BITS OF 3^n IN BINARY, WIEFERICH PRIMES AND A CONJECTURE OF ERDŐS

TAYLOR DUPUY AND DAVID WEIRICH

ABSTRACT. Let p and q be distinct primes. We show that on average the base q-expansions of the sequence $\{p^n\}_{n\geq 1}$ have digits which are equidistributed over $a \in \{0,1,\ldots,q-1\}$. A non-averaged version of equidistribution of $(p^n)_q$ as $n \to \infty$ implies a conjecture of Erdős stating that the ternary expansion of 2^n, $(2^n)_3$ omits a 2 for only finitely many n. Our result also implies the non-existence of "higher Weiferich primes".

1. INTRODUCTION

In [Erd79] Erdős conjectured that there are only finitely many powers of 2 whose ternary expansion omits a 2. We will refer to this conjecture as "Erdős' Conjecture".

Progress towards this conjecture has been in the form of upper bounds on the function

$$N(X) = \#\{n \leq X : \ (2^n)_3 \text{ omits a } 2\},$$

which, according to Erdős' conjecture, should approach a constant. We explain the notation: for a prime q and a number N we will let $(N)_q$ denote the base q expansion of N. We view a base q expansion as a string of numbers from the set $\{0,1,\ldots,q-1\}$. [1]. The best known bound on $N(X)$ is due to Narkiewicz [Nar80] who showed

$$N(X) \leq 1.62X^{\alpha_0},$$

where $\alpha_0 = \log_3(2) \approx 0.630$. We refer the reader to [Lag09] for readable proofs and Narkiewicz type bounds for certain dynamical generalizations of this problem. See in particular [Lag09, Theorem 1.4, Proof on page 20 of arxiv version] as well as a refinement of Erdő's conjecture [Lag09, Conjecture E].

For p and q distinct primes the present paper studies the structure of $(p^n)_q$ as $q \to \infty$. Computer experimentation has lead the authors to believe that base q digits of $(p^n)_q$ are equidistributed as $n \to \infty$. We will now formalize this statement: for $a \in \{0,\ldots,q-1\}$ let $d_n(a)$ be the number of a's appearing in $(p^n)_q$.

Conjecture 1. *For all p and q distinct primes and every $a \in \{0,\ldots,q-1\}$,*

$$\lim_{n\to\infty} \frac{d_n(a)}{n\log_q(p)} = \frac{1}{q}. \tag{1.1}$$

Remark 2. The equidistribution statement (1.1) in the case $p = 2$ and $q = 3$ implies Erdős' conjecture. To see this, one argues by contrapositive: Suppose Erdős' conjecture were false. This says 0 is limit point of the sequence $\{d_n(2)\}_{n\geq 0}$. This implies equation (1.1) is false.

[1] for example $(3)_2 = 11$

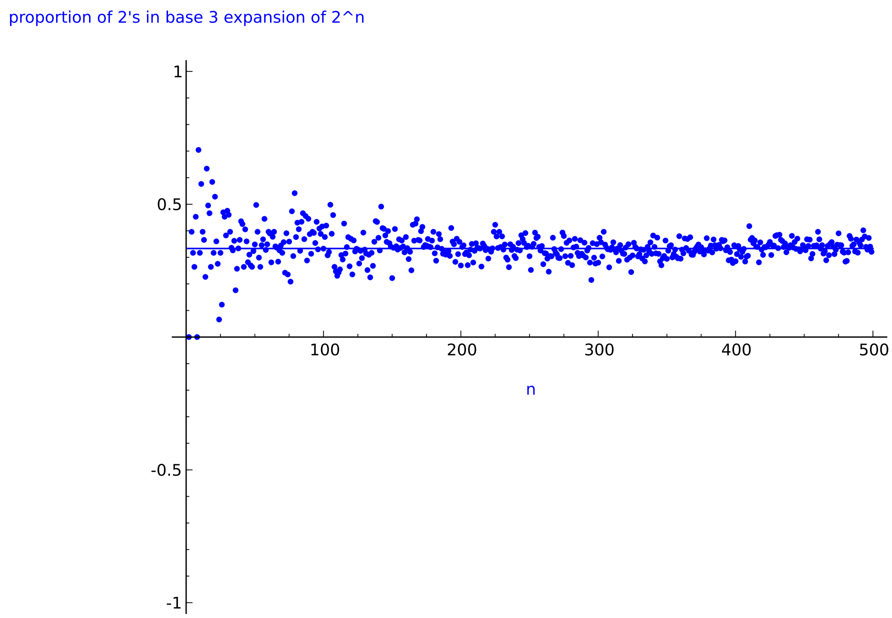

FIGURE 1. The number of 2s in $(2^n)_3$

When $p = 3$ and $q = 2$, Conjecture 1 says that the 0's and 1's appearing $(3^n)_2$ are equidistributed as $n \to \infty$. Table 1 contains the first several members of the sequence $\{(3^n)_2\}_{n\geq 0}$.

n	3^n	$(3^n)_2$	$d_n(0)$	$d_n(1)$
0	1	1	0	1
1	3	11	0	2
2	9	1001	2	2
3	27	11011	1	4
4	81	1010001	4	3

TABLE 1. The first few values of $d_n(a)$ for $p = 3$ and $q = 2$.

For $p = 2$ and $q = 3$ the graph of $\{\frac{d_n(2)}{\log_3(2^n)}\}_{n\geq 1}$ is provided in Figure 1.

The present paper proves an averaged version Conjecture (1.1). Before stating our result we fix some notation. Fix distinct primes p and q, a natural number $m \leq \log_q(p^n)$ and $a \in \{0, \ldots, q-1\}$. Define $d_{n,m}(a)$ and $d'_{n.m}(a)$ to be the number of a's in the first m digits and remaining digits of $(p^n)_q$ respectively[2], so that

$$d_n(a) = d_{n,m}(a) + d'_{n,m}(a).$$

Note that $d_{n,m}(a) = d_n(a)$ when $m = \lceil \log_q(p^n) \rceil$, the number of base q digits in $(p^n)_q$.

[2]The first digit of $(5)_3 = 12$ is 2.

The goal of this paper is to prove the following result:

Theorem 3. *Let p and q be distinct primes and let $a \in \{0, \dots, q-1\}$.*

(1) *For every $m \geq 0$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \frac{d_{n,m}(a)}{m} = \frac{1}{h_m} \sum_{n=1}^{h_m} \frac{d_{n,m}(a)}{m} \tag{1.2}$$

where $h_m = \#H_m$ and $H_m = \langle p \rangle \subset (\mathbb{Z}/q^m)^\times$.

(2) *In view of (1.2), the average proportion of a's in the first m digits is $A_m(a) := \frac{1}{h_m} \sum_{n=1}^{h_m} \frac{d_{n,m}(a)}{m}$. In the limit we have*

$$\lim_{m\to\infty} A_m(a) = 1/q.$$

A trivial consequence of our averaged equidistribution result is the following:

Corollary 4. *The existence of the limit $\lim_{n\to\infty} \frac{d_n(a)}{n \log_q(p)}$ (without necessarily knowing its value) when $p = 2$ and $q = 3$ implies Conjecture 1 and Erdős' conjecture.*

The proof of Theorem 3 goes a theorem about $(\mathbb{Z}/q^m)^\times$ and makes contact with the theory of so called Wieferich primes. We introduce some terminology to explain the main lemma used to prove Theorem 3.

Definition 5. A prime q is called **(classical) Wieferich** if one of the following equivalent conditions holds

(1) $2^{q-1} \equiv 1$ modulo q^2.
(2) The multiplicative order of 2 in $(\mathbb{Z}/q^2)^\times$ is $q-1$.
(3) The multiplicative group generated by 2 modulo q is isomorphic to the multiplicative group generated by 2 modulo q^2.

Such primes were first investigated in [Wie09] in connection to Fermat's Last Theorem. There he proved that if $x^q + y^q = z^q$ is a Fermat triple then q is Wieferich. It is an open problem whether there exists infinitely many Wieferich primes (even assuming the ABC conjecture). The infinitude of non Wieferich primes is implied by the ABC conjecture [Sil88]. The distribution of Wieferich primes is the subject of the Lang-Trotter conjecture. A review of these facts can be found in [Lan90]. We refer the reader to [CDP97] for details on numerical searches for Wieferich primes.

We generalize the notion of a Wieferich prime for the purposes of our discussion.

Definition 6. Let p and q be distinct primes. Let's call a prime q p-**Wieferich** at r if the multiplicative group generated by p modulo q^r is isomorphic to the group generated by p modulo q^{r+1}.

In this notation classical Weiferich primes are simply 2-Wieferich primes at 2. Note that table 1 for example shows that 2 is 3-Wieferich at 3 since the third column of digits is all zeros.

We can now state our main Lemma which we used to prove Theorem 3.

Theorem 7. *Let p and q be distinct primes.*

$$\#\{n : \ q \text{ is } p\text{-Wieferich at } n\} < \infty.$$

This theorem appears in the body as Theorem 14. The proof depends on a modest generalization of the structure theorem for q-adic unit groups $\mathbb{Z}_q^\times$. In particular we show that for m sufficiently large the groups generated by p modulo q^m contain subquotients of the form $(1+q^s\mathbb{Z})/(1+q^m\mathbb{Z}) \cong \mathbb{Z}/q^{m-s}$ with $s<m$ (see Theorem 14)

Acknowledgments. We would like to that C. Moore, R. Lemke Oliver and Carl Pomerance for helpful comments and suggestions.

2. p-Adic units

Let q be a prime. To describe the structure of $(\mathbb{Z}/q^r)^\times$ it is sufficient and conventient to describe the units of $\mathbb{Z}_q = \varprojlim \mathbb{Z}/q^m$, the q-adic integers. This section aims to make standard theorems in elementary number theory explicit for the purpose of later use.

Theorem 8 ([Ser73] Chapter 1)**.** (1) *The units of $\mathbb{Z}_q$ are factor as a direct sum (written multiplicatively here): $\mathbb{Z}_q^\times = T\cdot U$, with*

$$\begin{aligned} T &= \{x\in\mathbb{Z}_q; x^q = x\} \\ U &= 1+q\mathbb{Z}_q. \end{aligned}$$

(2) *We have the following isomorphisms:*
(a) $T\cong (\mathbb{Z}/q)^\times$, *for all q*
(b) $U \cong \begin{cases} (\mathbb{Z}_q,+), & q\neq 2 \\ \mathbb{Z}/2\oplus\mathbb{Z}_q, & q=2 \end{cases}.$

Comments on Theorem 8 part (1) *and part* (2a)*.* The group T is commonly referred to at the **Teichmuller elements** of $\mathbb{Z}_q$. The isomorphism $T\cong \mathbb{Z}/q$ is given by the so-called Teichmuller map described below. In what follows for $x\in\mathbb{Z}_q$ we let $\bar{x}$ denote its residue class in $\mathbb{Z}/q$.

Definition 9. The **Teichmuller map** $\tau : (\mathbb{Z}/q)^\times \to \mathbb{Z}_q^\times$ is defined by

$$\tau(\bar{x}) = \lim_{n\to\infty} x^{q^n}.$$

One can extend this map to all of $\mathbb{Z}_q$ and we note that $\tau(x)$ only depends on the residue class of x modulo q (it is a standard fact that this is well-defined). One notes that reduction modulo q and τ are inverse operations which tells us that the exact sequence

$$1\to 1+q\mathbb{Z}_q \to (\mathbb{Z}_q)^\times \to (\mathbb{Z}/q)^\times \to 1 \tag{2.1}$$

splits. This splitting proves part (1). $\square$

Examining the proof, we observe that the direct sum decomposition $\mathbb{Z}_q^\times = T\cdot U$ in Theorem 8 part (1) can be made effective.

Corollary 10. *The factorization of $\mathbb{Z}_q^\times = T\cdot U$ in Theorem 8 can be made explicit. For $x\in\mathbb{Z}_q^\times$ we have*

$$\begin{aligned} x &= \tau(x)(1+qa(x)), \\ a(x) &= (x/\tau(x)-1)/q. \in 1+q\mathbb{Z}_q \end{aligned}$$

The proof of part (2b) of Theorem 8 when $q\neq 2$ amounts to showing that $(1+q\mathbb{Z}_q)/(1+q^n\mathbb{Z}_q)$ is cyclic. The strategy is to pick some $\alpha\in 1+q\mathbb{Z}_q$ and show $\{\alpha^{q^i}\}_{i=0}^{n-1}$ are distinct modulo q^n. This will follow from the contraction property of the qth power map below (Lemma 13).

Our observation is that one can apply the same trick to smaller balls around the identity of $\mathbf{G}_m(\mathbb{Z}_q)$, i.e. to $\alpha \in 1 + q^r\mathbb{Z}_q$. The goal of the rest of this section is to prove the following strengthening of 2b of Theorem 8.

Theorem 11. *Suppose one of the following*

(1) $q > 2$, $s \geq 1$ *and* $r > s$.
(2) $q = 2$, $s \geq 2$ *and* $r > s$.

Then for all $\alpha \in (1 + q^s\mathbb{Z}_q) \setminus (1 + q^{s+1}\mathbb{Z}_q)$ *we have*

$$\langle \overline{\alpha} \rangle = (1 + q^s\mathbb{Z})/(1 + q^r\mathbb{Z}) \leq (\mathbb{Z}/q^r)^\times.$$

The group generated by $\overline{\alpha}$ *has order* q^{s-r}

Remark 12. The isomorphism in part (2b) of Theorem 8 is the case $s = 1$ of Theorem 11. Explicitly, the isomorphism in part (2b) of Theorem 8 is given by $x \mapsto \alpha^x$

Lemma 13 (Contraction Property of qth Power Map)**.** *Let* $q \geq 2$ *or* $s \geq 2$.

$$\alpha \in (1 + q^s\mathbb{Z}_q) \setminus (1 + q^{s+1}\mathbb{Z}_q) \implies \alpha^q \in (1 + q^{s+1}\mathbb{Z}_q) \setminus (1 + q^{s+2}\mathbb{Z}_q).$$

In this lemma we view $U = 1 + q\mathbb{Z}_q$ as the unit ball around the identity of the multiplicative group $\mathbf{G}_m(\mathbb{Z}_q) = \mathbb{Z}_q^\times$. Observing that we may decompose U into a annuli,

$$U = 1 + q\mathbb{Z}_q = \coprod_{s \geq 1} (1 + q^s\mathbb{Z}_q) \setminus (1 + q^{s+1}\mathbb{Z}_q)$$

the theorem says that the map $x \mapsto x^q$ contracts each annulus in this decomposition to the neighboring annulus one level closer to the identity.

Proof of Lemma 13. For any $a_s \in \mathbb{Z}_q \setminus 0$ one can verify the formulas

$$(1 + a_s q^s)^q = 1 + a_{s+1}q^{s+1}, \tag{2.2}$$

$$a_{s+1} := \frac{(1 + a_s q^s)^q - 1}{q^{s+1}} = \sum_{j=1}^{q} \frac{1}{q}\binom{q}{j} a_s^j q^{s(j-1)} \in \mathbb{Z}_q. \tag{2.3}$$

If $a_s \in \mathbb{Z}_q^\times$ then by examining (2.3) modulo q we see that $a_{s+1} \in \mathbb{Z}_q^\times$ under the hypothesis that $q \geq 2$ or $s \geq 2$.

In the case that $q = 2$ we have $a_{s+1} = a_s(1 + a_s 2^{s-1})$ using formula (2.3) and difference of squares. Here it is necessary to have $s \geq 2$ as $1 + a_s$ may be congruent to 0 modulo 2. □

Proof of Theorem 11. Suppose that $r > s \geq 2$ and q is any prime. Let $a_1 \in \mathbb{Z}_q^\times$ and define $\alpha = 1 + q^s a_1$. For every $t > 0$ we have

$$(\alpha)^{q^t} = (1 + qa_1)^{q^t} \in (1 + q^{s+t}\mathbb{Z}_q) \setminus (1 + q^{s+t+1}\mathbb{Z}_q) \tag{2.4}$$

by the contraction property (Lemma 13). Consider the reduction of (2.4) modulo q^r. Observe that

$$(\alpha)^{q^{r-s}} \equiv 1 \mod q^r$$

and $\alpha, \alpha^q, \ldots, \alpha^{q^{r-s}}$ are distinct.

Since $\#(1 + q^s\mathbb{Z})/(1 + q^r\mathbb{Z}) = \#\mathbb{Z}/q^{r-s} = q^{r-s}$ and $\alpha^{q^{r-s-1}} \neq 1$ modulo q^r, α must have order q^{r-s} as an element of $(\mathbb{Z}/q^r)^\times$ and hence generate all of $(1 + q^s\mathbb{Z})/(1 + q^r\mathbb{Z})$. □

3. Proof of Theorem 7

Using the results we proved in the section on p-adic units (Section 2), we are now able to prove Theorem 7.

Theorem 14 (No higher Weiferich primes)**.** *Let p and q be distinct primes. Let H_r be the cyclic group generated by p modulo q^r. There exists some s (depending on p and q) such that for all $r > s$ sufficiently large, the group H_r containes a subquotient isomorphic to the cyclic group $(1+q^s\mathbb{Z})/(1+q^r\mathbb{Z}) \cong (\mathbb{Z}/q)^{r-s}$.*

In particular if K_r denotes the kernel of the natural quotient map $H_r \to H_{r-1}$ then for all $r > s$ the Kernel K_r is nontrivial (which means q is not p-Weiferich at r).

Proof. We would like the show K_r is nontrivial. Observe the following reduction. Let U_r denote the reduction of U modulo q^r. By the factorization of $\mathbb{Z}_q^\times = T \cdot U$ (Theorem 8) it suffices to show that the kernel of $U_r \cap H_r \to U_{r-1} \cap H_{r-1}$ is nontrivial.

Since p^n is not torsion in $\mathbb{Z}_q$ we have

$$1 \neq \alpha := p/\tau(p) \in (1+q^t\mathbb{Z}_q) \setminus (1+q^{t+1})\mathbb{Z}_q.$$

For some t depending on p and q (c.f. corollary 10).

We claim that some power of α is congruent to 1 modulo q^2.

- **Case** $q \neq 2$**:** Raising α to the power q will achieve this by the contraction lemma (Lemma 13).
- **Case** $q = 2$ **:** If $t > 1$ we are ok. Suppose now that $t = 1$. Write $\alpha = 1 + 2a$. Suppose $n = 0 \mod 4$ and $n > 3$. We will show that $(1+2a)^n \in (1+4\mathbb{Z}_2)$. In this situation
$$\binom{n}{j}(2a)^j = 0 \mod 4$$
for $j \geq 3$. We now have
$$\alpha^n = 1 + \binom{n}{1}2a + \binom{n}{2}(2a)^2 \mod 4.$$
Since
$$\binom{n}{1}2a + \binom{n}{2}(2a)^2 = 2n(a + (n-1)a^2) = 0 \mod 4$$
we can see that $\alpha^n = (1+2a)^n \in 1 + 4\mathbb{Z}_2$. (It suffices to take $n = 4$)

This shows the claim.

We can now suppose there exists some power of α, which we will call β which is a member of $(1+q^s\mathbb{Z}_q) \setminus (1+q^{s+1}\mathbb{Z}_q)$ for some positive s.

We have $\langle\overline{\beta}\rangle = (1+q^s\mathbb{Z})/(1+q^r\mathbb{Z})$ for all $r > s$ by Theorem 11. Hence for all $r > s$, we have

$$\#(1+q^s\mathbb{Z}_q)/(1+q^r\mathbb{Z}_q) = q^{s-r},$$

by Theorem 11 the surjective map

$$(1+q^s\mathbb{Z}_q)/(1+q^r\mathbb{Z}_q) \to (1+q^s\mathbb{Z}_q)/(1+q^{r-1}\mathbb{Z}_q)$$

has nontrivial Kernel of size $\mathbb{Z}/q$. This proves that K_r is nontrivial for every $r > s > 2$. $\square$

4. Proof of Theorem 3

In what follows it will be convenient to think of elemenets in $\mathbb{Z}/q^n$ or $\mathbb{Z}_q = \varprojlim \mathbb{Z}/q^n$ in decimal form. For a sequence of elements $a_i \in \{0, \dots, q-1\}$ we use the notation

$$(a_n \dots a_2 a_1 a_0)_q := a_0 + a_1 q + a_2 q^2 + \cdots + a_n q^n.$$

Again, the digits of $(a_n \dots a_2 a_1 a_0)_q$ are ordered with a_0 being the first digit and a_1 being the last digit.

Lemma 15. *Fix p and q distinct primes. Let H_m be the multiplicative group generated by p in $(\mathbb{Z}/q^m)^\times$.*

1. *For every m the first m digits in the sequence $(p^n)_q$ is periodic in n. The period is the order of the subgroup H_m.*
2. $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \frac{d_{n,m}(a)}{m} = \frac{1}{\# H_m} \sum_{i=1}^{\# H_m} d_{n,m}(a)$

Proof. The first m digits of a number $a \in \mathbb{N}$ can be determined by $a \mod q^m$. Since $a \in (\mathbb{Z}/q^m)^\times$, the group of units, there exists some number h_m such that $p^{h_m} \equiv 1 \mod q^m$.

Part (2) follows from part (1). □

Remark 16. In the statement of Theorem 3, we used the notation

$$A_m(a) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \frac{d_{n,m}(a)}{m}.$$

This will appear again later.

Let p and q be distinct primes and let H_m be the group generated by p in $(\mathbb{Z}/q^m)^\times$. Define

$$K_m = \ker(H_m \to H_{m-1}).$$

Since

$$\ker((\mathbb{Z}/q^m)^\times \to (\mathbb{Z}/q^{m-1})^\times) = \{1 + q^{m-1}a \mod q^m : a \in \mathbb{Z}/q\} \cong \mathbb{Z}/q.$$

We know K_m is either isomorphic to $\mathbb{Z}/q$ or 1. This means that $\# K_m = 1$ or $\# K_m = q$. As sets we have the following description of K_m:

$$K_m = \begin{cases} \{00\cdots01_q, 10\cdots01_q, \dots, (q-1)0\cdots01_q\}, & \# K_m = q \\ \{00\cdots01_q\}, & \# K_m = 1 \end{cases}. \tag{4.1}$$

We will show that one can determine inductively the total number of bits equal to a in the sequence $\{p^n \mod q^m\}$ given the behavior of K_m. The following notation will become useful:

$$\begin{aligned} h_m &:= \# H_m, \\ k_m &:= \# K_m, \\ t_m(a) &:= \sum_{n=1}^{h_m} d_{n,m}(a), \text{ for } a \in \{0, \dots, q-1\}. \end{aligned}$$

Observe that $t_m(a)$ is just the total number of $a's$ appearing in the sequence $\{(p^n \mod q^m)_q\}_{n=1}^{h_m}$. Also observe that we also have the equiality $A_m(a) = t_m(a)/mh_m$. Here $A_m(a)$ was defined in part (2) of Theorem (3) to be the average number of a's in the first m bits of p^n as $n \to \infty$.

The following Lemma says we can determine distribution digits in H_m from the distribution of digits in H_{m-1}.

Lemma 17. *In the case $k_m = 1$ we have*

$$\begin{aligned} t_m(a) &= t_{m-1}(a), \\ h_m &= h_{m-1}. \end{aligned}$$

In the case $k_m = q$ ee have

$$\begin{aligned} t_m(a) &= q t_{m-1}(a) + h_{m-1}, \\ h_m &= q h_{m-1}. \end{aligned}$$

Proof. If $h \in H_{m-1}$ lets define $\widetilde{h} \in H_m$ to be lift of h where we tack a zero on the end. Observe that we have the partition

$$H_m = \bigcup_{h \in H_{m-1}} \widetilde{h} K_m.$$

- If $\#K_m = 1$ then every element of H_m (viewed as a string) is an element of H_{m-1} (as string) just with an extra 0 appended to the end. The equality $t_m(a) = t_{m-1}(a)$ is a trivial consequence of this.
- Suppose $\#K_m = q$. The equality $h_m = q h_{m-1}$ is trivial. We now work on showing $t_m(a) = q t_{m-1}(a) + h_{m-1}$. Let $b_{m-2} \dots b_1 1_q \in H_{m-1}$ and $c0 \dots 01_q = (cq^{m-1}+1)_q \in K_m$. We have

$$\begin{aligned} (0 b_{m-2} \dots b_1 b_0)_q \cdot (c0 \dots 01)_q &= (b_{m-1} b_{m-2} \dots b_1 b_0)_q \\ b_{m-1} &= c \cdot b_0 \mod q \end{aligned}$$

  We know that the b_0's are equidistributed over $\{1, \dots, q-1\}$ in H_{m-1} and that $c \in \{0, \dots, q-1\}$ uniquely determines the element of K_m.

  For each element of $(b_{m-2} \dots b_1 b_0)_q \in H_{m-2}$ we get a whole set of elements

$$\{(cb_0) b_{m-2} \dots b_1 b_0 : c \in \mathbb{Z}/p\}.$$

  If $c = 0$ then $cb_0 = 0$, and this only happens once. Since $(\mathbb{Z}/p)^\times$ is a group

$$\{a : a \in (\mathbb{Z}/p)^\times\} = \{a \cdot b_0 : a \in (\mathbb{Z}/p)^\times\}$$

  this implies that

$$\pi_{m,m-1}^{-1}(b_{m-2} \dots b_1 b_0) = \{b_{m-1} b_{m-2} \dots b_0 : b_{m-1} \in \mathbb{Z}/p\}.$$

  The result follows from the equality

$$H_m = \bigcup_{h \in H_{m-1}} \pi_{m,m-1}^{-1}(h).$$

  (Alternatively, one can argue from periodicity).

□

We now derive some formulas for $A_m(a)$. The main idea of this proof is that $k_m = q$ pulls digits of p^n toward equidistribution and $k_m = 1$ pulls the distribution of the bits of p^n toward having more zeros. In the situation where $k_m = q$ the "new bit" is completely equidistibuted. Note in particular that for all m we have $0 \leq A_m(a) \leq 1$ from which it is easy to see that if $k_m = q$ "pushes" $A(a,m)$ towards equidistibution $1/q$.

Lemma 18. (1) *For $m > 2$ we have*

$$A_m(a) = \left(1 - \frac{1}{m}\right) A_{m-1}(a) + \frac{1}{q(q-1)m}(k_m - 1)$$

(2) *Define $\bar{k}_m = k_m - 1$ for $m \geq 2$ and define $\bar{k}_1 = q$. For all $a \in \{1, \ldots, q-1\}$ we have*

$$A_m(a) = \frac{1}{q(q-1)} \frac{\bar{k}_1 + \bar{k}_2 + \cdots + \bar{k}_m}{m}. \tag{4.2}$$

Proof. We analyze the formula by cases:

$k_m = 1$: (the density of a's in H_m is strictly decreasing). We have $H_m \cong H_{m-1}$ and that elements of H_{m-1} give an element of H_m by just tacking a zero at the end. We have

$$\begin{aligned} t_m(a) &= t_{m-1}(a) \\ h_m &= h_{m-1} \end{aligned}$$

which implies

$$A_m(a) = \frac{t_m(a)}{mh_m} = \frac{m-1}{m} \cdot \frac{t_{m-1}(a)}{(m-1)h_{m-1}} = \left(1 - \frac{1}{m}\right) A_{m-1}(a).$$

$k_m = q$: (density of a's will approach the equilibrium) We have

$$\begin{aligned} t_m &= qt_{m-1} + h_{m-1} \\ h_m &= qh_{m-1} \end{aligned}$$

which gives

$$A_m(a) = \frac{t_m}{mh_m} = \frac{qt_{m-1} + h_{m-1}}{mh_m} = \left(1 - \frac{1}{m}\right) A_{m-1}(a) + \frac{1}{qm}.$$

We now solve the recurrence relation to give the formula in part 2. This proof is by induction. Fix some $a \in \{1, \ldots, q-1\}$. Note that $A_1(a) = 1/(q-1)$ since p generates the unit group mod q which has $(q-1)$ elements, so the base case is trivial. We now do the inductive step and suppose the formula holds for m and prove it for $m+1$.

$$\begin{aligned} A_{m+1}(a) &= \frac{n}{m+1} A_m(a) + \frac{\bar{k}_{m+1}}{q(q-1)(m+1)} \\ &= \frac{1}{q(q-1)(m+1)} \left[\bar{k}_1 + \bar{k}_2 + \cdots + \bar{k}_m\right] + \frac{\bar{k}_{m+1}}{q(q-1)(m+1)} \\ &= \frac{1}{q(q-1)} \frac{\bar{k}_1 + \bar{k}_2 + \cdots + \bar{k}_{m+1}}{m+1}, \end{aligned}$$

which proves our result. □

Remark 19. Note that (1) shows that $A_m(a) = \frac{1}{h_m} \sum_{n=1}^{h_m} \frac{d_{n,m}(a)}{m}$ only depends on whether a is zero or nonzero. This follows from $A_1(a) = 1/(q-1)$ for all $a \in \{1, \ldots, q-1\}$ as p generates $(\mathbb{Z}/q)^\times$ together with the recurrence.

Supposing $A_m(a)$ was completely independent of a we would have $qA_m(a) = \sum_{a=0}^{q-1} A_m(a) = 1$ which implies $A(m) = 1/q$. This would give an easy proof of our result.

We have now related the distribution of bits to the condition about "Weiferich primes".

Lemma 20. *With the notation as above we have*

$$\lim_{m\to\infty} A_m(a) = 1/q \iff 1 = \lim_{m\to\infty} \frac{1}{m(q-1)} \sum_{j=1}^{m} \bar{k}_j.$$

Proof. Follows directly from Lemma 18 part (2) and the definition of $A_m(a)$. □

We now prove that $\lim_{n\to\infty} \frac{1}{n(q-1)} \sum_{j=1}^{n} \bar{k}_j = 1$. To do this we need to study the multiplicative group generated by p modulo q^r.

Theorem 21. *With the notation as above and $\bar{k}_j = \#K_j - 1$ we have*

$$\lim_{n\to\infty} \frac{1}{n(q-1)} \sum_{j=1}^{n} \bar{k}_j = 1.$$

In particular this implies

$$\lim_{m\to\infty} A_m(a) = 1/q$$

for all $a \in \{0, \dots, q-1\}$.

Proof. Since $K_j \leq \ker(\mathbb{Z}/q^j \to \mathbb{Z}/q^{j-1}) \cong \mathbb{Z}/q$ it can only have order q or 1. By 14, K_j is nontrivial for all but a finite number of j and hence $\bar{k}_j$ must be equal to q for all but a finite number of j.

The second part follows from Lemma 20. □

5. Discussion

Let p and q be distinct primes and consider powers of p in base q as usual.

Remark 22. Let $f : \mathbb{N} \to \mathbb{N}$. Let q be a prime. We will introduce temporary notation for this remark: for $a \in \{0, \dots, q-1\}$, take $d_n(a)$ to be the number of a's appearing in $(f(n))_q$ (this paper is mostly concerned with $f(n) = p^n$). For "generic" exponential diophantine functions one expects (from experiments) that

$$\lim_{n\to\infty} \frac{d_n(a)}{\log_q(f(n))} = 1/q. \tag{5.1}$$

For example, when p, l and q are distinct primes one expects (5.1) for $f(n) = p^n + l^n$. Another example is when $f(n) = p^n + g(n)$ where $g(n)$ is a polynomial. More generally, if $f(n) = p^n + g(n)$ where $\log_q |g(n)| = o(\log_q(p^n))$ as $n \to \infty$, the truth of (5.1) with $f(n) = p^n$ implies the truth of equation (5.1) for $f(n) = p^n + g(n)$. This is because $g(n)$ will affect only a density zero proportion of the digits in the limit $n \to \infty$.

It is unclear how to characterize the subset of exponential diophantine functions should satisfy (5.1) even conjecturally. Figure 22 provides a graph of a sequence $\{d_n(a)\}$ when $f(n) \neq p^n$.

Remark 23. Our result in the case that $p = 2$ and $q = 3$ together with bounds of the form $N(X) \leq \beta X^\alpha$ for positive constants β and α do not appear strong enough to prove Erdős' conjecture.

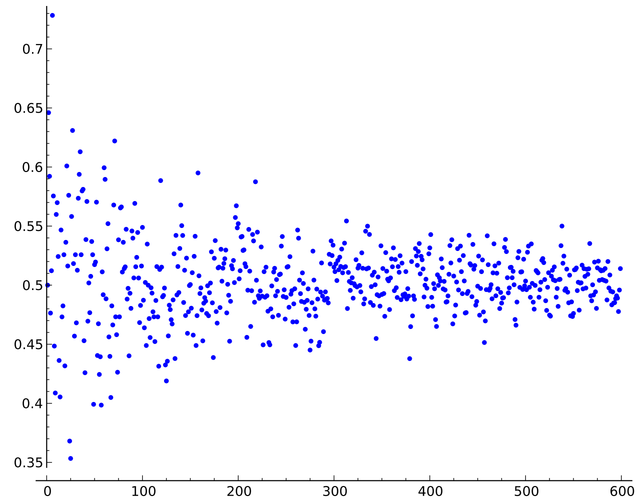

FIGURE 2. FIX ME

Remark 24. The exact sequence

$$1 \to (1+q\mathbb{Z})/q^m \to (\mathbb{Z}/q^m)^\times \to (\mathbb{Z}/q^{m-1})^\times \to 1$$

can be replaced by the sequence

$$1 \to (1+q^r\mathbb{Z})/(q^{r+s}) \to (\mathbb{Z}/q^{r+s})^\times \to (\mathbb{Z}/q^s)^\times \to 1$$

whenever $r \leq s$ (a nice choice is $r = s = 2^l$). This allows us to understand half of the digits of $(p^n)_{q,r+s}$ rather than a single bit as coming from an abelian group since $(1+q^r\mathbb{Z})/(1+q^{r+s}\mathbb{Z}) \cong \mathbb{Z}/q^s$ as abelian groups.

Let h_m be the order of p in $(\mathbb{Z}/q^m)^\times$. We know

$$h_m = Cq^m$$

for some constant C and sufficiently large m. This means that the largest number in $\langle p \rangle \subset (\mathbb{Z}/q^m)^\times$ is roughly p^{q^m} and $(p^{q_m})_q$ has length roughly q^m. This means we may hope to undersatnd $(p^n)_{q,\log_q(n)}$ using group theoretic methods (which is still a zero density proportion of the total digits in $(p^n)_q$). The authors do not know currently if

$$\lim_{n\to\infty} \frac{d_{n,\log_q(n)}(a)}{\log_q(n)} = \frac{1}{q} \tag{5.2}$$

or if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \frac{d_{n,\log_q(n)}(a)}{\log_q(a)} = \frac{1}{q}. \tag{5.3}$$

Note that Equation 5.2 implies Erdős' conjecture.

REFERENCES

[CDP97] Richard Crandall, Karl Dilcher, and Carl Pomerance. A search for wieferich and wilson primes. *Mathematics of Computation of the American Mathematical Society*, 66(217):433–449, 1997.

[Erd79] P Erdős. Some unconventional problems in number theory. *Acta Mathematica Hungarica*, 33(1):71–80, 1979.

[Lag09] Jeffrey C Lagarias. Ternary expansions of powers of 2. *Journal of the London Mathematical Society*, 79(3):562–588, 2009.

[Lan90] Serge Lang. Old and new conjectured diophantine inequalities. *Bulletin of the American Mathematical Society*, 23(1):37–75, 1990.

[Nar80] W Narkiewicz. A note on a paper of h. gupta concerning powers of 2 and 3, univ. *Beograd. Publ. Elecktrotech. Fak. Ser. Mat. Fiz*, (678-715):173–174, 1980.

[Ser73] Jean-Pierre Serre. *A course in arithmetic*, volume 97. Springer-Verlag New York, 1973.

[Sil88] Joseph H Silverman. Wieferich's criterion and the¡ i¿ abc¡/i¿-conjecture. *Journal of Number Theory*, 30(2):226–237, 1988.

[Wie09] ArthurP Wieferich. Zum letzten fermatschen theorem. *Journal für die reine und angewandte Mathematik*, 136:293–302, 1909.