-2} \circ T_1 \circ T_0 \circ T_1^{-1} \circ T_0(n) = n + 1$ for any n, and so trivially we have that

$$T_0^{-2} \circ T_1 \circ T_0 \circ T_1^{-1} \circ T_0(n) \underset{b}{\equiv} n + 1.$$

Let f be the order of $\frac{3}{2}$ modulo b, and let e be the order of 2 modulo b. Notice that since 3 is not congruent to 2 mod b, f is at least 2, and similarly e is at least 2. We also have $T_0 = T_0^{1-e}$ and $T_1 = T_1^{1-f}$ in G_b. Substituting, we obtain

$$T_0^{-2} \circ T_1 \circ T_0 \circ T_1^{-1} \circ T_0 = T_0^{-2} \circ T_1^{1-f} \circ T_0^{1-e} \circ T_1^{-1} \circ T_0^{1-e}.$$

Let $s_1 = (2, \underbrace{0, 0, \ldots, 0}_{f-2}, e-1, e-1)$, so that

$$v_{s_1}(n) = T_0^{-2} \circ T_1^{1-f} \circ T_0^{1-e} \circ T_1^{-1} \circ T_0^{1-e}(n) \underset{b}{\equiv} n + 1.$$

Notice that $l(s_1) = f$.

Now, let x be a positive integer relatively prime to 3, and define $s_2 = (0, 0) \cdot \underbrace{s_1 \cdot s_1 \cdot \cdots \cdot s_1}_{b}$. Since 2 is a primitive root mod every power of 3 (see, for instance, [5]), there is a positive integer k for which $s_2 \in \mathcal{E}(2^k x)$. Hence $s_2 \cdot (k) \in \mathcal{E}(x)$. It follows that any vector of the form $s_1 \cdot s_1 \cdot \cdots \cdot s_1 \cdot (k)$, where the number of copies of s_1 is at most b, is in $\mathcal{E}(x)$ as well.

Let $c = T_0^{-k}(x) \bmod b$, and let $s = \underbrace{s_1 \cdot s_1 \cdot \cdots \cdot s_1}_{(a-c) \bmod b} \cdot (k)$. Then we have

$$\begin{aligned} v_s(x) &\underset{b}{\equiv} c + (a - c) \\ &\underset{b}{\equiv} a \end{aligned}$$

and

$$\begin{aligned} l(s) &= ((a-c) \bmod b) \cdot f \\ &\leq (b-1)f. \end{aligned}$$

Finally, to see that $v_s(x) \underset{3}{\not\equiv} 0$, let $t = (a - c) \bmod b$ and let

$$s_3 = (0, 0) \cdot \underbrace{s_1 \cdot s_1 \cdot \cdots \cdot s_1}_{b-t}.$$

Then $s_3 \cdot s = s_2 \cdot (k)$, which is an admissible sequence for x, and so s_3 is admissible for $v_s(x)$. Since s_3 has length at least 1, we have that $v_s(x)$ is not divisible by 3, as desired. □

With this Lemma in hand, we can now prove Theorem 4.3.

Proof. First, by Lemma 4.3, we can back trace from x to some integer $y \underset{3}{\not\equiv} 0$ that is congruent to 0 modulo b via an admissible sequence of length at most $(b-1)f$. We can then apply T_0^{-n} to y to obtain an integer $z \in \mathbb{N} - 3\mathbb{N}$ that is congruent to 0 modulo $2^n b$.

We wish to back trace from z to an integer w with $w \underset{2^n b}{\equiv} a$ and $w \underset{3^m}{\equiv} a$. Following the arguments in [9], we can find a sequence $s \in \mathcal{E}(z)$ of length at most $(b-1)f + n + 1$ for which $v_s(z) \underset{2^n b}{\equiv} a$. Since 2 is a primitive root mod $3^{l(s)+m}$, there is a power of 2, say 2^k, such that $2^k z$ also has s as an admissible vector and moreover $v_s(2^k z) \underset{3^m}{\equiv} a$ (c.f. [9], Lemma 3.7). Thus, replacing s by $s \cdot (k)$, we can set $w = v_s(z)$, and we are done.

The total length of the back tracing sequence from x to w is then at most $(b-1)f + (b-1)f + n + 1 = 2(b-1)f + n + 1$, as desired. □

4.2 Back tracing when 2 is a primitive root of the modulus

We can improve this bound in some special cases, particularly when 2 is a primitive root modulo b. Since the only integers that can have a primitive root are 2, 4, p^r, and $2p^r$ where p is an odd prime, this implies that b must be a power of an odd prime.

Theorem 4.4. *Let r be a positive integer and p be an odd prime greater than 3 such that 2 is a primitive root modulo p^r. Let a be any residue modulo p^r relatively prime to p. Then for any positive integer x relatively prime to 3, there exists a admissible vector $s \in \mathcal{E}(x)$ for which $v_s(x) \underset{p^r}{\equiv} a$ and $v_s(x) \underset{3}{\not\equiv} 0$, such that*

$$l(s) \leq 1.$$

Proof. If x is relatively prime to p then since 2 is a primitive root, there exists k such that $2^k x \underset{p^r}{\equiv} a$. Clearly $2^k x \underset{3}{\not\equiv} 0$ since $x \underset{3}{\not\equiv} 0$. Thus taking $s = (k)$ gives the desired result.

If x is not relatively prime to p then since 2 is a primitive root mod 9 we can choose $k \geq 0$ such that $2^{k+1} x \underset{9}{\equiv} 4$. So $T_1^{-1} \circ T_0^{-k}(x) = \frac{2^{k+1}x-1}{3}$ is an integer that is relatively prime to both 3 and p. Thus there exists j such that $T_0^{-j} \circ T_1^{-1} \circ T_0^{-k}(x) = 2^j \left(\frac{2^{k+1}x-1}{3}\right) \underset{p^r}{\equiv} a$. Thus taking $s = (j, k)$ gives the desired result. □

Theorems 4.3 and 4.4 allow us to back trace to an integer in a desired congruence class mod b that is also not divisible by 3, so that we can continue back tracing to obtain more elements of the same congruence class. Thus, there is an infinite back tracing sequence $x_1, x_2, \ldots$ of elements in $\mathcal{G}$, satisfying $x_i = T(x_{i+1})$ for all i, that contains infinitely many elements congruent to a mod b. In section 5, we study infinite back tracing sequences in further depth.

5 Infinite back tracing and inverse limits

We first define infinite back tracing sequences in terms of inverse limits of *level sets* in $\mathcal{G}$, defined as follows.

Definition. Let x be a positive integer and let k be a nonnegative integer. The *kth level set of* x, which we denote $\mathcal{L}_k(x)$, is the set of all positive integers y for which $T^k(y) = x$.

Remark. This is a generalization of the notion of level set defined in [1], which referred only to the level sets of 1.

Definition. Let x be a positive integer. Consider the directed system $\{\mathcal{L}_k(x)\}_{k\geq 0}$ where the map from $\mathcal{L}_{k+1}(x)$ to $\mathcal{L}_k(x)$ is given by T. We define

$$\mathcal{I}_x = \varprojlim \mathcal{L}_k(x).$$

We also use the phrase *infinite back tracing sequence from* x to refer to an element of $\mathcal{I}_x$, or simply *infinite back tracing sequence* when x is understood.

Some of the elements of the sets $\mathcal{I}_x$ are rather simple to describe. For instance, recall that when x is divisible by 3, one can only ever apply T_0^{-1}, as the result will never be congruent to 2 mod 3. Thus the only infinite back tracing sequence from $x = 3y$ is $x, 2x, 4x, 8x, \ldots$. For this reason, we primarily are concerned with the elements of the $3x + 1$ graph which are not divisible by 3, and we define a modified version of the inverse limits for the pruned $3x + 1$ graph $\widetilde{\mathcal{G}}$, shown in Figure 5.1.

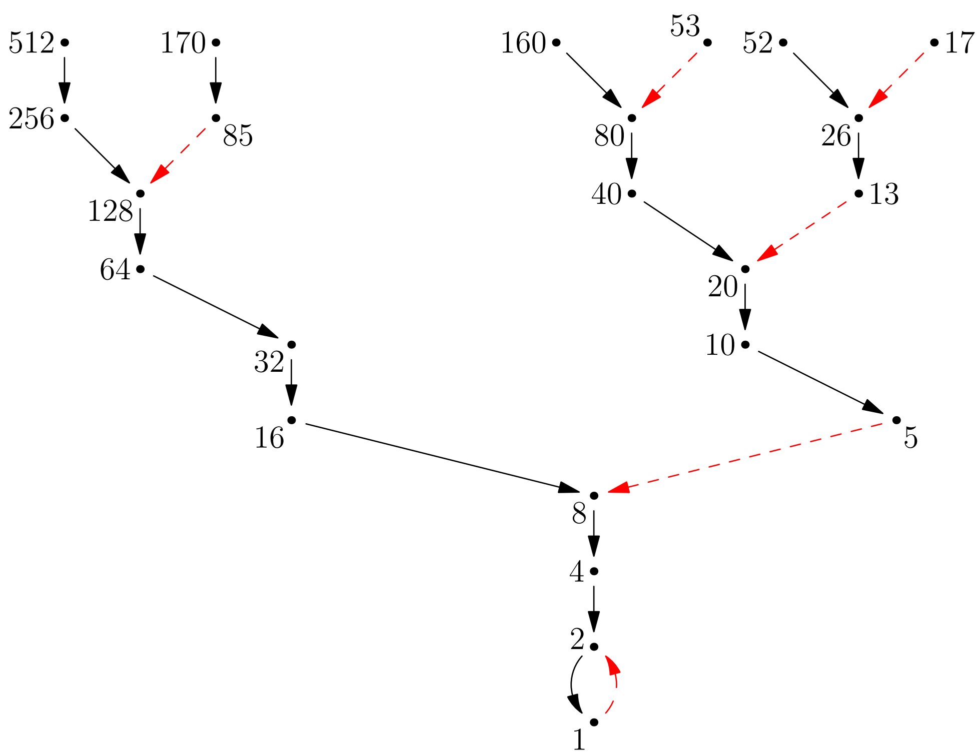

Figure 5.1: A portion of the pruned $3x + 1$ graph $\widetilde{\mathcal{G}}$ near 1.

Definition. Let $\widetilde{\mathcal{G}}$ denote the restriction of the $3x+1$ graph to the positive integers relatively prime to 3, and let x be one such positive integer. Let $\widetilde{\mathcal{L}}_k(x)$ be the set of all positive integers y in $\widetilde{\mathcal{G}}$ for which $T^k(y) = x$. Then we define

$$\widetilde{\mathcal{I}}_x = \varprojlim \widetilde{\mathcal{L}}_k(x).$$

Notice that $\widetilde{\mathcal{I}}_x$ is strictly contained in $\mathcal{I}_x$ for every x.

5.1 Structure of the inverse limits

We now investigate the structure of the inverse limit sets $\mathcal{I}_x$. To start, just as the (forward) parity vector of an integer determines its congruence class mod every power of 2, we can show that the parity of the values of an infinite back tracing sequence from x having infinitely many 1's determines the congruence class of x mod every power of 3, and hence determines the integer uniquely.

Definition. Let x be a positive integer. A *back tracing parity vector* (from x) is an infinite sequence of 0's and 1's that is congruent mod 2 to some infinite back tracing sequence from x.

Since we can expand any $(s_0, s_1, \ldots, s_n) \in \mathcal{F}$ to the parity vector

$$(\underbrace{0, \ldots, 0}_{s_0}, 1, \underbrace{0, \ldots, 0}_{s_1}, 1, \ldots, 1, \underbrace{0, \ldots, 0}_{s_n}, 1)$$

and vice versa, we can say that a finite back tracing parity vector is admissible for x if and only if the corresponding element of $\mathcal{F}$ is. An infinite back tracing parity vector is admissible for x if and only if every initial finite subsequence is.

Notice that for any 3-adic integer x, we can define $T_0^{-1}(x) = 2x$ and for $x \underset{3}{\equiv} 2$ we can also define $T_1^{-1}(x) = \frac{2x-1}{3}$. Hence the notion of a back tracing parity vector can be naturally extended to the 3-adic integers. Furthermore, for any positive integer k a 3-adic integer α is congruent to a unique ordinary integer a modulo 3^k, and thus a given back tracing vector is admissible for α if and only if it is admissible for a.

Theorem 5.1. *Let x be a 3-adic integer, and suppose v is a back tracing parity vector for x containing infinitely many 1's. If v is also a back tracing parity vector for the 3-adic integer y, then $x = y$.*

Proof. Let v_k be the smallest initial segment of the sequence v containing k 1's. Since v_k is admissible for both x and y, and since there is a unique congruence class modulo 3^k for which v_k is admissible, we must have $x \underset{3^k}{\equiv} y$. Since v_k exists and is finite for every k by assumption, it follows that $x \underset{3^k}{\equiv} y$ for every k and thus $x = y$. □

Since the positive integers embed naturally in the 3-adics, we can easily deduce a similar result for the positive integers for our purposes.

Corollary 5.2. *Let x be a positive integer, and suppose v is a back tracing parity vector for x containing infinitely many 1's. If v is also a back tracing parity vector for the positive integer y, then $x = y$.*

We now study properties of the back tracing vectors themselves. For the next result, we consider a back tracing parity vector as the binary expansion of a 2-adic integer.

Theorem 5.3. *Every back tracing parity vector, considered as a 2-adic integer, is either:*

(a) *a positive integer (i.e., only a finite number of the digits are nonzero),*

(b) *irrational, or*

(c) *immediately periodic (i.e., its binary expansion has the form* $\overline{v_0 \dots v_k}$ *where each* $v_i \in \{0,1\}$*).*

In particular, if the back tracing parity vector corresponds to an infinite back tracing sequence in $\mathcal{I}_1$*, and it is not the trivial cycle* 10101010…*, then it is either an integer or irrational.*

Proof. Let v be a back tracing parity vector for x.

It is known that a 2-adic is a rational number if and only if its binary expansion is eventually repeating (or immediately repeating). Thus, if the digits of v are never periodic, then v satisfies (b).

Now, suppose v is eventually repeating. If its repeating part contains only 0's, it satisfies (a). So, suppose its repeating part contains at least one 1. Let $v = v_0 v_1 \dots v_i \overline{v_{i+1} v_{i+2} \dots v_{i+j}}$, where one of $v_{i+1} \dots v_{i+j}$ is 1.

Since x is a positive integer and v is a back tracing parity vector for x, each initial segment of v must correspond to an admissible back tracing function for x. Thus, the value

$$x' = T_{v_0}^{-1} \circ \cdots \circ T_{v_i}^{-1}(x)$$

is an integer, and $\overline{v_{i+1} v_{i+2} \dots v_{i+j}}$ is a valid back tracing parity vector for x'.

Now, let

$$x'' = T_{v_{i+1}}^{-1} \circ \cdots \circ T_{v_{i+j}}^{-1}(x').$$

By a similar argument, x'' is an integer and $\overline{v_{i+1} v_{i+2} \dots v_{i+j}}$ is a valid back tracing parity vector for x''. By Theorem 5.1, it follows that $x'' = x'$. Thus, we have

$$x' = T_{v_{i+1}}^{-1} \circ \cdots \circ T_{v_{i+j}}^{-1}(x'),$$

which implies that $T^j(x') = x'$. Thus x' is a periodic point of T. But it is impossible to back trace from x into a cycle of T unless x itself is in the cycle. It follows that v is in fact immediately periodic, as desired. □

Notice that, to prove the nontrivial cycles conjecture, it suffices to show that the only periodic back tracing parity vector for any positive integer x is the 2-cycle $\overline{10}$.

The integer back tracing vectors are relatively easy to understand; they are formed by back tracing a finite number of steps, and then multiplying by 2 indefinitely. Occasionally one is forced into doing so, for T_1^{-1} can only be applied to integers congruent to 2 mod 3. If one first back traces to a multiple of 3, then multiplying by 2 will still result in a multiple of 3, and one can never apply T_1^{-1}.

The irrational back tracing parity vectors are not so easy to understand. As with most irrational numbers, it is difficult to write one down explicitly. However, we can bound the limiting fraction of 1's in the back tracing parity vectors as follows.

Lemma 5.4. *Let v be a back tracing parity vector of some positive integer x. Let k be the number of 1's among the first n digits of v and $p_n = k/n$. Then*

$$\limsup_{n\to\infty} p_n \leq \log_3(2) \approx 0.6309.$$

Proof. Let $f_n = T_{v_0}^{-1} \circ T_{v_1}^{-1} \circ \cdots \circ T_{v_{n-1}}^{-1}$ be the back tracing function corresponding to the first n digits of v. Then $f_n(x)$ is a positive integer for all n. Thus there is a minimum value among the values of $f_n(x)$. Let $f_{n_0}(x)$ be the first occurrence of this minimal value. Then for all $k \geq 0$, $f_{n_0}(x) \leq f_{n_0+k}(x)$.

Now, notice that $T_1^{-1}(y) < \frac{2}{3}y$ for all y. Therefore, if a function f is formed by composing i copies of T_1^{-1} and j copies of T_0^{-1}, we have $f(y) \leq \left(\frac{2}{3}\right)^i 2^j y$.

Let t be the number of occurrences of 1 among the first n_0 digits of v, and let r_k be the number of occurrences of 1 among the next $k+1$ digits. Then we have

$$f_{n_0}(x) \leq f_{n_0+k}(x) \leq (2/3)^r\, 2^{k+1-r} f_{n_0}(x)$$

and so $3^r \leq 2^{k+1}$. Taking the natural log of both sides, we find that $r\ln(3) \leq (k+1)\ln(2)$. Thus $r \leq (k+1)\log_3(2)$.

Finally, we have $p_{n_0+k} = \frac{r_k+t}{n_0+k} \leq \frac{(k+1)\log_3(2)+t}{n_0+k}$. Since t and n_0 are constant, the right hand side of this inequality tends to $\log_3(2)$ as k approaches infinity, and so the lim sup of the values of p_n is bounded above by this limit. $\square$

5.2 Greedy back tracing

While it is difficult to write down even one irrational infinite back tracing vector explicitly, there are several ways to obtain such vectors via a recursion. In particular, we can use a greedy algorithm that tries to keep the elements of the sequence as small as possible at each step, with the hopes of gaining insight into the structure of $\mathcal{G}$ by partitioning it into a union of the following greedy sequences.

Definition. Let x be a positive integer. The *greedy back tracing sequence* for x, denoted $\text{Greedy}(x)$, is the sequence of positive integers $a_0, a_1, \ldots$ defined recursively by $a_0 = x$ and for all $i > 0$

$$a_{i+1} = \begin{cases} T_1^{-1}(a_i) & \text{if } T_1^{-1}(a_i) \not\equiv_3 0 \\ T_0^{-1}(a_i) & \text{otherwise} \end{cases}$$

We also write V_x to denote $\text{Greedy}(x)$ taken mod 2, the back tracing parity vector of $\text{Greedy}(x)$.

It is easily verified that the recursion for $\text{Greedy}(x)$ can also be written as $a_0 = x$ and for all $i > 0$

$$a_{i+1} = \begin{cases} T_1^{-1}(a_i) & \text{if } a_i \equiv_9 2 \text{ or } a_i \equiv_9 8 \\ T_0^{-1}(a_i) & \text{otherwise} \end{cases}$$

We now show that for x relatively prime to 3, the back tracing parity vector V_x corresponding to $\text{Greedy}(x)$ has infinitely many 1's, and therefore that, for instance, V_4 is irrational.

Lemma 5.5. *Let x be a positive integer relatively prime to 3. Then V_x can have at most three 0's in a row at any point in the sequence.*

Proof. It suffices to show if y is an odd positive integer relatively prime to 3, the greedy algorithm applies T_0^{-1} at most three times before applying a T_1^{-1}.

Suppose y is an odd positive integer relatively prime to 3. Then it is congruent to one of 1, 2, 4, 5, 7, or 8 mod 9.

Case 1. Suppose y is congruent to 2 or 8 mod 9. Then the greedy algorithm applies T_1^{-1}, and we are done.

Case 2. Suppose y is congruent to 1 or 4 mod 9. Then the greedy algorithm determines that the next integer in the sequence is $T_0^{-1}(y) = 2y$, which is congruent to 2 or 8 mod 9. At this point, T_1^{-1} is applied, and we are done.

Case 3. Suppose y is congruent to 5 mod 9. Then the greedy algorithm applies T_0^{-1} to yield an integer congruent to 1 mod 9. Then, T_0^{-1} is applied again to obtain an integer congruent to 2 mod 9, and T_1^{-1} is applied.

Case 4. Suppose y is congruent to 7 mod 9. Then the greedy algorithm applies T_0^{-1} to yield an integer congruent to 5 mod 9, and by the above argument, two more T_0^{-1}'s are used before applying T_1^{-1}. □

We immediately obtain the following fact about greedy vectors.

Corollary 5.6. *Let x be a positive integer, let k be the number of 1's among the first n digits of v and $p_n = k/n$. Then*

$$\liminf_{n\to\infty} p_n \geq \frac{1}{4}.$$

Proof. If the nth term of V_x is 1, then each 1 in the first n terms is preceded by no more than three 0's, by Lemma 5.5. It follows that $p_n \geq 1/4$ in this case.

Otherwise, the nth term is 0, and the first n terms end in a string of k zeroes, where $1 \leq k \leq 3$. The first $n-k$ terms, however, have the property that each 1 is preceded by at most three 0's, so there are at least $(n-k)/4 \geq \frac{n-3}{4}$ ones among the first n terms. As n approaches infinity, the lower bound approaches $n/4$, and so $\liminf_{n\to\infty} p_n \geq \frac{1}{4}$, as desired. □

This gives a lower bound on the limiting percentage of 1's in a greedy back tracing vector. Since we are greedily choosing to apply T_1^{-1} whenever possible, it would be of interest to determine whether the greedy algorithm does maximize the percentage of 1's in a back tracing parity vector starting from x, and what that percentage is precisely. We leave this as an open problem for further study.

Having studied infinite back tracing sequences in some depth, we return to the problem of finding arithmetic progressions in the $3x + 1$ graph.

5.3 A strongly sufficient arithmetic progression

We can obtain surprising information about infinite back tracing sequences when we look modulo certain integers. We begin by proving the following remarkable fact.

Theorem 5.7. *Let x be a positive integer relatively prime to 3. Then every infinite back tracing sequence in $\widetilde{\mathcal{I}}_x$ contains a positive integer congruent to 2 mod 9.*

Proof. We first draw a directed graph to represent of the action of T_0 and T_1 on the elements of $\mathbb{Z}/9\mathbb{Z}$ relatively prime to 3, as shown in figure 5.2. Denote this directed graph by $\widetilde{\Gamma}_9$. Notice that any infinite back tracing sequence in $\widetilde{\mathcal{I}}_x$, taken mod 9, defines a sequence of residues traced out by an infinite path along the arrows in $\widetilde{\Gamma}_9$ in the *reverse* direction (against the arrows). We call such a path a *reverse path*.

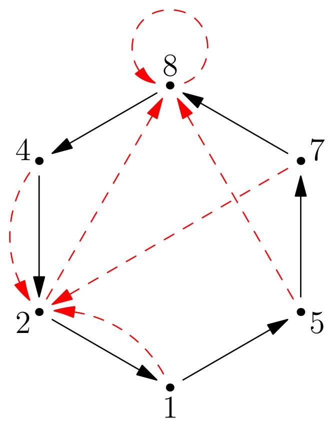

Figure 5.2: The action of T_0 and T_1 on the residues mod 9 relatively prime to 3.

Let $v \in \widetilde{\mathcal{I}}_x$ be an arbitrary back tracing sequence avoiding multiples of 3, and let P be the corresponding reverse path on Γ_9. Assume to the contrary that v is does not contain an integer congruent to 2 mod 9. Then the path P avoids the node labeled 2, and so it lies entirely in the subgraph Γ_9' shown in Figure 5.3.

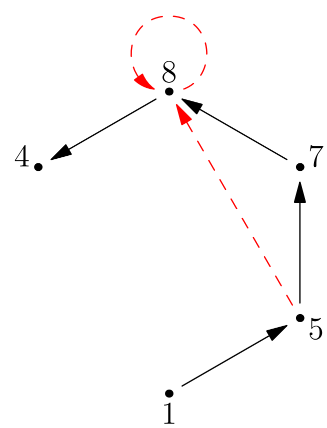

Figure 5.3: The subgraph Γ_9' of $\widetilde{\Gamma}_9$ formed by deleting the node labeled 2.

Now, since the path P in is infinite, it cannot contain the nodes 7, 5, or 1, since if we travel backwards along the edges from these nodes we must end up at 1, from which we cannot travel further. Furthermore, if P begins at the node 4, then it must travel to the node 8, where it is locked into the red loop at 8. But by Theorem 5.3, this is impossible. Thus P is the cyclic path $8, 8, 8, \ldots$.

Since P must consist only of red arrows, the back tracing parity vector corresponding to v is the all-1's vector, which is a valid back tracing parity vector for the integer -1. But by Theorem 5.1, it can therefore not be a valid back tracing parity vector for any other 3-adic integer, and in particular, it cannot be a valid back tracing parity vector for any positive integer. Thus, we have a contradiction, and so v must in fact contain an integer congruent to 2 mod 9. $\square$

Thus, for the arithmetic sequence $S = \{2 + 9n\}$, not only can we back trace from any $x \underset{3}{\not\equiv} 0$ to an element of S, but we cannot avoid doing so no matter how we back trace from x. We say such sets S are *strongly sufficient in the backward direction*, or simply *backward sufficient*.

Notice that the same argument applies to the forward direction: by looking at (ordinary, not reverse) paths in $\widetilde{\Gamma}_9$, we see that any forward T-orbit must contain an integer congruent to 2 mod 9.

Corollary 5.8. *The T-orbit of every positive integer contains an integer congruent to* 2 *mod* 9.

This essentially "proves the Collatz conjecture mod 9", and we say that $S = \{2 + 9n\}$ is *strongly sufficient in the forward direction*, or simply *forward sufficient*. We define these notions precisely in the next section.

6 Strong sufficiency and directed graphs

We define strong sufficiency in both the forward and backward directions, and also for the special case of nontrivial cycles, as follows.

Definition. Let S be a set of positive integers. Then

- S is *forward sufficient* if every divergent T-orbit contains an element of S.
- S is *cycle sufficient* if every nontrivial cycle contains an element of S.
- S is *backward sufficient* if for every positive integer x relatively prime to 3, every element of $\widetilde{\mathcal{I}}_x$ having an irrational back tracing parity vector contains an element of S.
- S is *strongly sufficient* if it is forward sufficient, cycle sufficient, and backward sufficient.

Notation. For simplicity in what follows, we write $\{a_1, \ldots, a_k \bmod d\}$ to denote the set of positive integers congruent to one of $a_1, a_2, \ldots, a_k$ mod d. We sometimes drop the brackets when the notation is clear. So we would say that $\{2 \bmod 9\}$, or 2 mod 9, is strongly sufficient.

Remark. In the case of 2 mod 9, we did not need to restrict our claim to the divergent orbits, the nontrivial cycles, and the aperiodic infinite back tracing sequences, because the cycle $\overline{1, 2}$ itself contains an element congruent to 2 mod 9. However, there are many sets S that intersect those T-orbits and back tracing sequences that do not end in $\overline{1, 2}$, but do not intersect *every* T-orbit simply because S does not contain 1 or 2. For this reason, we throw away the back tracing sequences and orbits that end in $\overline{1, 2}$ in our definition of strong sufficiency.

Notice also that it suffices to prove the $3x+1$ conjecture for the elements of any single strongly sufficient set, and that for any strongly sufficient set S, $S \cup \{1\}$ and $S \cup \{2\}$ are sufficient sets. Moreover, not only does every positive integer x *merge* with an element of $S \cup \{1\}$ (or $S \cup \{2\}$), but it actually *contains* one in its T-orbit and in every element of $\widetilde{\mathcal{I}}_x$. Hence the term "strong sufficiency."

Remark. Obtaining strongly sufficient sets give us a promising way in which to approach the nontrivial cycles conjecture. In particular, suppose we can show that a fixed, finite set of residues $a_1, \ldots, a_k$, modulo a set of arbitrarily large values of n, is strongly sufficient for each of these moduli n. Then any nontrivial cycle, being bounded, must contain one of the positive integers $a_1, \ldots, a_k$, and so we would only need to verify that the finite list of positive integers $a_1, \ldots, a_k$ have a T-orbit that contains 1.

In light of this remark, we begin a search for strongly sufficient sets. To do so, we first define a generalization of the directed graph Γ_9.

6.1 The graphs Γ_d and $\widetilde{\Gamma}_d$

We define the graph Γ_d to be the $3x+1$ graph $\mathcal{G}$ taken modulo d, as follows.

Definition. For a positive integer k, define Γ_k to be the two-colored directed graph on $\mathbb{Z}/k\mathbb{Z}$ such that

- there is a black arrow from r to s if and only if there exist positive integers x and y with $x \underset{k}{\equiv} r$ and $y \underset{k}{\equiv} s$ with $T_0(x) = y$, and
- there is a red arrow from r to s if and only if there exist positive integers x and y with $x \underset{k}{\equiv} r$ and $y \underset{k}{\equiv} s$ with $T_1(x) = y$.

As with the $3x+1$ graph $\mathcal{G}$, since we are primarily interested in the portions of T-orbits and infinite back tracing sequences whose elements are all relatively prime to 3, we also consider the pruned graph $\widetilde{\mathcal{G}}$ taken modulo d.

Definition. For $d \underset{3}{\equiv} 0$, the pruned graph $\widetilde{\Gamma}_d$ is the subgraph of Γ_d formed by deleting the nodes divisible by 3 (along with all of their adjacent edges). When $d \underset{3}{\equiv} 0$, we define $\widetilde{\Gamma}_d = \Gamma_d$.

Notice that when we refer to a node z in $\widetilde{\Gamma}_d$ or Γ_d we identify the congruence class z with the integer in $\{0, 1, \ldots, d-1\}$ that is in that class. For b relatively prime to 2 and 3, the graph Γ_b is a natural representation of the action of T_0 and T_1 on $\mathbb{Z}/b\mathbb{Z}$ in the group G_b. Examples are given in Figure 6.1.

We now demonstrate several basic properties of the graphs Γ_d for various d.

Proposition 6.1. *Let d be a positive integer.*

(a) *If d is even, each even node has two black arrows and no red arrows coming from it, each odd node has two red arrows and no black arrows coming from it. If d is odd, each node has exactly one red and black arrow coming from it.*

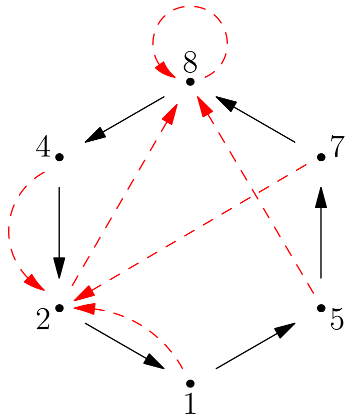

Figure 6.1: The digraphs $\widetilde{\Gamma}_7 = \Gamma_7$ (top left), $\widetilde{\Gamma}_8 = \Gamma_8$ (bottom), and Γ_9 (top right), which strictly contains $\widetilde{\Gamma}_9$.

(b) *If $d \underset{3}{\equiv} 0$, each node congruent to* 2 *modulo* 3 *has exactly one black arrow and at least red arrow pointing to it, and each other node has one black arrow and no red arrows pointing to it. If $d \underset{3}{\not\equiv} 0$, then every node has one black and one red arrow pointing to it.*

(c) *If d is relatively prime to* 2 *and* 3*, then in Γ_d, the black arrows form disjoint cycles on the vertices, as do the red arrows. There is one black loop at* 0*, and each of the other black cycles have length dividing the order of* 2 *mod d. There is one red loop at $d-1$, and each of the other red cycles have length dividing the order of* $3/2$ *mod d.*

(d) *If $d = 3^m$ for some m, then in Γ_d, the black arrows form a single cycle on the nodes which are relatively prime to* 3*. For $i = 1, \ldots, m-1$, there is also a cycle of black arrows consisting of the nodes divisible by 3^i but not by 3^{i+1}, and a black loop at the node* 0*.*

(e) *If $d = 3^m$ for some m, then in Γ_d, the red arrows form a rooted oriented tree with $3^m - 1$ as the root and all arrows oriented towards the root, plus a red loop at the root. The length of the shortest red path from any leaf to the root is m.*

Proof. Claim (c) follows from Lemma 3.3 and the fact that T_0 and T_1 generate a permutation group on $\mathbb{Z}/b\mathbb{Z}$.

For claim (a), note that dividing by 2 modulo some even d can be done in two ways: either $2r \mapsto r$ or $2r \mapsto r + \frac{1}{2}d$. Thus, if we send a congruence class x to $x/2$ or to $(3x+1)/2$ modulo d, in both cases we have exactly two possible results for the congruence class of $T(x)$ mod 2^n. It follows that each even node has two red arrows and two black arrows coming from it. If d is odd, then 2 is invertible modulo d, and so $T(x)$ is well defined.

For claim (b), note that $T_0^{-1}(x) = 2x$ is a well-defined function on $\mathbb{Z}/d\mathbb{Z}$ for all d, but $T_1^{-1}(x) = (2x-1)/3$ is well-defined if and only if d is not divisible by 3. Thus, if d is not divisible by 3, there is one red and one black arrow pointing to every node. If d is divisible by 3, however, then only those x congruent to 2 modulo 3 can have a red arrow pointing to it.

For part (d), it is known that 2 is a primitive root mod 3^m (see [5]), so the black arrows behave as described.

We now prove part (e). To do so, we consider the purely red back tracing paths starting at integers congruent to -1 mod 3^m. Suppose $x \underset{3^m}{\equiv} -1$, and let $M \geq m$ be the largest positive integer such that $x \underset{3^M}{\equiv} -1$. Then we can write $x = 3^M k - 1$ where k is relatively prime to 3.

Back tracing along a red arrow from x, we have that $T_1^{-1}(x) = (2x-1)/3 = 2 \cdot 3^{M-1}k - 1$, so $T_1^{-1}(x)$ is congruent to -1 mod $M-1$. If $M-1 \geq m$, we have that it is also congruent to -1 mod m. Thus, for the first $M-m$ steps in back tracing along red arrows, we follow a self-loop in Γ_{3^m} from $3^m - 1$ to itself. In particular, this loop exists in the graph, since there are positive integers congruent to -1 mod 3^M for any $M > m$.

Now, choose an integer x such that $M = m$, that is, m is the maximum positive integer for which $x \underset{3^m}{\equiv} -1$. Then by a similar argument, $T_1^{-1}(x) \underset{3^{m-1}}{\equiv} -1$, and by induction we have

$$(T_1^{-1})^k(x) \underset{3^{m-k}}{\equiv} -1$$

for all $k \geq 0$. Thus $(T_1^{-1})^{m-1}(x) \underset{3}{\equiv} -1$. It follows that $(T_1^{-1})^m(x)$ is congruent to either 0 or 1 mod 3, and so we cannot back trace using T_1^{-1} any further from here.

Note also that for each step in this process, the maximum M for which $(T_1^{-1})^k(x) \underset{3^M}{\equiv} -1$ is monotone decreasing by 1 at each step. Thus we can partition the congruence classes mod 3^m into grades based on the value of M, with the final grade consisting of those residues congruent to 0 or 1 mod 3, and we see that each element in the back tracing sequence from x is in a distinct grade. Moreover, each of these sequences, starting from $M = m$, has length m. It follows that the red arrows do indeed form a tree oriented towards the root at $3^m - 1$, with the shortest path from any leaf to the root having length m. □

6.2 Vertex minors and strong sufficiency

Using the digraphs Γ_d for various d, we can obtain several new strongly sufficient sets. In order to do so, we first give a graph-theoretic criterion for strong sufficiency.

Proposition 6.2. *Let $d \in \mathbb{N}$, and let $a_1, \ldots, a_k$ be k distinct residues mod d. Define Γ'_d to be the subgraph of $\widetilde{\Gamma}_d$ formed by deleting the nodes labeled $a_1, \ldots, a_k$ and all arrows connected to them and define Γ''_d to be the graph formed by deleting any edge from Γ'_d that is not contained in any cycle in Γ'_d. If Γ''_d is a disjoint union of cycles and isolated vertices, and each of the cycles have length less than* 630,138,897*, then the set*

$$a_1, \ldots, a_k \bmod d$$

is strongly sufficient.

Proof. Suppose Γ''_d is a disjoint union of cycles and isolated vertices, and each of the cycles have length less than 630,138,897.

We first show that $\{a_1, \ldots, a_k \bmod d\}$ is forward sufficient and cycle sufficient. Assume for contradiction that there is a positive integer x whose T-orbit, taken mod d, does *not* end in the cycle $\overline{1,2}$ and also avoids the set $\{a_1, \ldots, a_k \bmod d\}$. Since we are interested in the long-term behavior of this orbit, we may assume without loss of generality that x and all the elements of its T-orbit are relatively prime to 3, and hence trace out an infinite path P in $\widetilde{\Gamma}_d$. Since the T-orbit of x avoids $\{a_1, \ldots, a_k \bmod d\}$, it follows that P lies entirely in Γ'_d.

Now, for any edge e in Γ'_d that is not contained in any cycle, the path P contains e at most once. Thus, some infinite tail of the path P does not contain e, and so there is a T-orbit of some positive integer whose corresponding path does not contain e. We can thus assume without loss of generality that P does not pass through e.

Using the same argument on each such edge e, we can assume that P lies on the subgraph Γ''_d formed by deleting these edges. Since P is an infinite path, it must be contained in one of the loops of Γ''_d. Thus P is periodic. Its parity vector is also periodic, determined by the color of the edges on the loop, and so the T-orbit of x is periodic, corresponding to a nontrivial cycle with period equal to the length of the loop. But by our assumptions, the length of the loop is less than 630,138,897, and it is not the cycle $\overline{1,2}$. But there are no such positive integer cycles (see [11]), and so we have a contradiction.

For strong sufficiency in the backward direction, the same argument can be applied to the graph formed by reversing the arrows in $\widetilde{\Gamma}_d$, and hence in Γ'_d. □

Remark. If we remove the bound 630,138,897 on the length of the loops, the criterion shows that the set is forward and backward sufficient, but not necessarily cycle sufficient.

Using Proposition 6.2, we have obtained the list of strongly sufficient sets shown in Table 1.

6.3 Forward, backward, and cycle sufficiency

The sets in Table 1 are all strongly sufficient. In this section, we use more powerful tools to obtain sets that are not necessarily strongly sufficient, but are strongly sufficient in the forward or backward direction or cycle sufficient.

We require some known results on the limiting percentage of odd numbers in a T-orbit. In [3], Eliahou showed that if a T-cycle of positive integers of length n contains r odd positive integers (and $n-r$ even positive integers), and has minimal element m and maximal element M, then

$$\frac{\ln(2)}{\ln\left(3+\frac{1}{m}\right)} \leq \frac{r}{n} \leq \frac{\ln(2)}{\ln\left(3+\frac{1}{M}\right)}.$$

In [7], Lagarias showed a similar result for divergent orbits: the percentage of odd numbers in any divergent orbit is *at least* $\ln(2)/\ln(3) \approx .6309$.

We also require a similar bound for infinite back tracing sequences. Let x be a positive integer relatively prime to 3 and let $x = x_0, x_1, x_2, \ldots$ be an infinite back tracing sequence in $\widetilde{\mathcal{I}}_x$. Suppose further that the sequence is not periodic. Then by Proposition 5.3, its back tracing parity vector is either irrational or has only finitely many 1's.

In the case that the parity vector is irrational, note that every positive integer occurs at most a finite number of times in the sequence (otherwise, the sequence must be a cycle containing that integer). In particular, there is some N such that for all $n > N$, $x_n > x$. Now, consider the function f defined by the composition of the first n applications of T_0^{-1}

Strongly sufficient sets					
0 mod 2	1, 4 mod 9	1, 2, 6 mod 7	3, 4, 7 mod 10	2, 7, 8 mod 11	4, 5, 12 mod 14
1 mod 2	1, 8 mod 9	0, 1, 3 mod 8	3, 6, 7 mod 10	3, 4, 5 mod 11	4, 6, 11 mod 14
1 mod 3	4, 5 mod 9	0, 1, 6 mod 8	3, 7, 8 mod 10	3, 4, 8 mod 11	4, 11, 12 mod 14
2 mod 3	4, 7 mod 9	2, 4, 7 mod 8	4, 5, 7 mod 10	3, 4, 9 mod 11	6, 7, 8 mod 14
1 mod 4	5, 8 mod 9	2, 5, 7 mod 8	5, 6, 7 mod 10	3, 4, 10 mod 11	6, 8, 9 mod 14
2 mod 4	7, 8 mod 9	0, 1, 4 mod 10	5, 7, 8 mod 10	3, 6, 10 mod 11	7, 8, 12 mod 14
2 mod 6	4, 7 mod 11	0, 1, 6 mod 10	0, 1, 5 mod 11	1, 7, 10 mod 12	8, 9, 12 mod 14
2 mod 9	5, 6 mod 11	0, 1, 8 mod 10	0, 1, 8 mod 11	1, 8, 11 mod 12	1, 5, 7 mod 15
0, 3 mod 4	6, 8 mod 11	0, 2, 4 mod 10	0, 1, 9 mod 11	2, 4, 11 mod 12	1, 5, 11 mod 15
0, 1 mod 5	6, 9 mod 11	0, 2, 6 mod 10	0, 2, 5 mod 11	4, 7, 10 mod 12	1, 5, 13 mod 15
0, 2 mod 5	1, 5 mod 12	0, 2, 7 mod 10	0, 2, 8 mod 11	1, 3, 4 mod 13	1, 5, 14 mod 15
1, 3 mod 5	2, 5 mod 12	0, 2, 8 mod 10	0, 4, 5 mod 11	1, 4, 6 mod 13	1, 7, 8 mod 15
2, 3 mod 5	2, 8 mod 12	0, 4, 7 mod 10	0, 4, 8 mod 11	1, 8, 11 mod 13	1, 8, 13 mod 15
1, 4 mod 6	2, 10 mod 12	0, 6, 7 mod 10	0, 4, 9 mod 11	2, 3, 7 mod 13	1, 8, 14 mod 15
1, 5 mod 6	4, 5 mod 12	0, 7, 8 mod 10	1, 2, 7 mod 11	2, 6, 7 mod 13	1, 10, 11 mod 15
4, 5 mod 6	5, 8 mod 12	1, 3, 4 mod 10	1, 3, 5 mod 11	3, 4, 9 mod 13	1, 10, 13 mod 15
2, 3 mod 7	7, 8 mod 12	1, 3, 6 mod 10	1, 3, 8 mod 11	3, 4, 10 mod 13	2, 5, 7 mod 15
2, 5 mod 7	8, 11 mod 15	1, 3, 8 mod 10	1, 3, 9 mod 11	3, 7, 10 mod 13	2, 5, 11 mod 15
3, 4 mod 7	1, 8 mod 18	1, 4, 5 mod 10	1, 3, 10 mod 11	3, 10, 11 mod 13	2, 5, 13 mod 15
4, 5 mod 7	2, 8 mod 18	1, 5, 6 mod 10	1, 5, 7 mod 11	4, 6, 9 mod 13	2, 5, 14 mod 15
4, 6 mod 7	2, 11 mod 18	1, 5, 8 mod 10	1, 7, 8 mod 11	4, 6, 10 mod 13	2, 7, 8 mod 15
1, 4 mod 8	7, 8 mod 18	2, 3, 4 mod 10	1, 7, 9 mod 11	4, 8, 9 mod 13	2, 7, 10 mod 15
1, 5 mod 8	8, 10 mod 18	2, 3, 6 mod 10	2, 3, 5 mod 11	6, 7, 10 mod 13	2, 8, 13 mod 15
2, 3 mod 8	8, 14 mod 18	2, 3, 7 mod 10	2, 3, 7 mod 11	6, 10, 11 mod 13	2, 8, 14 mod 15
2, 6 mod 8	10, 11 mod 18	2, 3, 8 mod 10	2, 3, 8 mod 11	7, 8, 9 mod 13	2, 10, 11 mod 15
3, 4 mod 8	5, 11 mod 21	2, 4, 5 mod 10	2, 3, 9 mod 11	8, 9, 11 mod 13	2, 10, 13 mod 15
3, 5 mod 8	0, 1, 3 mod 7	2, 5, 6 mod 10	2, 3, 10 mod 11	8, 10, 11 mod 13	2, 10, 14 mod 15
4, 6 mod 8	0, 1, 5 mod 7	2, 5, 7 mod 10	2, 5, 7 mod 11	3, 4, 10 mod 14	4, 5, 11 mod 15
5, 6 mod 8	0, 1, 6 mod 7	2, 5, 8 mod 10	2, 6, 7 mod 11	4, 5, 6 mod 14	4, 10, 11 mod 15

Table 1: Some strongly sufficient sets. Each entry reveals a new property of the divergent T-orbits and nontrivial cycles. For instance, every divergent T-orbit, nontrivial cycle, and aperiodic infinite back tracing sequence in $\widetilde{\mathcal{G}}$ contains an element congruent to either 5 or 11 mod 21.

or T_1^{-1} in this back tracing sequence. Then since $T_1^{-1}(y) \leq 2y/3$ for any positive integer y, we have that $f(x) \leq \left(\frac{2}{3}\right)^r \cdot 2^{n-r}x$ where r is the number of 1's among the first n digits of the back tracing parity vector. It follows that

$$x < x_n = f(x) = \left(\frac{2}{3}\right)^r \cdot 2^{n-r}x$$

and therefore

$$1 < \left(\frac{2}{3}\right)^r \cdot 2^{n-r}.$$

Taking the natural log of both sides and solving for r/n, we obtain

$$\frac{r}{n} \leq \frac{\ln 2}{\ln 3}.$$

In the case that the parity vector has only finitely many 1's, there is clearly an N for which the same inequality holds for all $n > N$. Thus, the percentage of odd numbers in any aperiodic infinite back tracing sequence is *at most* $\ln(2)/\ln(3) \approx .6309$.

We summarize these results in the following proposition.

Proposition 6.3. *Let* $\rho = \ln(2)/\ln(3) \approx .6309$.

(a) *The percentage of odd numbers in any divergent orbit is at least* ρ.

(b) *The percentage of odd numbers in any nontrivial cycle with minimal element* m *and maximal element* M *is bounded below by* $\frac{\ln(2)}{\ln\left(3+\frac{1}{m}\right)}$ *and above by* $\frac{\ln(2)}{\ln\left(3+\frac{1}{M}\right)}$.

(c) *The percentage of odd numbers in any aperiodic infinite back-tracing sequence is at most* ρ.

Using this as a tool, we obtain the following result.

Theorem 6.4. *The arithmetic sequence* $\{20 \bmod 27\}$ *is forward sufficient and cycle sufficient.*

Proof. Consider the graph $\widetilde{\Gamma}_{27}$, drawn in Figure 6.2.

Now, suppose for contradiction that there is a nontrivial T-cycle or divergent T-orbit of positive integers which does not contain an integer congruent to 20 mod 27. Consider the path P on Γ_{27} formed by taking this T-orbit mod 27, starting at the first element which is not divisible by 3. Consider the subgraph Γ'_{27} of $\widetilde{\Gamma}_{27}$ formed by deleting the node 20 and all its adjacent edges. Since the path P does not contain the node 20 by assumption, we see that P lies entirely within Γ'_{27}.

Notice that in Γ'_{27}, the node 10 has no arrows coming into it, so it cannot occur more than once in the path P. Similarly, the nodes 5, 16, 13, and 26 cannot occur more than once in P. Thus, some infinite tail P' of the path P must lie in the subgraph Γ''_{27} shown in Figure 6.3.

Note that by Proposition 6.3, any nontrivial cycle has $m \geq 3$ and hence its percentage of odd elements is at least $\ln(2)/\ln(3 + 1/3) \approx 0.576$, and the percentage of odd elements in

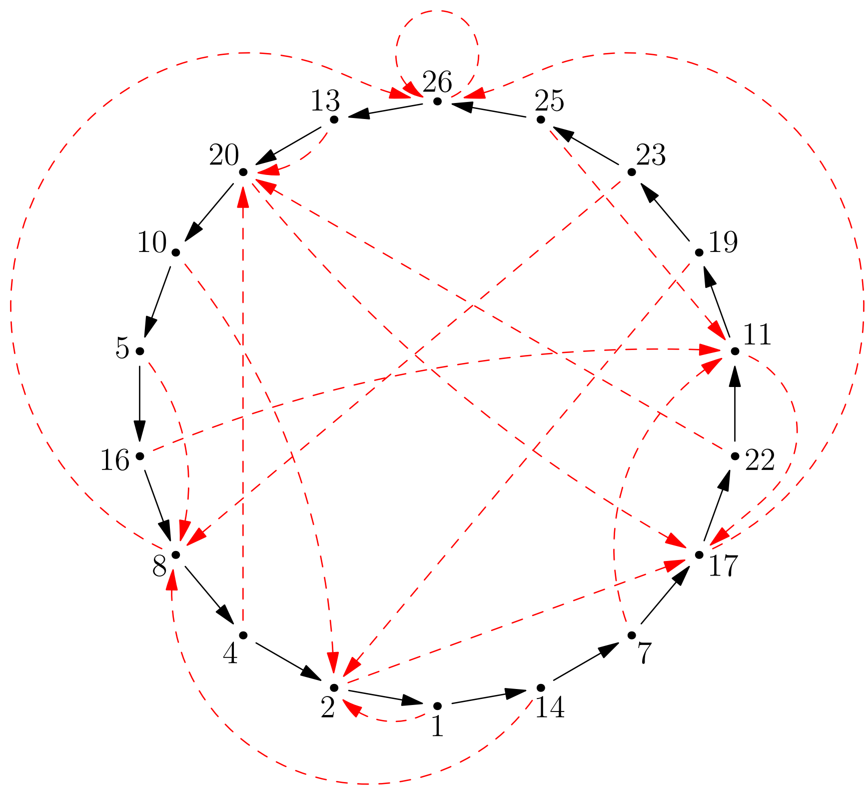

Figure 6.2: The action of T_0 and T_1 on the residues mod 27 relatively prime to 3.

any divergent T-orbit is at least 0.6309. We show that the fraction of red arrows followed by any infinite path in Γ''_{27} is at most 0.5, hence obtaining a contradiction.

We first note that any consecutive path of red arrows in Γ''_{27} has length at most 2. Moreover, any path of 2 consecutive red arrows (either from 19 to 2 to 17 or from 1 to 2 to 17) must be followed by at least 2 consecutive black arrows (from 17 to 22 to 11). It follows that the path P' has at most 50 percent red arrows, as desired. $\square$

The key observation in the proof above is that the graph Γ''_{27} essentially has too many black edges. We can use similar methods to obtain simple graph-theoretic criteria for strong sufficiency in the forward and backward directions and for cycle sufficiency.

Definition. A *simple cycle* in a directed graph is a directed path $A_1, \ldots, A_k$ of nodes for which $A_i = A_j$ if and only if $\{i, j\} = \{1, k\}$.

Proposition 6.5. *Let $d \in \mathbb{N}$, and let $a_1, \ldots, a_k$ be k distinct residues mod d. Let Γ'_d be the subgraph of $\widetilde{\Gamma}_d$ formed by deleting the nodes labeled $a_1, \ldots, a_k$ and all arrows connected to them, and let Γ''_d be the graph formed from Γ'_d by deleting any edge which is not contained in any cycle of Γ'_d.*

(a) *If the fraction of red arrows in every simple cycle of Γ''_d is less than $\ln(2)/\ln(3)$, then $\{a_1, \ldots, a_k \bmod d\}$ is forward sufficient.*

(b) *If the fraction of red arrows in every simple cycle of Γ''_d is greater than $\ln(2)/\ln(3)$, then $a_1, \ldots, a_k \bmod d$ is backward sufficient.*

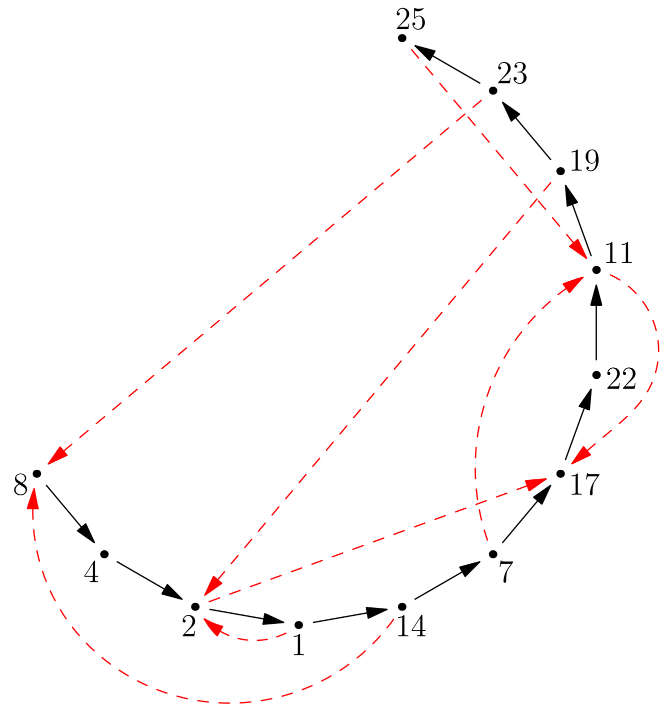

Figure 6.3: The graph Γ''_{27}.

(c) *If the fractions of red arrows in the simple cycles of each connected component H of Γ''_d are either all greater than $\ln(2)/\ln(3)$ or all less than $\ln(2)/\ln(3+1/m)$ where $m = 2^{60}$, then $a_1, \ldots, a_k \bmod d$ is cycle sufficient.*

Proof. Let $\rho = \ln(2)/\ln(3)$. First, suppose the fraction of red arrows in every simple cycle of Γ''_d is less than ρ. We show that every infinite path in Γ''_d must also have its limiting fraction of red arrows less than ρ, showing that $a_1, \ldots, a_k \bmod d$ is forward sufficient. Let P be an infinite path $A_1, A_2, \ldots$ of nodes in Γ''_d.

Since P is infinite and Γ''_d has a finite number of nodes, some node must occur infinitely many times in P. Call this node A. We show that the fraction of red arrows in the portion of P between any two consecutive occurrences of A is less than ρ. Let $A, B_1, \ldots, B_n, A$ be such a sub-path of P, and call this sub-path X.

To show that the fraction of red arrows along the path X must be less than ρ, we induct on an invariant which we call the *complexity* of X. Define the *complexity* of X to be the number of pairs of equal nodes in the sequence $A, B_1, \ldots, B_n, A$. For instance, the complexity of the sequence A, B, C, B, C, A is 3, and the complexity of the sequence $A, B, C, D, B, C, D, B, D, A$ is 8.

For the base case, suppose X has complexity 1. Then all of $B_1, \ldots, B_n$ are distinct, and so X is a simple cycle. By our hypothesis, the fraction of red arrows in X is less than ρ.

Let $n \geq 1$, and assume for strong induction that if X has complexity at most n then the fraction of red arrows in the path X is less than ρ. Suppose X has complexity $n+1$. Choose a node B other than A which occurs twice in X. Then we can write $X = u, B, v, B, w$ for some sequences of nodes u, v, and w.

Now, notice that the complexity of the sub-path B, v, B of X is strictly less than that of X, since it does not contain the two copies of A on each end. Letting a be the number of red arrows along this path and b the total number of arrows, we have that $a/b < \rho$ by the induction hypothesis.

Let X' be the cyclic path formed by deleting this cycle from X to form the sequence of nodes u, B, w. Then the complexity of X' is also less than that of X, so if c is the number of red arrows along X' and e is the total number of arrows, we have that $c/e < \rho$ by the induction hypothesis.

Finally, we have that $(a+c)/(b+e)$ is the fraction of red arrows in the entire path X. It is well-known that this *Farey sum*, also known as the *mediant* of the fractions a/b and c/e, must lie between a/b and c/e. Hence it must also be less than ρ. This completes the induction, proving the first claim.

The second claim is analogous. For the third claim, note that the $3x+1$ conjecture has now been verified for the positive integers less than 2^{60}, so any nontrivial cycle must have its minimal element m and maximal element M both greater than 2^{60}. Furthermore, any infinite periodic path lying in Γ''_d must lie entirely in one of the connected components of Γ''_d.

Assume that for all connected components H of Γ''_d, the fractions of red arrows in the simple cycles of H are either all greater than $\ln(2)/\ln(3)$ or all less than $\ln(2)/\ln(3+1/m)$ where $m = 2^{60}$. Suppose to the contrary that there is an infinite periodic path P in Γ''_d, and let H be the connected component containing it. If the simple cycles in H have fractions of red arrows less than $\ln(2)/\ln(3+1/m)$, then by the above argument, the fraction of red arrows in P is also less than $\ln(2)/\ln(3+1/m)$, contradicting Proposition 6.3. If instead the simple cycles in H have fractions of red arrows greater than $\ln(2)/\ln(3)$, then the fraction of red arrows in P is also greater than $\ln(2)/\ln(3) > \ln(2)/\ln(3+1/M)$, again contradicting Proposition 6.3. This completes the proof. $\square$

Using Proposition 6.5, we have obtained, with the use of a computer, several examples of forward sufficient, backward sufficient, and cycle sufficient sets that do not appear in Table 1. We list these results in Tables 2, 3, and 4.

7 Self duality and folding in Γ_{2^n}

We now use properties of the 2-adic dynamical system $T : \mathbb{Z}_2 \to \mathbb{Z}_2$ to provide a better understanding of the graphs Γ_{2^n}. We will use these insights to find more strongly sufficient sets from the ones we have already found.

7.1 Self color duality

The graphs Γ_{2^n} exhibit a surprising and beautiful self-duality.

Definition. Let Γ be any directed graph having each edge colored either red or black. The *color dual* of Γ is the graph formed by replacing all red edges with black edges and vice versa.

Forward sufficient sets				
$3 \bmod 4$	$11, 15 \bmod 16$	$0, 1, 6 \bmod 13$	$3, 9, 12 \bmod 14$	$7, 15, 19 \bmod 20$
$5 \bmod 6$	$4, 13 \bmod 18$	$0, 2, 3 \bmod 13$	$3, 9, 13 \bmod 14$	$9, 11, 15 \bmod 20$
$3 \bmod 8$	$11, 17 \bmod 18$	$0, 2, 6 \bmod 13$	$4, 5, 13 \bmod 14$	$11, 15, 18 \bmod 20$
$6 \bmod 8$	$13, 17 \bmod 18$	$0, 3, 9 \bmod 13$	$4, 11, 13 \bmod 14$	$11, 15, 19 \bmod 20$
$4 \bmod 9$	$5, 14 \bmod 21$	$0, 3, 10 \bmod 13$	$5, 6, 7 \bmod 14$	$10, 14, 17 \bmod 21$
$8 \bmod 9$	$5, 17 \bmod 24$	$0, 6, 9 \bmod 13$	$5, 6, 9 \bmod 14$	$13, 14, 17 \bmod 21$
$5 \bmod 12$	$11, 14 \bmod 24$	$0, 6, 10 \bmod 13$	$5, 7, 12 \bmod 14$	$14, 17, 20 \bmod 21$
$8 \bmod 18$	$11, 17 \bmod 24$	$0, 8, 9 \bmod 13$	$5, 7, 13 \bmod 14$	$3, 10, 17 \bmod 22$
$20 \bmod 27$	$11, 19 \bmod 24$	$1, 3, 7 \bmod 13$	$5, 9, 12 \bmod 14$	$3, 17, 20 \bmod 22$
$0, 3 \bmod 7$	$14, 20 \bmod 24$	$1, 3, 11 \bmod 13$	$5, 9, 13 \bmod 14$	$3, 17, 21 \bmod 22$
$0, 5 \bmod 7$	$14, 22 \bmod 24$	$1, 6, 7 \bmod 13$	$6, 7, 11 \bmod 14$	$4, 15, 19 \bmod 22$
$1, 7 \bmod 8$	$14, 23 \bmod 24$	$1, 6, 11 \bmod 13$	$6, 9, 11 \bmod 14$	$5, 16, 17 \bmod 22$
$4, 5 \bmod 11$	$17, 23 \bmod 24$	$2, 3, 4 \bmod 13$	$7, 8, 13 \bmod 14$	$8, 17, 19 \bmod 22$
$4, 8 \bmod 11$	$10, 17 \bmod 27$	$2, 3, 11 \bmod 13$	$7, 11, 12 \bmod 14$	$12, 13, 19 \bmod 22$
$2, 11 \bmod 12$	$13, 17 \bmod 27$	$2, 4, 6 \bmod 13$	$7, 11, 13 \bmod 14$	$4, 17, 22 \bmod 24$
$7, 10 \bmod 12$	$13, 22 \bmod 27$	$2, 6, 11 \bmod 13$	$8, 9, 13 \bmod 14$	$7, 17, 20 \bmod 24$
$7, 11 \bmod 12$	$17, 26 \bmod 27$	$2, 8, 11 \bmod 13$	$9, 11, 12 \bmod 14$	$7, 17, 22 \bmod 24$
$5, 11 \bmod 15$	$22, 26 \bmod 27$	$3, 7, 9 \bmod 13$	$9, 11, 13 \bmod 14$	$7, 19, 20 \bmod 24$
$3, 11 \bmod 16$	$1, 3, 9 \bmod 10$	$3, 9, 11 \bmod 13$	$1, 3, 7 \bmod 16$	$7, 19, 22 \bmod 24$
$6, 7 \bmod 16$	$1, 5, 9 \bmod 10$	$6, 7, 9 \bmod 13$	$1, 3, 9 \bmod 16$	$7, 19, 23 \bmod 24$
$6, 14 \bmod 16$	$3, 7, 9 \bmod 10$	$6, 9, 11 \bmod 13$	$1, 3, 14 \bmod 16$	$8, 17, 23 \bmod 27$
$7, 9 \bmod 16$	$5, 7, 9 \bmod 10$	$3, 6, 7 \bmod 14$	$2, 7, 12 \bmod 16$	$8, 17, 25 \bmod 27$
$7, 11 \bmod 16$	$1, 2, 5 \bmod 11$	$3, 6, 9 \bmod 14$	$2, 11, 12 \bmod 16$	$10, 11, 13 \bmod 27$
$9, 12 \bmod 16$	$1, 2, 8 \bmod 11$	$3, 7, 10 \bmod 14$	$2, 12, 14 \bmod 16$	$10, 11, 26 \bmod 27$
$9, 14 \bmod 16$	$1, 5, 9 \bmod 11$	$3, 7, 12 \bmod 14$	$5, 11, 16 \bmod 18$	
$9, 15 \bmod 16$	$1, 8, 9 \bmod 11$	$3, 7, 13 \bmod 14$	$7, 9, 15 \bmod 20$	
$11, 14 \bmod 16$	$0, 1, 3 \bmod 13$	$3, 9, 10 \bmod 14$	$7, 15, 18 \bmod 20$	

Table 2: Some forward sufficient sets obtained using the first criterion in Proposition 6.5.

Backward sufficient sets			
$2, 4 \bmod 8$	$1, 3, 5 \bmod 16$	$2, 4, 12 \bmod 16$	$1, 4, 20 \bmod 24$
$2, 5 \bmod 8$	$1, 3, 8 \bmod 16$	$2, 5, 12 \bmod 16$	$1, 5, 13 \bmod 24$
$1, 8 \bmod 12$	$1, 3, 10 \bmod 16$	$2, 5, 13 \bmod 16$	$1, 8, 20 \bmod 24$
$2, 4 \bmod 18$	$1, 4, 12 \bmod 16$	$2, 8, 12 \bmod 16$	$2, 4, 20 \bmod 24$
$2, 5 \bmod 24$	$1, 5, 13 \bmod 16$	$2, 8, 13 \bmod 16$	$2, 8, 20 \bmod 24$
$1, 4, 10 \bmod 12$	$1, 8, 13 \bmod 16$	$2, 10, 12 \bmod 16$	
$1, 3, 4 \bmod 16$	$2, 3, 10 \bmod 16$	$1, 4, 10 \bmod 18$	

Table 3: Some backward sufficient sets obtained using the second criterion in Proposition 6.5.

Cycle sufficient sets
$1, 3 \bmod 16$
$2, 12 \bmod 16$

Table 4: Some cycle sufficient sets obtained using the third criterion in Proposition 6.5.

Definition. A graph is *self color dual* if it is isomorphic to its color dual up to a relabeling of the vertices.

We give a complete classification of the self color dual graphs Γ_k.

Theorem 7.1. *The graph Γ_k is self color dual if and only if $k = 2^n$ for some positive integer n.*

To prove this, we require some terminology and background. Define $\mathbb{Z}_2$ to be the ring of 2-adic integers equipped with the usual 2-adic metric. The map T can be extended to be defined on $\mathbb{Z}_2$. Define the *parity vector function*

$$\Phi^{-1} : \mathbb{Z}_2 \to \mathbb{Z}_2$$

to be the map sending x to the T-orbit of x taken mod 2. Bernstein [2] shows that the inverse parity vector function Φ is well-defined, that is, the parity vector of a 2-adic uniquely determines the 2-adic. Moreover, Lagarias [7] shows that T is conjugate to the binary shift map

$$\sigma : \mathbb{Z}_2 \to \mathbb{Z}_2,$$

the map sending a 2-adic binary expansion $a_0a_1a_2a_3\ldots$ to the shifted 2-adic $a_1a_2a_3\ldots$, via the parity vector function Φ^{-1}. That is, $T = \Phi \circ \sigma \circ \Phi^{-1}$.

In [4], Hedlund shows that there are exactly two continuous *autoconjugacies* of the shift map (conjugacies from σ to σ), namely the identity map and the "bit complement" map $V : \mathbb{Z}_2 \to \mathbb{Z}_2$ given by

$$V(a_0a_1a_2\ldots) = b_0b_1b_2\ldots$$

where $b_i = 1 - a_i$ for all i. For instance, $V(100100100\ldots) = 011011011\ldots$.

In [8], the second author uses Hedlund's result to demonstrate that there are exactly two continuous autoconjugacies of T with itself. The identity map is one such map. The other, denoted $\Omega : \mathbb{Z}_2 \to \mathbb{Z}_2$, is the map

$$\Omega := \Phi \circ V \circ \Phi^{-1}.$$

We will use Ω to demonstrate self-color-duality in Γ_{2^n}. We use the fact that V is an involution, and hence Ω is an involution as well, that is, $\Omega^2 = 1$. In particular, we have that

$$T_1 \circ \Omega = \Omega \circ T_0$$

and

$$\Omega \circ T_0 = T_1 \circ \Omega$$

where these maps are defined.

Finally, a map $f : \mathbb{Z}_2 \to \mathbb{Z}_2$ is called *solenoidal* if it induces a permutation on $\mathbb{Z}/2^n\mathbb{Z}$ for all n. It is known ([4], [7], [8]) that the maps V, Φ, Φ^{-1}, and hence Ω are all solenoidal. Note that Ω therefore induces an involution on $\mathbb{Z}/2^n\mathbb{Z}$ as well.

We now have the tools to prove Theorem 7.1.

Proof. Let $n \geq 1$ and let $\Gamma^*_{2^n}$ denote the color dual of Γ_{2^n}. Let $\Gamma^\Omega_{2^n}$ denote the graph formed from Γ_{2^n} by replacing each node label a with $\Omega(a)$ mod 2^n. We show that $\Gamma^\Omega_{2^n} = \Gamma^*_{2^n}$, from which it follows that $\Gamma^*_{2^n}$ is isomorphic to Γ_{2^n} up to a relabeling of the nodes.

Suppose that in $\Gamma^*_{2^n}$, there is a red arrow from a to b. Then in Γ_{2^n}, there is a black arrow from a to b. It follows that there are positive integers x and y congruent to a and b mod 2^n respectively for which $T_0(x) = y$. Therefore $\Omega(T_0(x)) = \Omega(y)$, and hence $T_1(\Omega(x)) = \Omega(y)$. Thus, in Γ_{2^n}, there is a red arrow from $\Omega(a)$ to $\Omega(b)$. Since Ω is an involution, in $\Gamma^\Omega_{2^n}$, there is a red arrow from $\Omega(\Omega(a)) = a$ to $\Omega(\Omega(b)) = b$.

Similarly, if there is a black arrow from a to b in $\Gamma^*_{2^n}$ then there is a black arrow from a to b in $\Gamma^\Omega_{2^n}$.

For the reverse direction, suppose that in $\Gamma^\Omega_{2^n}$ there is a red arrow from a to b. Then in Γ_{2^n}, there is a red arrow from $\Omega(a)$ to $\Omega(b)$. Thus there are positive integers x and y congruent to a and b mod 2^n respectively for which $T_1(\Omega(x)) = \Omega(y)$. Thus $\Omega(T_0(x)) = \Omega(y)$, and since Ω is an involution, we have $T_0(x) = y$. It follows that there is a red arrow from a to b in $\Gamma^*_{2^n}$.

A similar argument shows that if there is a black arrow from a to b in $\Gamma^\Omega_{2^n}$ then there is a black arrow from a to b in $\Gamma^*_{2^n}$. This shows that Γ_{2^n} is self color dual.

To prove that no other Γ_k is self color dual, let $k = 2^n b$ where b is an odd positive integer greater than 1 and assume that $\Gamma_{2^n b}$ is self color dual. Then there exists a graph isomorphism $\rho : \Gamma_{2^n b} \to \Gamma_{2^n b}$ mapping red arrows to black ones and vice versa.

For any node z in $\Gamma_{2^n b}$ define $\hat{T}_0(z)$ to be the set of nodes w such that there is a black arrow from z to w and $\hat{T}_1(z)$ to be the set of nodes w such that there is a red arrow from z to w. Furthermore, for any nonnegative integer k define $\hat{T}^k(z)$ to be the set of nodes that can be reached starting from z by a path of length k. Clearly the graph isomorphism ρ must preserve the number of nodes that can be reached in such a manner, i.e.

$$\left|\hat{T}^k(z)\right| = \left|\hat{T}^k(\rho(z))\right| \tag{7.1}$$

for any z and k.

Suppose a node z has a black arrow from z to itself. Then by the proof of Proposition 6.1, if $n > 0$ then z is even and there exists an even integer $2a$ congruent to z modulo $2^n b$ such that either $T_0(2a) = a$ or $T_0(2a + 2^n b) = a + 2^{n-1}b$ is congruent to z, and thus to $2a$, modulo $2^n b$. Thus either $a \underset{2^n b}{\equiv} 0$ or $a \underset{2^n b}{\equiv} 2^{n-1}b$ so that in both cases $2a$, and thus z, must be congruent to 0 modulo $2^n b$. A similar argument shows that the only node z that has a red arrow from z to itself is -1. Since any color reversing graph isomorphism must map these nodes to each other, $\rho(-1) = 0$ and $\rho(0) = -1$.

We now show by finite induction that for any $k \in \{0, 1, \ldots, n\}$, $\hat{T}^k(-1)$ is the set of all nodes z such that $z \underset{2^{n-k}b}{\equiv} -1$. For the base case, notice that $\hat{T}^0(z) = \{z\}$ so that in particular $\hat{T}^0(-1) = \{-1\}$, i.e. the set of nodes that are congruent to -1 modulo $2^n b$. If $n = 0$ then we are done. If not, let $k < n$ and assume that $\hat{T}^k(-1) = \left\{ z \mid z \underset{2^{n-k}b}{\equiv} -1 \right\}$ which is a set of odd nodes. Then $\hat{T}^{k+1}(-1)$ is the set of nodes obtained by following a red arrow from a node $z \in \hat{T}^k(-1)$. Since $z \underset{2^n b}{\equiv} -1 + 2^{n-k}bj$ for some j, and $T_1\left(-1 + 2^{n-k}bj\right) = -1 + 3 \cdot 2^{n-(k+1)}bj \underset{2^{n-(k+1)}b}{\equiv} -1$ it follows that z is in the set of all nodes that are congruent to -1 modulo $2^{n-(k+1)}b$. Conversely if w is congruent to -1 modulo $2^{n-(k+1)}b$, then $w = -1 + 2^{n-(k+1)}bl$ for some l and thus is congruent modulo $2^{n-k}b$ to

$$-1 + 2^{n-(k+1)}bl + 2^{n-k}b = -1 + 3 \cdot 2^{n-(k+1)}bl = T_1\left(-1 + 2^{n-k}bl\right)$$

Since $-1 + 2^{n-k}bl \underset{2^{n-k}b}{\equiv} -1$ it is congruent to an element of $\hat{T}^k(-1)$ and so there is a red arrow from an element of $\hat{T}^k(-1)$ to w. Thus $\hat{T}^{k+1}(-1)$ is the set of nodes that are congruent to -1 modulo $2^{n-(k+1)}b$, which completes the induction.

A similar argument shows that $\hat{T}^k(0)$ is the set of all nodes z such that $z \underset{2^{n-k}b}{\equiv} 0$ for all $k \in \{0, 1, \ldots, n\}$. Since the graph isomorphism ρ must map the set of nodes that are reachable by a path of length n from -1 to the set of nodes reachable by a path of length n from $\rho(-1) = 0$ we have that ρ maps the set of nodes congruent to -1 modulo b to those congruent to 0 modulo b.

Now b is odd, so 2 is invertible modulo b. Let $z \underset{b}{\equiv} -1$ and not congruent to -1 modulo $2b$. Then z is even and $T_0(z) \underset{b}{\equiv} -\frac{1}{2}$. Conversely, if $w \underset{b}{\equiv} -\frac{1}{2}$ then $w = T_0(z)$ for some even $z \underset{b}{\equiv} -1$. Thus every node $z \underset{b}{\equiv} -1$ that is not congruent to -1 modulo $2b$ has a black arrow from z to a node w congruent to $-\frac{1}{2}$ modulo b and every node $w \underset{b}{\equiv} -\frac{1}{2}$ has such an arrow pointing to it. Since we have seen that all other nodes congruent to -1 modulo b only have arrows pointing to other such nodes, $\hat{T}^{n+1}(-1)$ consists of all nodes congruent to either -1 or $-\frac{1}{2}$ modulo b.

Similar arguments show that $\hat{T}^{n+2}(-1)$ consists of all nodes congruent to either -1, $-\frac{1}{2}$, or $-\frac{1}{4}$ modulo b and that $\hat{T}^{n+2}(0)$ consists of all nodes congruent to either 0, $\frac{1}{2}$, $\frac{1}{4}$, or $\frac{5}{4}$ modulo b.

If b is odd and greater than 5, directly counting these nodes shows that $\left|\hat{T}^{n+2}(-1)\right| = 3 \cdot 2^n$ while $\left|\hat{T}^{n+2}(\rho(-1))\right| = \left|\hat{T}^{n+2}(0)\right| = 4 \cdot 2^n$ contradicting 7.1. If $b = 3$, directly counting these

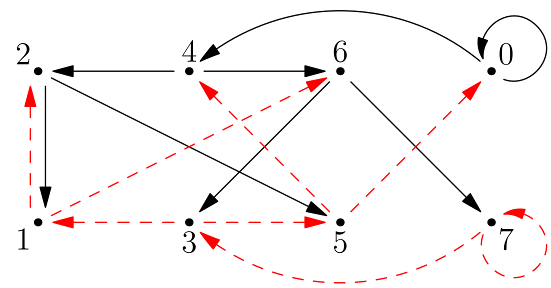

Figure 7.1: The digraph Γ_8. Self color duality is evident by reflecting about the horizontal.

nodes shows that $\left|\hat{T}^{n+2}(-1)\right| = 2 \cdot 2^n$ while $\left|\hat{T}^{n+2}(\rho(-1))\right| = \left|\hat{T}^{n+2}(0)\right| = 3 \cdot 2^n$ again contradicting 7.1.

Finally, suppose $b = 5$. Then $\hat{T}^{n+1}(-1) \setminus \hat{T}^n(-1)$ consists of the nodes congruent to $-\frac{1}{2}$ modulo 5 (i.e. the nodes congruent to 2 mod 5). All arrows from these nodes point to a node congruent to $-\frac{1}{4}$ modulo 5 (i.e. the nodes congruent to 1 mod 5). No node congruent to 1 mod 5 is in $\hat{T}^n(-1)$ as these are all congruent to -1 mod 5. Similarly, $\hat{T}^{n+1}(0) \setminus \hat{T}^n(0)$ consists of the nodes congruent to $\frac{1}{2}$ modulo 5 (i.e. the nodes congruent to 3 mod 5).

Since ρ is a graph isomorphism, it must map the set of nodes $\hat{T}^n(-1)$ to $\hat{T}^n(0)$ and $\hat{T}^{n+1}(-1) \setminus \hat{T}^n(-1)$ to $\hat{T}^{n+1}(0) \setminus \hat{T}^n(0)$ and also preserve the property that no arrow coming from a node in $\hat{T}^{n+1}(0) \setminus \hat{T}^n(0)$ can map to a node in $\hat{T}^n(0)$. But since $T(3) = 5$, the node 3 has a red arrow mapping it to the node 5, which is in $\hat{T}^n(0)$ as these are all the nodes congruent to 0 mod 5. This is a contradiction, which completes the proof. $\square$

To illustrate Theorem 7.1, the graph Γ_8 is shown in Figure 7.1, with each odd residue drawn directly above its image under Ω.

7.2 Folding

An *endomorphism* of a map $f : \mathbb{Z}_2 \to \mathbb{Z}_2$ is any map $h : \mathbb{Z}_2 \to \mathbb{Z}_2$ for which $f \circ h = h \circ f$. Note that an endomorphism is not necessarily invertible, and so while all autoconjugacies of T are endomorphisms of T, there may be endomorphisms which are not autoconjugacies.

In [10], the fourth author classified and studied all continuous endomorphisms of T having solenoidal parity vector functions, and in [6], the first author and Kraft studied the remaining continuous endomorphisms of T. It is natural to ask whether these endomorphisms yield further insights into the structure of the graphs Γ_{2^n}.

The simplest example of a continuous endomorphism of T which is not an autoconjugacy is defined in [10] as follows. Let $D : \mathbb{Z}_2 \to \mathbb{Z}_2$ be the *discrete derivative* map, given by $D(a_0a_1a_2\ldots) = d_0d_1d_2\ldots$ where $d_i = |a_i - a_{i+1}|$ for all i. Then

$$R := \Phi \circ D \circ \Phi^{-1}$$

is an endomorphism of T.

Unlike Ω, the function R is not solenoidal, since D is not solenoidal. However, the value of x mod 2^n determines the value of $D(x)$ mod 2^{n-1} for all n. In particular, D induces a

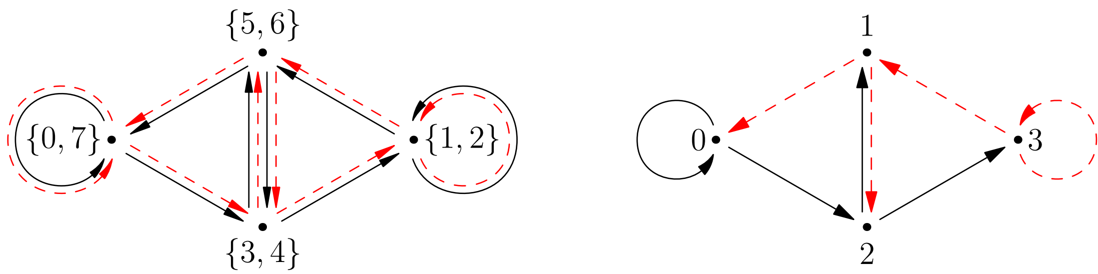

Figure 7.2: At left, the graph formed by identifying the pairs of nodes in Γ_8 that map to each other under Ω. At right, the graph Γ_4.

2-to-1 map $\mathbb{Z}/2^n\mathbb{Z} \to \mathbb{Z}/2^{n-1}\mathbb{Z}$, with $D(x) = D(V(x))$ for all x. Thus R also induces a 2-to-1 map $\mathbb{Z}/2^n\mathbb{Z} \to \mathbb{Z}/2^{n-1}\mathbb{Z}$, with $R(x) = R(\Omega(x))$ for all x. We therefore obtain the following.

Proposition 7.2. *Let $n \geq 2$ be a positive integer.*

(a) *For any $x, y \in \mathbb{Z}/2^n\mathbb{Z}$, there is a black edge between $R(x) \bmod 2^{n-1}$ and $R(y) \bmod 2^{n-1}$ in $\Gamma_{2^{n-1}}$ if and only if there is a path of length two in Γ_{2^n} from x to y that consists of either two black or two red edges.*

(b) *For any $x, y \in \mathbb{Z}/2^n\mathbb{Z}$, there is a red edge between $R(x) \bmod 2^{n-1}$ and $R(y) \bmod 2^{n-1}$ in $\Gamma_{2^{n-1}}$ if and only if there is a path of length two in Γ_{2^n} from x to y that consists of one black and one red edge.*

In other words, Γ_{2^n} "folds" onto $\Gamma_{2^{n-1}}$ by identifying Ω-pairs and using D to define the edges. For $n = 3$, the graph Γ_8 shown in Figure 7.1 can be folded to obtain the graph Γ_4, by identifying the Ω-pairs of nodes and drawing in new edges according Proposition 7.2. (See Figure 7.2.)

More generally, we can fold the graphs Γ_{2^n} onto any Γ_{2^t} for $t \leq n$ in a similar manner using the endomorphisms studied in [6]. For each $k \geq 2$, define $M_k : \mathbb{Z}_2 \to \mathbb{Z}_2$ to be the map given by $M_k(a_0a_1a_2\ldots) = m_0m_1m_2\ldots$ where

$$m_i = a_i + a_{i+1} + \cdots + a_{i+k-1} \bmod 2$$

for all i. Then

$$H_{M_k} := \Phi \circ M_k \circ \Phi^{-1}$$

is an endomorphism of T. Note that $M_2 = D$ and $H_{M_2} = R$.

The aim of this section is to prove the following result, which enables us to obtain more strongly sufficient sets modulo powers of 2.

Theorem 7.3. *Suppose $a_1, a_2, \ldots, a_l \bmod 2^n$ satisfy the criterion for strong sufficiency of Proposition 6.2 for $d = 2^n$. Let q be the length of the largest cycle in Γ''_{2^n}, and let k be any positive integer satisfying $kq \leq 630{,}138{,}897$. Then the preimage of $\{a_1, a_2, \ldots, a_l\}$ under H_{M_k} modulo 2^{n+k-1} also satisfies the criterion from Proposition 6.2, and is therefore a strongly sufficient set.*

Example. In Table 1, we see that 1 mod 4 satisfies the criterion for strong sufficiency of Proposition 6.2. We see from Figure 7.2 that the inverse image of $\{1 \bmod 4\}$ under H_{M_2} is the set $\{3, 4 \bmod 8\}$, which is therefore also strongly sufficient.

Indeed, $\{3, 4 \bmod 8\}$ also appears in Table 1. We can therefore unfold this set another time under H_{M_2}, which shows that $\{7, 8, 9, 10 \bmod 16\}$ is also strongly sufficient.

To prove Theorem 7.3, we introduce the notation established in [6], for any $a \in \mathbb{Z}_2$, write $[a]$ to denote the equivalence class of a under the equivalence relation $a \sim b$ if and only if $H_{M_k}(a) = H_{M_k}(b)$, i.e. $[a] = H_{M_k}^{-1}(\{H_{M_k}(a)\})$. Notice that H_{M_k} restricts to a well-defined surjective map $\overline{H}_{M_k} : \mathbb{Z}/2^{n+k-1}\mathbb{Z} \to \mathbb{Z}/2^n\mathbb{Z}$ for any n and $k > 0$. We also use $\overline{a}$ to denote either the congruence class of $a \in \mathbb{Z}_2$ modulo 2^n or 2^{n+k-1} when the power of 2 is understood, and using this notation we have $\overline{H_{M_K}(a)} = \overline{H}_{M_k}(\overline{a})$. Hence, $\sim$ also restricts to an equivalence relation $\overline{\sim}$ on $\mathbb{Z}/2^{n+k-1}\mathbb{Z}$, in which residues $\overline{a}$ and $\overline{b}$ are equivalent if and only if $\overline{H}_{M_k}(\overline{a}) = \overline{H}_{M_k}(\overline{b})$. We also denote the $\overline{\sim}$-equivalence class of $\overline{a}$ by $[\overline{a}]$, i.e. $[\overline{a}] = \overline{H}_{M_k}^{-1}\left(\left\{\overline{H}_{M_k}(\overline{a})\right\}\right)$. Notice that with this notation we have $[\overline{a}] = \overline{[a]}$.

Throughout this section, we write $x \twoheadrightarrow y$ to indicate that there is an arrow (either red or black) from x to y in the digraph $\mathcal{G}$, $\Gamma_{2^{n+k-1}}$, or Γ_{2^n}.

Lemma 7.4. *There is an arrow $\overline{H}_{M_k}(\overline{x}) \twoheadrightarrow \overline{H}_{M_k}(\overline{y})$ in Γ_{2^n} if and only if there exist $\overline{a} \in [\overline{x}]$ and $\overline{b} \in [\overline{y}]$ for which $\overline{a} \twoheadrightarrow \overline{b}$ in $\Gamma_{2^{n+k-1}}$. In this situation such arrows between the elements of $[\overline{x}]$ and $[\overline{y}]$ form a bijection between $[\overline{x}]$ and $[\overline{y}]$.*

Proof. Suppose $\overline{a} \in [\overline{x}]$ and $\overline{b} \in [\overline{y}]$ such that $\overline{a} \twoheadrightarrow \overline{b}$ in $\Gamma_{2^{n+k-1}}$. Then there exist $a, b \in \mathbb{Z}_2$ in the congruence classes $\overline{a}$ and $\overline{b}$ modulo 2^{n+k-1} respectively, with $T(a) = b$. Then since H_{M_k} is an endomorphism of T,

$$
\begin{aligned}
T\left(H_{M_k}(a)\right) &= H_{M_k}\left(T(a)\right) \\
&= H_{M_k}(b).
\end{aligned}
$$

Thus $H_{M_k}(a) \twoheadrightarrow H_{M_k}(b)$ in $\mathcal{G}$ and thus $\overline{H}_{M_k}(\overline{a}) \twoheadrightarrow \overline{H}_{M_k}(\overline{b})$ in Γ_{2^n}.

Conversely, suppose $\overline{H}_{M_k}(\overline{x}) \twoheadrightarrow \overline{H}_{M_k}(\overline{y})$ in Γ_{2^n}. Then $H_{M_k}(x) \twoheadrightarrow H_{M_k}(y)$ in $\mathcal{G}$. Therefore $T\left(H_{M_k}(x)\right) = H_{M_k}(y)$. Thus $H_{M_k}\left(T(x)\right) = H_{M_k}(y)$ and thus $[y] = [T(x)]$. So taking $a = x$ and $b = T(x)$ we have $a \twoheadrightarrow b$ in $\mathcal{G}$ and consequently $\overline{a} \twoheadrightarrow \overline{b}$ in $\Gamma_{2^{n+k-1}}$ and $\overline{a} \in [\overline{a}] = [\overline{x}]$ and $\overline{b} \in \left[\overline{T(x)}\right] = [\overline{y}]$.

Let s, t be nodes in $\Gamma_{2^{n+k-1}}$ such that $s \twoheadrightarrow t$. Then $s = \overline{a}$ for some a and $t = \overline{T(a)}$. Since T restricts to a bijection between $[a]$ and $[T(a)]$ by the proof of Lemma 23 in [6] it induces a bijection from $\overline{[a]} = [s]$ to $\overline{[T(a)]} = [t]$ in $\Gamma_{2^{n+k-1}}$. $\square$

Lemma 7.5. *Suppose $x_1, \ldots, x_j$ are nodes in Γ_{2^n} such that the subgraph induced by these nodes is a cycle. Then the subgraph induced by $\overline{H}_{M_k}^{-1}(\{x_1, \ldots, x_j\})$ is a union of disjoint cycles in $\Gamma_{2^{n+k-1}}$.*

Proof. Consider a sequence of nodes $x_1, \ldots, x_j$ in Γ_{2^n} that form a cycle in Γ_{2^n}, such that the only arrows between the x_i's are the arrows forming the cycle. Then by Lemma 7.4, the arrows from nodes of $\overline{H}_{M_k}^{-1}(\{x_i\})$ and $\overline{H}_{M_k}^{-1}(\{x_{i+1}\})$ form a bijection between these sets for any $1 \leq i \leq j$ (where we set $x_{j+t} := x_t$ for convenience).

Now, given a node z in one of the sets $\overline{H}_{M_k}^{-1}(\{x_i\})$, there is exactly one arrow $z \twoheadrightarrow z_1$ for some $z_1 \in \overline{H}_{M_k}^{-1}(\{x_{i+1}\})$. Moreover, since the only arrows between the x_i's are the arrows forming the cycle, there are no arrows from z to any other node in $\overline{H}_{M_k}^{-1}(\{x_1, \ldots, x_j\})$. Similarly, there is a unique arrow $z_1 \twoheadrightarrow z_2$ for some $z_2 \in \overline{H}_{M_k}^{-1}(\{x_{i+2}\})$, and there are no other arrows from z_1 into $\overline{H}_{M_k}^{-1}(\{x_1, \ldots, x_j\})$.S We continue this process to define a sequence of nodes $z \twoheadrightarrow z_1 \twoheadrightarrow z_2 \twoheadrightarrow z_3 \twoheadrightarrow \cdots$.

Since there are only a finite number of nodes in $\bigcup_{i=1}^{j} \overline{H}_{M_k}^{-1}(\{x_i\})$, the sequence $z, z_1, z_2, z_3 \ldots$ must be eventually repeating, say with minimum period m. Suppose $z_t \neq z$ is the first entry at which the sequence repeats. Then both $z_{t-1} \twoheadrightarrow z_t$ and $z_{t-1+m} \twoheadrightarrow z_t$, and so $z_{t-1}$and z_{t-1+m} must both lie in $\overline{H}_{M_k}^{-1}(\{x_{i+t-1}\})$. But $z_{t-1} \neq z_{t-1+m}$ by minimality, contradicting Lemma 7.4. It follows that $z_t = z$, and so the sequence $z, z_1, z_2, z_3 \ldots$ is a cycle.

Similarly, choosing a node z' not in this cycle, there is a cycle $z' \twoheadrightarrow z'_1 \twoheadrightarrow z'_2 \twoheadrightarrow z'_3 \twoheadrightarrow \cdots$ that must be disjoint from the previous cycle. Continuing in this manner, we see that $\overline{H}_{M_k}^{-1}(\{x_1, \ldots, x_j\})$ is a union of disjoint cycles in $\Gamma_{2^{n+k-1}}$. □

We now have the tools to prove our main result on folding.

Proof. [Proof of Theorem 7.3] Suppose $a_1, a_2, \ldots, a_l \bmod 2^n$ satisfy the criterion for strong sufficiency of Proposition 6.2 for $d = 2^n$. Let $\Gamma''_{2^{n+k-1}}$ be the graph formed by deleting the preimage of $\{a_1, a_2, \ldots, a_l\}$ under $\overline{H}_{M_k}$ and all edges which are attached to those nodes, followed by the edges which are not part of a cycle in the remaining graph. Since the edges that remain are part of cycles, the cycles containing them map to cycles in Γ''_{2^n}. Since the inverse image of such a cycle is a disjoint union of cycles by Lemma 7.5, $\Gamma''_{2^{n+k-1}}$ consists of a disjoint union of cycles.

Let q be the length of the largest cycle in Γ''_{2^n}, and suppose $kq \leq 630{,}138{,}897$. By Corollary 22 in [6], none of the cycles in $\Gamma''_{2^{n+k-1}}$ has length greater that kq. Thus the maximum cycle in $\Gamma''_{2^{n+k-1}}$ has length less than 630,138,897. Thus $H_{M_k}^{-1}(\{a_1, a_2, \ldots, a_l\})$ is a set of nodes that satisfies the criterion of Proposition 6.2, and is therefore a strongly sufficient set. □

8 Example: $1, 3 \bmod 16$

We conclude with an example that illustrates and links several of the main results in this paper. In Table 4, we see that $1, 3 \bmod 16$ and $2, 12 \bmod 16$ are sets that satisfy the third criterion of Proposition 6.5. Figure 8.1 shows Γ''_{16} for the set $1, 3 \bmod 16$.

Notice that the connected component containing 0 has the property that every red arrow must be followed by at least two black arrows, and the connected component containing 15 has the opposite property: every black arrow must be followed by at least two red arrows in any infinite path. Hence, it does indeed satisfy the third criterion of Proposition 6.5, and so every nontrivial cycle must contain an element congruent to 1 or 3 mod 16, i.e. $1, 3 \bmod 16$ is a cycle sufficient set.

Notice further that the components exhibit the self color duality in Γ_{16}: the connected component of 0 maps to the connected component of 15 under Ω, and in fact one component

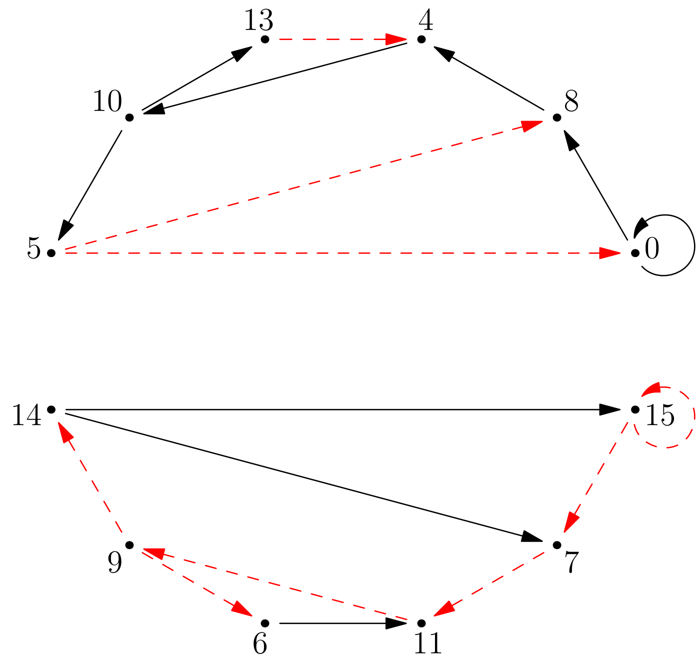

Figure 8.1: The graph Γ''_{16} obtained after removing the nodes 1 and 3 from Γ_{16} and then removing any nodes or edges that are not contained in a cycle.

can be reflected onto the other, with the colors of the arrows reversed, matching each node with its Ω-dual.

Finally, notice that $\Omega(1) \underset{16}{\equiv} 2$ and $\Omega(3) \underset{16}{\equiv} 12$. By the self color duality of Γ_{16}, it follows that removing the nodes 2 and 12 from Γ_{16}, and then removing the nodes and edges not contained in any cycle, results in the same graph Γ''_{16} shown in Figure 8.1. Thus $2, 12 \bmod 16$ is a cycle sufficient set as well.

9 Acknowledgments

The authors would like to thank Gina Monks for her love and support throughout this research project.

References

[1] Applegate, D. and Lagarias, J. C., *Density Bounds for the* $3x+1$ *Problem I. Tree-Search Method*, Math. Comp. 64 (1995), pp. 411-426.

[2] Bernstein, D. J., *A non-iterative* 2*-adic statement of the* $3x+1$ *conjecture*, Proc. Amer. Math. Soc. **121** (1994), 405-408.

[3] Eliahou, S., *The* $3x+1$ *problem: new lower bounds on nontrivial cycle lengths*, Discrete Math. **188** (1993), 45-56.

[4] Hedlund, G., *Endomorphisms and automorphisms of the shift dynamical system*, Math. Systems Theory **3** (1969), 320-375.

[5] Hua, L. K., *Introduction to Number Theory*, Springer-Verlag, 1982, ISBN: 3-540-10818-1.

[6] Kraft, B., Monks, K., *On Conjugacies of the* $3x+1$ *Map Induced by Continuous Endomorphisms of the Shift Dynamical System*, Discrete Math. **310** (2010), 1875-1883.

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

[8] Monks, K. G., Yasinski, J., *The Autoconjugacy of the* $3x+1$ *function*, Discrete Math. **275** (2004), 219-236.

[9] Monks, K. M., *The sufficiency of arithmetic progressions for the* $3x+1$ *conjecture*, Proc. Amer. Math. Soc., 134 (10), October (2006), 2861-2872.

[10] Monks, M., *Endomorphisms of the shift dynamical system, discrete derivatives, and applications*, Discrete Math. **309** (2009), 5196-5205.

[11] Sinisalo, M. K., *On the minimal cycle lengths of the Collatz sequences*, preprint, Univ. of Oulu, Finland, 2003.

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